FEDERAL RESERVE BANK *of* ATLANTA WORKING PAPER SERIES

Long-Run Intergenerational Effects of Social Security

Daniel K. Fetter, Lee M. Lockwood, and Paul Mohnen

Working Paper 2026-10
July 2026

Abstract: Both historically and today, support of aging parents has largely taken the form of in-kind transfers that require physical proximity, such as housing and caregiving. If Social Security substitutes for such support, it can relax constraints on where recipients' children live and work. We investigate the long-run intergenerational effects of the early Social Security program, exploiting within-occupation, cross-industry differences in coverage, and a new dataset linking parents to their children's later-life outcomes. We find that sons whose parents had greater predicted coverage moved farther from their childhood homes, earned more, and lived in better neighborhoods late in life. We find no such effects for daughters, who tended to provide forms of support less easily replaced by Social Security. The gains considerably exceeded the associated Social Security benefits for the average family, with migration to better-matched labor markets a likely key driver. We propose that the early program enabled families to realize gains from migration that were back-loaded, uncertain, and difficult to contract on.

JEL classification: H55, N32, J61, R23, J14, D15

Key words: Social Security, social insurance, old-age support, intergenerational transfers, coresidence, geographic mobility, internal migration, intergenerational mobility, linked census data, New Deal, public pensions

https://doi.org/10.29338/wp2026-10

Daniel K. Fetter is with Dartmouth College and NBER. Lee M. Lockwood is with the University of Virginia and NBER. Paul Mohnen is with the Federal Reserve Bank of Atlanta. For helpful comments and discussions, the authors thank Ran Abramitzky, Amy Finkelstein, Sasha Indarte, Olivia Mitchell, Matt Notowidigdo, Jim Poterba, Shanthi Ramnath, Arthur Seibold, Kent Smetters, Jessica Van Parys, Heidi Williams, as well as participants at the AEA, Atlanta Fed, Berkeley, Boston College Center for Retirement Research, Boston Fed, Chicago Fed, Clemson, Dartmouth, Emory, Harvard, GMU, Hitotsubashi, IPUMS, KU Leuven, Lehigh, LMU-Munich, Michigan RDRC, Minneapolis Fed, NBER (Aging, DAE, ESS, Economics of Mobility, TAPES), Philadelphia Fed, Princeton, UC Davis, UCLA, UIUC, UNL, Rutgers, Stanford, SEA, SOLE, Toronto, Trinity College Dublin, Wesleyan, Wharton, and Yale. They thank David Lee, Tamri Matiashvili, Elisa Heinrich Mora, Maxwell Rong, Amanda Wahlers, and Qiyi Zhao for excellent research assistance. Fetter and Lockwood gratefully acknowledge financial support from National Science Foundation (NSF) Grant SES-1628860. Fetter gratefully acknowledges financial support from the Stanford Institute for Economic Policy Research. Lockwood gratefully acknowledges financial support from the Bankard Fund for Political Economy. Mohnen gratefully acknowledges financial support from the Alfred P. Sloan Foundation Pre-Doctoral Fellowship on the Economics of an Aging Workforce, awarded through the NBER. The LIFE-M project was generously supported by the National Science Foundation (SMA 1539228), the National Institute on Aging (R21 AG05691201), the University of Michigan Population Studies Center Small Grants (R24 HD041028), the Michigan Center for the Demography of Aging (MiCDA, P30 AG012846-21), the University of Michigan Associate Professor Fund, the Michigan Institute on Research and Teaching in Economics (MITRE), the Russell Sage Foundation (1911-19560), and the Washington Center for Equitable Growth. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

1 Introduction

The introduction and expansion of Social Security in the United States was associated with profound changes in the economic relationships between aging parents and their adult children. The decline in co-residence shown in Figure 1 is one notable example, which past work (e.g., Costa, 1999; McGarry and Schoeni, 2000) has linked to the rise of government old-age support. It has long been recognized that reallocation of support from families to the state may affect not only recipients but also their children. Yet while key forms of family old-age support—including shared housing, help with household tasks, and caregiving—require parents and children to live together or nearby, seminal theoretical analyses of these intergenerational effects (e.g., Barro, 1974; Becker, 1974) largely abstract from the *spatial* dimension of family support.

This spatial dimension points to a channel through which Social Security may affect recipients' children: To the extent Social Security displaces support that requires parents and children to live in close proximity, it may relax a constraint on children's location choices—freeing them to live and work in a broader set of places, with long-run implications for their careers and other outcomes. We posit, however, that common types of support differed in their substitutability with cash benefits—shared housing being more substitutable and hands-on caregiving less so, for example. An implication is that this channel is likely weaker for daughters than for sons, as daughters have historically provided more of the caregiving and other in-person assistance that we argue was more difficult to purchase on the market (e.g., Horowitz, 1985; Grigoryeva, 2017).

Motivated by these observations, we provide new evidence on the long-run effects of Social Security on recipients' children. We combine two key ingredients. First, we introduce a novel source of variation in Social Security eligibility in the early years of the program. This approach leverages the fact that employment in some industries was covered when the program was introduced in 1935, while similar employment in other industries was only covered in 1950 or later. For example, work as a janitor in a private, for-profit firm was covered in 1935, while work as a janitor in a nonprofit hospital or school was not covered until 1950. As a result, individuals with similar jobs sometimes differed, solely because of their industry of employment, in the likelihood of ever becoming eligible for Social Security or in the age at which they became eligible for benefits. Moreover, differences in coverage year had differential "bite" across cohorts since eligibility to receive retirement benefits was conditional on being 65 or older; being in 1935- versus 1950-covered employment mattered much less for cohorts that turned 65 in the early 1950s than for cohorts that turned 65 ten years earlier, for example. This allows us to compare the long-run outcomes of otherwise-similar children whose parents differed in the likelihood and timing of Social Security eligibility.

Second, we build a new dataset linking Social Security death records to the 1930 and 1950

US Censuses. The 1930 Census allows us to predict the age at which fathers would become eligible for Social Security based on their year of birth and employment information prior to the passage of the Social Security Act. We show that this predicted eligibility strongly predicts fathers' actual claiming behavior, and is uncorrelated with households' pre-program characteristics in 1930 and with selection into the linked sample. Social Security death records contain information on children's last known ZIP code of residence, which we show to be correlated with lifetime earnings and use as a proxy measure for their late-life socioeconomic status. The 1950 Census allows us to study children's migration patterns by observing their location 10 years after the first monthly Social Security benefits were issued, when the children in our sample were in their early- to mid-adulthood. We also provide complementary evidence from the 1973 Survey of Occupational Changes in a Generation (OCG), which contains direct measures of children's income and occupational standing in adulthood and, crucially, sufficient retrospective information to determine their fathers' likely Social Security eligibility age.

Our main finding is that sons whose parents were more likely to ever be eligible for Social Security benefits, or were likely eligible at a younger age, lived in better neighborhoods near the end of their lives, as measured by ZIP-level income or home values. In contrast, we find no evidence of such effects for daughters, consistent with the prediction that they might be less affected due to differences in the types of support they tended to provide to their parents. Using the 1973 OCG, we also show that these sons had better-paying jobs and higher earnings in adulthood, consistent with the interpretation that residence in higher-income ZIP codes late in life reflects higher late-career earnings for sons.

The results imply large, long-lasting financial gains. Under our baseline assumptions, a 10-year difference in fathers' predicted earliest eligibility age—roughly the difference between 1935- and 1950-covered fathers from the 1877 birth cohort—led to a full-career average earnings gain of roughly 12 percent and an expected present value gain (in 2020 dollars) of about \$130,000 per son. Aggregating across all children of the average recipient family, we estimate that the total gain was about 6 times the associated Social Security benefits (95% CI: [2.2, 10.6]). This exceeded recipients' earnings losses, implying an increase in total family earnings. As a result, our point estimates suggest that these early Social Security expansions paid for themselves (net cost per dollar of benefits of $-$0.55), though given the statistical and modeling uncertainty, we cannot rule out that they had a positive net cost (95% CI: [$-$1.77, \$0.67]).

A range of evidence points to migration to better-matched labor markets as the main driver of these gains. Social Security led sons, but not daughters, to move farther from their parents by 1950, consistent with the displacement of location-dependent family support. These moves were closely tied to long-run gains: Exposed sons were less likely to remain in their parents' 1930 census division and end up in the bottom two quintiles of the late-life ZIP income distribution, and more likely to move across census divisions and end up in the top two quintiles. The direction and timing of migration offer further supporting evidence. Sons moved to states where workers with their education level earned more, and they were less likely to move long distances after 1950, suggesting that earlier moves helped them achieve better labor market matches. These patterns are difficult to explain through increased financial transfers from parents alone.

Our paper makes several contributions. A large literature in economics documents that Social Security and other old-age support programs substituted for family support, including co-residence, caregiving, and private transfers (Costa, 1997, 1998, 1999; McGarry and Schoeni, 2000; Engelhardt et al., 2005; Mukherjee, 2022; though cf. Ruggles, 2007). Much less is known, however, about whether and how this substitution affected the lives of recipients' children. Children's long-run outcomes are a necessary input to a comprehensive welfare evaluation of these programs, and our novel linkages following children of early recipients to the end of their lives enable us to trace out these effects. Our results suggest that the proximity requirement of much family support has first-order consequences for children's location and lifetime outcomes.

We also speak to the canonical analysis of pay-as-you-go programs in overlapping generations economies (e.g., Samuelson, 1958; Diamond, 1965; Feldstein, 1974; Auerbach and Kotlikoff, 1987). A foundational result is that expansions of such programs confer a windfall gain on the initial old, often at the considerable expense of subsequent generations, in large part by crowding out private saving and reducing the capital stock; in the case of Social Security, the large statutory redistribution to early recipients is well documented (Moffitt, 1984; Leimer, 1994; Geanakoplos et al., 1998; Feldstein and Liebman, 2002). The literature has long recognized that government old-age support may crowd out family transfers: The dynastic models of Barro (1974) and Becker (1974) treat such links as a central margin, and Becker's broader framework allows nonpecuniary interactions within the family. Yet even these seminal analyses abstract from the spatial dimension of family support. Just as Social Security is thought to crowd out private saving, our evidence suggests that, at least for these early cohorts, it also displaced in-kind support that recipients would otherwise have received from their adult children. But the two forms of crowding out have opposite implications for later generations: Whereas a lower capital stock is presumed to impose long-run losses, displacing proximity-based family support relaxed constraints on where recipients' children could live and work, generating large, long-lasting gains.

Our findings also have implications for the efficiency and welfare effects of Social Security. A key efficiency cost of Social Security is thought to be its negative impact on labor supply by inducing recipients to retire earlier (see Coile, 2015, for a review). But in our setting, Social Security did not simply increase old-age support while reducing recipients' work. It changed the form of old-age support, thereby relaxing location constraints on recipients' children. As noted above, once one accounts for the effects on recipients' children, these expansions increased total family earnings—enough that they may have paid for themselves. So, at least for these early expansions, what is widely thought to be a major efficiency cost was outweighed by a previously underemphasized efficiency benefit: By relaxing the location constraints of family support, Social Security improved the matching of recipients' children to labor markets.

By introducing a new research design to estimate the causal impact of eligibility for Social Security among early recipients, our paper adds to the quasi-experimental literature on Social Security. A fundamental challenge in studying Social Security is its near-universal nature and the fact that the timing of receipt and benefit amounts are functions of endogenous decisions that workers make. Previous studies have exploited features of the program and policy changes that generate discontinuities to isolate the causal impact of Social Security on various outcomes of interest, such as labor supply, saving, and mortality. These include the Social Security "Notch" (e.g., Krueger and Pischke, 1992; Engelhardt et al., 2005; Gelber et al., 2016; Mukherjee, 2022), the early eligibility age of 62 (Fitzpatrick and Moore, 2018), the eligibility age of 60 for survivor benefits (Coyne et al., 2024), the introduction of the early eligibility age (Engelhardt et al., 2022), and staggered increases in the full retirement age (Mastrobuoni, 2009). Our research design, which exploits how Social Security was rolled out in the US, makes use of significantly larger differences in treatment of otherwise similar individuals. Moreover, studying an era in which government old-age support was minimal and family support was extensive sheds new light on how government programs displace family support.

A related literature examines how the introduction of pension programs in developing countries affects the next generation (e.g., Jensen, 2004; Edmonds et al., 2005; Ardington et al., 2009; Huang and Zhang, 2021; Bau, 2021). Separately, a large literature studies the gains from and barriers to internal migration in developing countries, particularly migration from rural to urban areas (e.g., Munshi and Rosenzweig, 2016; Lagakos et al., 2023). Recent work at the intersection of these literatures shows that proximity-based support to elderly parents can deter even high-return migration, and that public pensions or institutions that facilitate upfront transfers to parents can relax this constraint (Gai et al., 2025; Bau et al., 2026). Our findings raise a natural question: If the gains from migration were so large, why did families not realize them on their own, even without Social Security? We posit that when migration would displace co-resident or in-person support, families often could not capture the gains: The returns to migration were back-loaded, uncertain, and hard to insure, while the family support a move would displace was subject to indivisibilities and missing markets that concentrated its cost on individual children. By documenting this channel in the historical US and tracing children's outcomes over the long run, we show that family-based location constraints can remain economically important even in a developed economy and that relaxing them can generate large, persistent gains for the next generation.

Overall, our paper provides a new perspective on the introduction and early expansion of Social

Security. Our findings suggest that it was not so much a large net transfer to the initial recipients at the expense of future generations as a change in the form of old-age support—away from location-constrained family support—that enabled significant long-run benefits for recipients' children by facilitating migration to better-matched labor markets. This accords with work by historians such as Gratton (1996), who points to the significant costs of a family-based old-age support system and notes that in public opinion surveys, Social Security was as popular with the young as with the old. Several features of our setting—minimal pre-existing government support, considerable family support that was substitutable with cash, and high returns to migration—were especially conducive to large intergenerational effects, so the magnitudes we estimate may not generalize to other settings, including marginal changes to the program today. The underlying mechanism, however—that government old-age support can relax location constraints for recipients' children—plausibly operates wherever family old-age support is location-specific and government programs can substitute for it. As such, our findings may hold important lessons for contemporary old-age support policies such as long-term care programs, given that informal caregiving to the elderly by family members remains prevalent in the US today (Barczyk and Kredler, 2018; Ko, 2022; Mommaerts, 2025).

2 Background

2.1 Eligibility and benefits in the early Social Security Program

The Social Security Act was passed in 1935, but it was under the 1939 Amendments to the Social Security Act that the first monthly retirement and survivors' benefits under Old Age and Survivors Insurance (henceforth, "Social Security") were paid, in 1940. Under the 1939 and subsequent amendments, a worker became eligible to receive monthly retirement benefits by working for a sufficient number of quarters in employment covered by Social Security, provided he or she earned above a threshold earnings level in that quarter. For example, under the 1939 Amendments, a worker would be eligible for retirement benefits if he or she earned a sufficient amount in covered employment in half of the quarters after 1936 (or the quarter he or she turned 21) and prior to death or age 65, subject to a minimum of six quarters.

The requirement for a sufficient length of employment meant that relatively few of the birth cohorts who were already at advanced ages in 1940 were eligible for monthly Social Security payments. The major source of expansion of the program in subsequent decades was the aging-in of cohorts who had worked a sufficient length of time to be eligible for benefits. The growth of the program as it matured is shown in Appendix Figure A.1, which shows the share of men receiving retirement benefits over time, separately by 5-year birth cohorts. Only about 5 percent of the 1871–75 birth cohort received monthly payments in 1940 at ages 65–69, rising to about 42 percent of those surviving to 1955 (ages 80–84). The 1896–1900 birth cohorts, born 25 years later, received Social Security at far higher rates: nearly 70 percent of men collected benefits in 1965, at ages 65–69, and 96 percent of those surviving to 1980 collected them then, at ages 80–84.

It will be central to our empirical analysis that eligibility also grew as the definition of covered employment was broadened. The original act and the 1939 Amendments covered wage and salary workers only, excluding the self-employed. Wage and salary workers in certain sectors were also excluded from coverage: in particular, domestic workers in private homes, agricultural wage workers, employees of nonprofit organizations, and local, state, and federal government employees.[1] For types of employment covered in the original Act, coverage began on January 1, 1937.[2]

Amendments to the Social Security Act, particularly between 1950 and 1956, significantly expanded the definition of covered employment. Our analysis focuses on wage and salary workers rather than the self-employed, so focus on those extensions.[3] The first extension of coverage to additional categories of wage workers was under the 1950 Amendments. Under these amendments, coverage was made compulsory for regularly employed agricultural wage workers and domestic workers. The amendments also extended coverage on an elective basis to most employees of nonprofit organizations and to employees of state and local governments who were not under an existing retirement system. Importantly, coverage was not retroactive: the new additions to covered employment applied only to work performed in 1951 and later. Hence, a worker working solely in 1950-covered employment would not begin to accumulate quarters of coverage until 1951. Amendments in 1954 extended coverage to state and local employees already under a retirement system, still on an elective basis, and amendments in 1956 further liberalized the provisions under which state and local government employees could be covered. Although coverage was on an elective basis for nonprofits, take-up was fairly rapid: as of 1951, 56% of employees of nonprofit organizations were covered under Social Security, rising to 76% by 1960 (Subcommittee on Social Security, 1976, p. 27). For state and local governments, take-up was slightly slower than for nonprofits, but still rose to high levels: about 11 percent of all state and local government

[1] See Appendix B.1 for a more detailed description of the initial exclusions and subsequent expansion. Nonprofit organizations were excluded due to concerns that it would jeopardize their tax-exempt status (McCamman, 1948), and state and local government employees were excluded because of perceived constitutional limitations on federal taxation of state and local governments (Clark et al., 2011). Other exclusions included railroad workers, who were covered under a separate system under the Railroad Retirement Act, and some other groups of smaller size, such as crew members of ships and fishermen.

[2] Work done at age 65 or older was not covered under the original act, but was covered under the 1939 Amendments, with an effective date of January 1, 1939. Hence, work during at least some part of 1937 and 1938 did not contribute to eligibility for workers in the 1873 and earlier birth cohorts.

[3] Coverage of self-employment was also important to the growth of Social Security, although it does not play a role in our empirical analysis. The 1950 Amendments extended coverage on a compulsory basis to the self-employed, except for certain professional groups and farmers. Self-employed farmers and farm managers were largely brought into the system in 1954, with a further expansion in 1956. Self-employed professionals who had not been covered in 1950 were mostly covered in either 1954 or 1956.

employees were covered by Social Security at the beginning of 1951 (the first year in which they could be), and by 1958 the majority of state and local government employees (54 percent) were covered by Social Security (Kestenbaum, 1982).

Under the 1939 Amendments, a worker who had worked a sufficient number of quarters in covered employment was eligible to receive monthly retirement benefits after reaching age 65, subject to an earnings test that withheld benefits in months in which earnings exceeded $15 (in 1939, this amount corresponded to about $280 in 2020 dollars). In addition, conditional on being 65 or older, the spouse of a retired male worker received spouse's benefits (50 percent of the primary benefit) and the widow of a male worker received widow's benefits (75 percent of the primary benefit). The primary monthly amount was a function of the beneficiary's earnings in covered employment. Replacement rates for the median earner were around 20 percent through the 1940s, and 25 to 30 percent through the 1950s and most of the 1960s; for a low earner replacement rates were somewhat higher, around 25 to 30 percent in the 1940s and 40 to 50 percent in the 1950s and 1960s (SSA Office of the Chief Actuary, 2013, Supplementary Table V.C7). Tax rates for both the employer and employee were 1 percent beginning in 1937; these rose to 1.5 percent in 1950 and to 2 percent in 1954, and gradually rose in later years.

As has been frequently noted (e.g., Moffitt, 1984; Schieber and Shoven, 1999; Gruber, 2013), the initial generation of beneficiaries paid little in taxes, received significant benefits over their lifetimes, and hence received large net transfers from Social Security. Consider a married couple with a husband born in 1880 and a wife born in 1883. Suppose that the wife did not work on the market, but the husband worked continuously until he was 65 in employment covered under the 1935 act, and in each year earned the average for all workers in covered employment. Also suppose that both husband and wife lived exactly their remaining life expectancy at age 65. The net present value at retirement of future benefits the couple received, minus taxes paid, is on the order of $100,000 in 2020 dollars. By comparison, the median home value in 1940 was about $55,000.[4]

2.2 Transfers and shared resources between the elderly and their adult children

While our empirical results are independent of the underlying mechanisms, our analysis is motivated and informed by the historical economic relationships between parents and their adult children. We emphasize two points in particular. First, many of the common forms of support required spatial proximity. Second, daughters and sons differed in the type of support they provided, and we posit that the forms of support daughters provided relatively more intensively (such as caregiving) were less readily substitutable with Social Security payments than the forms sons provided relatively more intensively (such as shared housing).

[4]See census.gov/programs-surveys/decennial/tables/time-series/coh-values/values-unadj.txt (accessed 3/18/2021).

Co-residence was one of the major forms of support (or shared resources) between adult children and their parents in the first half of the 20th century. We have seen in Figure 1 that in 1930, roughly 27 percent of women 65 and older, and about 12 percent of men 65 and older, lived with their children and were not household heads; co-residence rates of the elderly fell significantly in the subsequent decades, coinciding with the introduction and expansion of government old-age support over the 20th century.[5] Because we are interested in effects of Social Security on adult children, it is also useful to consider the share of adult children co-residing with parents, and hence what an adult's expectations might have been regarding their future living arrangements. Panel (a) of Figure 2 shows the share of native-born adults who were either household head or the spouse of the household head and who had one or more parents or parents-in-law in the household, tracing individual 10-year birth cohorts over their life cycles. For each birth cohort, co-residence rises between ages 25–34 and 35–44 and then declines, presumably due to parents' mortality. For all cohorts shown here, this form of co-residence peaked at ages 35 to 44. For cohorts that reached their mid-20s by 1940, about 7 percent had a parent or parent-in-law in their household at ages 35–44, serving as a lower bound on the share of children who ever lived as household heads with their parents in the household. Note that since all children within a family are included in the base, with an average of 4 to 5 children per family, the share of families in which children ever lived with their elderly parents was significantly higher. Consistent with evidence that Social Security led to reductions in co-residence of the elderly (e.g., McGarry and Schoeni, 2000), co-residence patterns changed significantly for cohorts reaching age 25 in 1950 or later, who lived with their parents at much lower rates.

Systematic data on other types of non-monetary support from adult children to their elderly parents is more scarce. But evidence from sociological studies in the 1950s and early 1960s, when many elderly cohorts were still not eligible for Social Security, indicates that such support was widespread (Schorr, 1960). Notably, many of the major forms of non-monetary help that children provided their parents other than shelter required both time and spatial proximity. These types of help included physical care, shopping, and performing household tasks (Sussman, 1965). Help during illness was particularly common (Streib, 1958; Sharp and Axelrod, 1956).

Monetary transfers from children to their parents were present, but connections within extended families did not primarily take this form (see e.g. Schorr, 1960). In a supplement to the 1952 Current Population Survey, Steiner and Dorfman (1957) found that less than 1 percent of couples and unmarried men over 65, and only 2.3 percent of unmarried women, received any regular monetary contributions from people outside their household. Irregular monetary contributions—such as pay-

[5] The decline prior to 1940 was most likely not due to Social Security retirement benefits, for which regular payments did not begin until 1940, but Old Age Assistance did play an important role (Fetter, Lockwood and Mohnen, 2018). We focus on co-residence in which parents were not household heads so it does not reflect changes in children's propensity to leave their childhood home.

ing a doctor or hospital bill in cases of serious illness—were more prevalent. Nevertheless, Steiner and Dorfman (1957) regarded these irregular monetary contributions as being less important than co-residence as a form of support that adult children provided their parents.

Daughters and sons differed in their connections with aging parents, and we posit that the degree to which Social Security would easily substitute for daughters' and sons' forms of support was not symmetric. Panel (b) of Figure 2 shows that until 1940, co-residence rates of older parents with daughters and sons were not substantially different, and in fact in 1900, co-residence with sons was slightly more common than with daughters.[6] On the other hand, it has been documented in more recent periods that non-housing forms of support, such as cooking, cleaning, or caregiving, are disproportionately provided by daughters (Horowitz, 1985; Grigoryeva, 2017), and this was likely the case historically as well: In the 1957 National Survey of the Aged, Shanas (1961) found that people 65 and older were about twice as likely to name a daughter as to name a son as the person they would ask for help during an extended illness.

A critical difference between these forms of support is how readily a cash transfer like Social Security could substitute for them. Our hypothesis is that much of the in-kind support sons provided relatively more intensively was fungible with cash. Shared housing, together with the food and other living expenses bundled with it, was substitutable with cash in the sense that parents with a guaranteed income of their own could instead pay for housing and everyday expenses on the market and live independently, as the literature on income and independent living among the elderly indicates (e.g., Costa, 1997, 1998, 1999). The hands-on care that daughters provided relatively more intensively was different: Social Security's cash benefits substituted poorly for caregiving, which was harder to purchase on the market, or which families may have preferred not to purchase. Even when care could be purchased, doing so likely would not free a child to move away as substituting cash for shared housing would, due to the costs of monitoring the quality of purchased care from afar. For both reasons, Social Security should have relaxed the location constraints of sons—whose support to parents was more substitutable with cash—more than those of daughters.

Our discussion has highlighted upstream transfers, but downstream transfers, and resource-sharing more generally, were also prevalent. There was a significant degree of co-residence of adult children with older parents as the household heads, and regardless of headship status, co-residence may have been a matter not of transfers in one direction or another but rather of mutual economic benefit (Schorr, 1960). Older parents also provided a substantial amount of non-housing services, such as help caring for young children or help during illness (Schorr, 1960; Sussman, 1965). Finally, downstream financial transfers were more prevalent than upstream transfers (Sussman, 1953; Streib, 1958; Schorr, 1960; Sussman, 1965), but they were largely made when both

[6] It is noteworthy, and consistent with our findings, that over the middle decades of the 20th century, sons' co-residence fell significantly faster than daughters'.

children and parents were relatively young (such as in the early years of children's marriage).

3 Data

3.1 Linked Numident-census data

Our primary analysis relies on two data sources: the full count US Censuses from 1920, 1930 and 1950 (Ruggles et al., 2024), and the public version of the Social Security Numerical Identification (Numident) File released by the National Archives and Records Administration. Contained in the Numident are Social Security death records for 49 million individuals, which include information on individuals' last known ZIP code of residence. Our goal is to link individuals in death records to their fathers in the 1930 Census, which allows us to predict their fathers' future Social Security eligibility based on their age and employment information. Because death records only contain basic information on individuals (full name, date of birth, and place of birth), the basis of our linked sample is a second set of Numident records: Social Security application (Form SS-5) records. Most SS-5 records can be linked to death records via Social Security Numbers (SSNs) and they additionally list parents' full names. This allows us to link individuals and their parents as families to households in the 1930 Census, improving both the quality and the number of links we are able to make. Since some children may have left their parents' household by 1930, we augment our sample by combining analogous links to the 1920 Census and links between households in the 1920 and 1930 Censuses. We also link children in SS-5 records to the 1950 Census, which allows us to observe their location in early-to-mid adulthood and study their migration patterns.[7]

Appendix Figure A.2 provides an overview of how we construct our linked sample. We start by leveraging linkages between siblings in SS-5 records based on parent names made by Bailey et al. (2026), which allows us to reconstruct families in the SS-5 data. We then link these families to households in the 1920 and 1930 Censuses based on parent names and information on children (first and middle name, age, and birthplace). Next, we link households in the 1930 Census to households in the 1920 Census using the same information as the previous links, except that we can additionally exploit parents' age and birthplace. Lastly, we link children in SS-5 records to themselves in the 1950 Census based on name, age, and birthplace information. Crucially, because we can observe both married and birth names for women in SS-5 records, we are able to find them in 1950 regardless of their marital status at the time.[8]

Each of these linkages involves linking records primarily based on names, which is notoriously

[7] The public Numident Files have been independently linked to census data using a variety of approaches, including Goldstein et al. (2023), Green (2024), Althoff et al. (2025), Ruggles et al. (2025a), and Bailey et al. (2026).

[8] We can infer women's birth names from their father's last name. For women who were already married when they first filed Form SS-5, the last name they list is their married name in most cases. Women who subsequently got married typically filed another Form SS-5 to update their last name information for administrative purposes.

challenging. In this paper, we adopt the supervised machine learning (ML) approach from the Longitudinal, Intergenerational Family Electronic Micro-Database (LIFE-M) Project (Bailey et al., 2022, 2023). Like other supervised ML approaches, this method involves making manual linking decisions for a random sample of records, which are then used to train an algorithm to make similar linking decisions for millions of records. A key feature of the LIFE-M training process is that human trainers examine an entire set of potential candidates for each record and decide whether or not to make a link and which link to make. Human input is useful as there are often multiple plausible candidates due to name commonality, and deciding which candidate to select or concluding there is too much ambiguity is not something that can easily be automated. A valid concern is that the quality of the model links is only be as good as the quality of the training data (Abramitzky et al., 2021). We mitigate this in two ways: (1) by instructing trainers to only make links when they are confident, and (2) by aggregating decisions from multiple trainers since links that are independently chosen are more likely to be objectively good links.[9]

The resulting hand-linked data is used to discipline an algorithm to mimic human decisions. We use the two-stage model from Bailey et al. (2023). In the first stage, the probability that there exists a link in the set of potential candidates is modeled using a random forest model. In the second stage, the probability that any link in the set is the true link, conditional on there being a link in the set, is modeled using a logit-style discrete choice model satisfying the property that probabilities sum to one within each set. These probabilities are then multiplied to obtain the unconditional match probability for every pair of records. Model-based links are determined using a threshold rule. In general, lowering the threshold raises the share of human links that the model is able to reproduce (recall rate), but lowers the share of model links that agree with human links (precision rate). This creates a natural trade-off between making more links and fewer false links, which training data allows us to control. In practice, we select the threshold associated with a 97 percent precision rate. Finally, using the trained model and the chosen threshold, linking is scaled to the full sample of records. More details on each specific link, including how we generate sets of potential candidates, what information is shown to trainers, and how well our models approximate trainer decisions, are provided in Appendix A.

After combining links between SS-5 records, death records and the 1920 and 1930 Censuses, we restrict the sample to our population of interest: children born between 1895 and 1930 whose fathers were born between 1875 and 1888 and covered by Social Security in 1935 or 1950 based their employment information in 1930, excluding self-employed workers.[10] The resulting linked

[9]Following the LIFE-M process, two trainers make linking decisions for a random sample of records. Cases where they disagree are shown to three additional trainers (mixed in with a random sample of other cases). Only links chosen by the two original trainers or by four out of five trainers are considered links for the purpose of training the model.

[10]We focus on fathers born in 1875–1888 for two reasons. First, including men born earlier would lower our sample's coverage since their children are more likely to have left the household by 1920. Including later cohorts

sample contains around 1.6 million children belonging to around 860,000 families. Appendix Figure A.3 plots the cohort distribution of fathers and children in the population and in our linked sample. The implied coverage rate of fathers in our population of interest is 22 percent while the approximate coverage rate of their children born by 1930 is 18 percent.[11] We are also able to link around 600,000 of children in our main sample to the 1950 Census. Importantly, our dataset covers sons and daughters at similar rates. This is because household linking is gender-neutral while the availability of married names enables us to find women in 1950 even if they change their last name.

Two key factors contribute to the incomplete coverage of our death-census linked sample. First, only a subset of children appear in the public version of the Numident, which only covers individuals who ever applied for a Social Security Number, died prior to 2007, and whose deaths were not state-reported. In practice, this file has near-universal coverage of deaths that occurred between 1988 and 2007, and relatively low coverage of deaths that occurred prior to 1988 (Goldstein and Breen, 2022). Second, human trainers are only able to confidently link a subset of records, and the model is only able to replicate a subset of those links at a 3 percent error rate.

Incomplete coverage raises two types of selection concerns: (1) which families appear in our linked sample, and (2) which children from those families appear in our sample. Appendix Table A.1 assesses the representativeness of our linked sample in terms of 1930 household characteristics. Overall, our sample is fairly representative of our population of interest, though as is common in historical record linking, Black and lower-socioeconomic status families are underrepresented as they tend to have more common names. To account for these differences, we re-weight families in our linked sample using inverse propensity score weights generated using the procedure from Bailey et al. (2020).[12] The resulting weighted means in column 3 are very close to the population means in column 1, even for characteristics we did not explicitly target.

Intuitively, selection of families into our linked sample is relatively mild because most families in our population of interest had at least one child who ever applied for an SSN and died between 1988 and 2007. However, our sample clearly contains the subset of children who satisfy these criteria. In particular, as is evident in Appendix Figure A.3, our linked sample is skewed towards later cohorts of children as they were more likely to survive until 1988. This reflects the selection of cohorts into SS-5/death records, prior to linking (see Appendix Figure A.4). On the other hand, selection in terms of place of birth or place of death is less of a concern given that SS-5/death

would have a similar effect since their children are more likely to be born after 1930. Second, as will be evident in Section 4, the variation in Social Security coverage we leverage is most salient for the 1875–1888 birth cohorts.

[11] We cannot compute the true coverage rate for children since our population of interest conditions on fathers' employment status in 1930 (and not all children live with their fathers in 1930). Instead, 18 percent is the coverage rate of children whose fathers were born in 1875–1888 regardless of their Social Security coverage status, which is based on the number of children aged under 10 (or 15) with fathers from those cohorts in the 1910–1930 Censuses.

[12] The weighting variables are fixed effects for father year of birth, race, place of birth, and literacy status, as well as fixed effects for household home ownership, radio ownership, farm status, and urban status.

records have excellent geographic coverage (see Appendix Figure A.5). Selection into our linked sample matters for internal validity only to the extent that it is correlated with the Social Security policy variation we exploit. In Section 4.4, we show that this is not the case.

The Numident also includes a third set of files: Social Security claims records for 25 million individuals. We build a separate dataset linking men born in 1875–1888 in the claims records to themselves in the 1930 Census. We will use this sample to test whether fathers' predicted eligibility for Social Security based on their 1930 employment information is correlated with their subsequent claiming behavior. These linkages are made based on name, age, birthplace, and place of residence information using the same supervised ML approach (see Appendix A.4 for more details). Ultimately, we are able to link around 360,000 men covered by Social Security in 1935 or 1950 at a 97 percent precision rate, or around 10 percent of the population. This lower coverage rate reflects both the coverage of claims records and the more limited information we can use for linking. We similarly re-weight this sample for representativeness (see Appendix Table A.2).

3.2 1973 Survey of Occupational Changes in a Generation

In the analysis, we also present complementary evidence on sons' labor market outcomes in adulthood using data from the 1973 Survey of Occupational Changes in a Generation. The 1973 OCG was a supplement to the 1973 Current Population Survey (CPS) that focused on civilian males aged 20–65 and collected detailed information on respondents, their parents, and their spouses, with a special focus on occupational histories (Blau et al., 1994). Two features of the OCG are uniquely useful for our purposes. First, it contains retrospective information on respondents' father's year of birth and occupation, industry, and self-employment status when respondents were aged 16. This allows us to predict fathers' likely Social Security eligibility age. Second, a subset of respondents were linked to their SSA earnings records covering the period 1937–1975 (Social Security Administration, 2008).[13]

In addition to containing a rich set of outcomes, another advantage of the OCG relative to our main linked sample is that it is a nationally representative sample. Therefore, it does not suffer from selection issues tied to the incomplete coverage of public Numident records or the fact that we can only link a subset of them to the 1930 Census using name and demographic information. The main drawback of the OCG is small sample sizes: After restricting to sons born in 1907–1930 whose fathers were born in 1875–1894 and covered by Social Security in 1935 or 1950 based on

[13]Because there are no unique identifiers to merge the OCG file with the Social Security Longitudinal Earnings file (SSLEF), we use an alternative approach. We first merge the SSLEF with another file, the Social Security Records: Exact Match Data (EMD) file (Social Security Administration, 2005). This is done based on row order since demographic variables that appear in both files (sex, race, month of birth, and year of birth) match for over 99% of observations using this approach. We then identify 34 variables that appear in both the EMD and OCG files, restrict to observations that are uniquely identified by these variables (over 99.9% of the sample), and merge on these variables.

their retrospective employment information, we are left with around 3,000 observations.[14]

4 Empirical Strategy

4.1 Parameterizing variation in Social Security eligibility

The expansion of the definition of covered employment over time, and its interaction with birth cohort, forms the basis of our empirical approach. To fix ideas, consider two married couples, in both of which the husband was born in 1877 and the wife was three years younger. Suppose that in one couple, the husband worked in employment covered under the original 1935 Act, in the other the husband's employment was not covered until 1950, and in both couples the wife did not have any covered employment. Between 1937 and the middle of 1939, the 1935-covered worker could accrue the 10 quarters of coverage he needed under the 1939 Amendments to become permanently fully insured (and hence eligible to receive benefits upon reaching age 65). Upon reaching age 65 in 1942, he could begin receiving primary benefits; once his wife reached age 65 in 1945 she could additionally begin receiving spousal benefits; and if he predeceased her, she would receive widow's benefits until the end of her life.[15] In contrast, if the 1950-covered worker did not switch to 1935-covered employment, it was not until 1951 that he could begin accruing quarters of coverage, and not until the middle of 1952, around age 75, that he would accrue the six quarters of coverage he needed under the 1950 Amendments to become eligible to receive benefits. Hence, even conditional on working long enough to be eligible, the 1950-covered couple would receive benefits much later in life, and hence for a shorter period of time.

The difference between 1935- and 1950-covered workers, however, became significantly smaller for cohorts born even a few years later. If the two workers above were born in 1888 (and their wives in 1891), for example, then even if the 1950-covered worker remained in 1950-covered employment, both workers could accrue the 6 quarters of coverage they needed under the 1950 Amendments by the middle of 1952, around age 64 (although the 1935-covered worker could, of course, accrue them earlier), and both could begin receiving benefits upon reaching age 65, with identical timing of primary, spousal, and survivor's benefits.[16]

To implement the idea of this comparison, we use the 1935 Act and the 1939 and 1950 Amend-

[14] 1907 is the earliest birth cohort in the OCG (aged 65 when surveyed but turning 66 in 1973), while 1930 is chosen as an upper bound for consistency with our main linked sample and to measure sons' occupations when they are at least in their 30s. We slightly expanded the set of father cohorts to increase sample sizes.

[15] In 1940, remaining life expectancy at age 65 was about 12 years for men and 13.5 years for women (Grove and Hetzel, 1968).

[16] Benefit amounts could differ between the two, but we focus on differences in eligibility because while the Numident claims data do allow us to establish a first stage for claiming benefits, we do not have data that would allow us to establish a first stage for benefit amounts. It is also worth noting that the "New Start" provision of the 1950 amendments tended to equalize benefits across workers from the two groups.

ments to calculate the earliest age at which an individual from a given birth cohort and with a given employment type could begin collecting monthly benefits, under the assumption that those in 1950-covered employment do not engage in any 1935-covered employment. This earliest age is the primary parameterization of "exposure" to Social Security in our analysis, and is illustrated for a broad range of birth cohorts in Figure 3; as noted earlier, our sample restricts attention to fathers born from 1875 through 1888. Note that men were eligible to claim before age 65 starting only in 1961, which is relevant only for cohorts born in 1897 or later.

4.2 Comparing similar workers with differing Social Security eligibility

We get empirical traction on this parameterization of likely Social Security eligibility by observing an individual's age (and hence birth cohort) and employment information in the 1930 Census, five years prior to the Social Security Act. For all workers, the Census recorded an occupation, industry, and class of worker (e.g. self-employed, working for wages or salary, or unpaid). We classify all combinations of these three variables into a year when this type of employment was covered by Social Security. Full details are provided in Appendix B.1, and Appendix Table A.3 shows that in the 1940 Census, our coding is highly predictive of workers reporting that they had a Social Security Number. A priori, one would expect that the types of workers who were covered in 1935 would systematically differ from those covered in 1950; the self-employed, domestic workers, and agricultural wage workers would likely have different underlying characteristics than wage and salary workers in manufacturing, for example. To the extent that these differences are constant across birth cohorts, one could net these differences out in a specification of the form

$$y_f = \alpha_{b(f)} + \delta_{g(f)} + \beta \cdot (\text{Social Security exposure})_{b(f)g(f)} + \Lambda' \mathbf{x}_f + \varepsilon_f, \qquad (1)$$

where y is an outcome for family f, $\alpha_{b(f)}$ and $\delta_{g(f)}$ are fixed effects for father's birth cohort b and father's coverage category g, and $\mathbf{x}_f$ a vector of controls (fixed effects for father's race and state or country of birth). For expositional purposes, we define Social Security exposure as the negative of the earliest age fathers could receive Social Security benefits based on their employment information in 1930 so that the coefficient of interest, β, captures the impact of fathers being predicted to be eligible for Social Security benefits one year *earlier*.

To isolate more compelling variation in Social Security exposure, however, our empirical approach is based on a finer comparison. In particular, the exclusion of certain industries in the original act generates variation that allows us to exclude the self-employed and to make comparisons of wage and salary workers with the same occupation, but working in different industries. As noted earlier, employees of nonprofit organizations, state and local governments, and private households were not covered under the 1935 Act or 1939 Amendments, even though work in the same occupation in other sectors was covered. For example, work as a janitor for a manufacturing firm would have been covered in 1935, but work as a janitor in a school, a nonprofit hospital, or a church would not have been covered until at least 1950. Differing treatment of similar workers did not escape the notice of the Social Security Board: Arthur J. Altmeyer, its Chairman, noted in 1944 that many employees of nonprofit institutions were not covered, despite the fact that their "skills, tasks, and earnings. . . do not usually distinguish them from comparable employees in commerce or industry" (Altmeyer, 1944).

To compare workers in the same occupation but with coverage status differing due to industry of employment, we exclude self-employed workers from our sample and estimate a specification of the form

$$y_f = \alpha_{b(f)o(f)} + \delta_{g(f)} + \beta \cdot (\text{Social Security exposure})_{b(f)g(f)} + \Lambda' \mathbf{x}_f + \varepsilon_f, \tag{2}$$

where, relative to equation (1), we include fixed effects for the interaction of father's birth cohort and father's 1930 3-digit occupation (subscripted by o) in order to focus on variation in earliest age of eligibility arising from work in different industries. Relative to simply controlling for occupation fixed effects, the interaction term accommodates the fact that having the same 1930 occupation may mean different things for workers born in different years. In this setup, the three coverage categories in $\delta_{g(f)}$ are: covered in 1935, covered in 1950 via domestic work, and covered in 1950 via work in the nonprofit or state and local government sector. Observations are weighted using the inverse propensity score weights from Appendix Table A.1, and standard errors are clustered at the 3-digit industry level. Equation (2) serves as our preferred empirical specification when analyzing family or parental outcomes. When analyzing the impact of father's Social Security eligibility on children's outcomes, we estimate a variant of equation (2)

$$y_{if} = \theta_{s(i)c(i)} + \alpha_{b(f)o(f)} + \delta_{g(f)} + \beta \cdot (\text{Social Security exposure})_{b(f)g(f)} + \Lambda' \mathbf{x}_f + \varepsilon_{if}, \tag{3}$$

where the unit of analysis are children i and we additionally include fixed effects for the interaction of children's sex and birth cohort (subscripted by s and c). In our main analysis, these fixed effects account for any average differences in late-life neighborhood quality that may result from the uneven cohort coverage of Numident records (Appendix Figure A.4). When analyzing outcomes in the 1973 OCG or in the 1950 Census, they additionally account for life cycle effects since each cohort's outcomes are measured at a different age.[17]

As described in detail in Appendix Section B.1, we use the industry classification of the Census to identify employees who are less likely than others to have been covered by the original 1935

[17] In equation (3), individuals are weighted by our baseline family-level weights normalized by the number of children that appear in our linked sample. This avoids mechanically placing more weight on larger families.

Social Security Act. Work for private households was identified as a separate industry category. Nonprofit organizations were concentrated most heavily in hospitals, educational institutions, and welfare and religious organizations, in addition to some other nonprofit membership organizations. Hospitals and educational services were also important areas of state and local government employment, in addition to more general state and local public administration. Hence, we treat wage and salary workers in these industries as not covered until 1950.[18] Note that while most employment in hospitals and educational institutions was in nonprofit organizations or in state and local governments, there were a handful of for-profit organizations in these sectors, as we discuss in Appendix B.1 and C. Any misclassification of this sort should tend to attenuate our results.

The occupations that form the basis of our comparison are those that were present in both for-profit and either nonprofit or public enterprises, and to a lesser extent private households.[19] Appendix Table A.4 shows the 20 largest occupation categories in the population and our linked samples in which there is non-trivial variation in Social Security coverage status (defined here as the smaller group comprising at least 1 percent of that occupation). Among the largest categories are miscellaneous types of laborers, managers and clerical workers that appear in many industries, and also carpenters, truck drivers, and janitors.

For interpretation of our estimates, it is useful to note that some workers we classify as not covered by Social Security until at least 1950—state and local government workers in particular—may have been covered by existing retirement systems. To the extent that this was the case, it is likely to attenuate our estimates. But it also bears emphasis that other pensions were less common for the workers we focus on than aggregate statistics would suggest. While 46 percent of all state and local government employees were covered by a retirement system in 1942 (United States Bureau of the Census, 1943, Table 2), the types of employees most often covered by these retirement systems are either excluded from our sample (e.g., policemen, firefighters) or have little effect on the estimates because these occupations seldom appear in the private sector (e.g., teachers). Government employees that form the primary basis for our comparisons—such as custodial and maintenance employees—were less likely to be covered under existing retirement systems. Similarly, about a third of private, nonprofit institutions of higher education had retirement plans in the 1940s, but maintenance and nonprofessional employees were generally not eligible (McCamman,

[18]As described in Section 2.1, nonprofit employees and state and local workers not covered under an existing retirement system were covered on an elective basis starting in 1950, while state and local workers already covered under a retirement system were not covered by Social Security until 1954. In the Census we do not observe whether or not workers are covered by a retirement system, nor can we generally distinguish between employees of nonprofit organizations and employees of state or local governments. Hence, we treat all such workers as covered in 1950. As we will show, our results are robust to treating all such workers as being covered in 1954 instead.

[19]Note that many of the wage workers covered in 1950 for whom it is less clear that there would be similar 1935-covered workers do not play a significant role in our comparisons. In particular, most agricultural wage workers and many domestic workers are in their own occupation codes in the census classification. They are included in our sample, but with no variation in their coverage status, they do not contribute significantly to our estimates of interest.

1948). Only about 8 percent of nonprofit hospitals had pension plans at the beginning of the 1940s (Hayhow, 1940). Finally, the benefits under retirement systems for both public and nonprofit employment typically differed from Social Security in important ways. Most relevant to our analysis, these plans tended to have either no provision or less generous provisions for widow's benefits (McCamman, 1953). Widows are a particularly relevant case for our study, since co-residence and other types of family old-age support were especially common among widows. Further details on retirement systems for government and nonprofit organizations can be found in Appendix C.

4.3 Predicted Social Security eligibility and subsequent claiming behavior

The discussion above suggests two testable predictions. First, workers whose pre-Social Security employment would be covered in 1935 should be more likely to ever claim Social Security than workers whose employment would only be covered in 1950, and conditional on claiming they should claim at earlier ages. Second, these differences should be smaller for later birth cohorts than earlier birth cohorts, at least over some range. We test these predictions using our sample of Social Security claims records linked to the 1930 Census. For individuals found in the claims data, we can observe the age at which they claimed Social Security benefits. While we cannot distinguish between individuals who never claimed Social Security and individuals who did but were not linked to the claims data (or simply do not appear in it), ever claiming should be positively correlated with the probability of being linked to a claims record. Hence, we can establish the existence of a "first stage," although tests based on linkage to the claims records will understate the true differences in the probability of claiming, as we discuss below.

Table 1 reports estimates from equation (2), and shows that the variation in Social Security exposure we leverage is strongly consistent with these predictions. In column 1, the sample is the 1930 population of men born in 1875–1888 and covered in 1935 or 1950 (excluding the self-employed), and the dependent variable is an indicator for being linked to a claims record. Eligibility one year earlier leads to a highly statistically significant 0.34 percentage point increase in the likelihood of being linked to a claim for monthly benefits, relative to a baseline mean of 9.3 percent. In column 2, we focus on the subset of men linked to a claims record and the dependent variable is the age at which they claimed Social Security benefits. Conditional on being linked to a claims record, eligibility one year earlier leads to claiming monthly benefits 0.11 years earlier.

In Figure 4, we combine both predictions by estimating equation (2) with dependent variables $\mathbf{1}(\text{linked to claim at age } a)$, for ages a ranging from 65 to 85, and $\mathbf{1}(\text{linked to claim made in year } t)$, for t ranging from 1940 to 1960. The pattern of results is encouragingly consistent with the predictions that underlie our empirical strategy. In particular, panel (a) shows that men we predict to be eligible at earlier ages have a greater likelihood of claiming, with the largest effects at earlier ages (between 65 and 70). Furthermore, since men working only in 1950-covered employment would have started to become eligible starting in 1952, it is consistent with our predictions that those who were eligible earlier had a higher likelihood of filing retirement claims during the 1940s, but a *lower* likelihood of filing claims in years from 1952 through 1954.

The finding that fathers' 1930 employment is predictive of their later claiming behavior implies that a significant share of workers whose work had not been covered in 1935 did not switch into 1935-covered employment. Indeed, we can verify directly that only 20 percent of men in 1950-covered employment in 1930 had switched to 1935-covered employment by 1940, based on links from the Census Linking Project (Abramitzky et al., 2025). Some of these cases may also be due to imperfect classification of either the 1930 or 1940 occupation and industry. Although there are many potential explanations of workers' stickiness in employment type, we suspect that job mobility over the 1930s was limited.

We do not formally treat the estimates from Table 1 or Figure 4 as a first stage—and in particular do not use them to scale our reduced-form estimates. These estimates understate the true difference in probability of claiming retirement benefits for two reasons: the Numident claims dataset contains only a subset of all claims, and not all claims are successfully linked to a Census record. The total number of primary claims for men in the 1875–1888 cohorts that we have linked to the Census represents about 14.5 percent of the total number of benefit awards to this group reported in the *Social Security Bulletin*, suggesting that the true size of the first stage is significantly higher. In addition, we are linking primary (retirement) claims to the Census, but not survivors' claims. As a result, even a complete linkage of every retirement claim made by these cohorts would not fully capture the variation in benefits relevant to families' outcomes: if an eligible father died before claiming retirement benefits but his widow subsequently claimed survivors' benefits, we would record the father as having no linked claim while still regarding the children as being "treated." Nevertheless, they do establish that there is a strong relationship between our parameterization of Social Security eligibility and subsequent claiming behavior, even when restricting comparisons to fine occupational categories.

4.4 Selection into the linked sample

Due to the coverage of Numident records and the fact that we can only link a selected subset of them to the censuses, the families and children that appear in our linked sample are a selected subset of the population. Even though we re-weight families in our sample to resemble the population in terms of observable characteristics, non-random selection could potentially affect the internal validity of our estimates if it is correlated with the variation we exploit. To assess the extent to which selection is a concern, we examine all households in our population of interest in 1930—those headed by a man born in 1875–1888 and covered by Social Security in 1935 or 1950 based on his employment information (excluding self-employed workers)—and test whether appearing in our linked sample is correlated with our Social Security eligibility variation. The results are reported in Table 2. Reassuringly, we find no evidence that fathers' earliest age of Social Security eligibility is correlated with the likelihood of the household appearing in our linked sample, the likelihood of any son or daughter appearing in our linked sample, or the number of sons or daughters appearing in our linked sample.[20]

4.5 Placebo checks in the 1930 Census

A key check on whether our empirical approach allows for valid causal comparisons is to test whether our variation is correlated with household characteristics in 1930, five years prior to the introduction of Social Security. The set of characteristics we examine are home ownership, home value, radio ownership, urban and farm status, number of children present in the household, father and mother literacy status, and mother labor force participation status. Table 3 reports estimates based on two approaches, first testing for differences in the full population of households in panel (a), and then testing for differences within the subset of households contained in our linked sample in panel (b). The fact that all the coefficients are close to zero and statistically insignificant confirms that conditional on making comparisons between households with fathers in the same occupation, there were no significant pre-existing differences between households whose fathers were predicted to be eligible for Social Security earlier versus later.

5 Results

5.1 Children's late-life neighborhood characteristics

In this section, we estimate the impact of fathers' Social Security eligibility on children's late-life socioeconomic status as proxied by the "quality" of their last known ZIP code of residence, which has also been used in other work (Althoff and Reichardt, 2024; Green, 2024). We measure neighborhood quality using ZIP-level mean adjusted gross income (AGI) in 2001, median home values in 2000, or a neighborhood quality index that captures a broader set of indicators. Partly following Bailey et al. (2024), this index additionally includes the following ZIP characteristics: (1) the homeownership rate in 2000, (2) median rents in 2000, (3) the share of college-educated individuals in 2000, (4) the share of individuals under the poverty line in 1999, (5) the share of households receiving public assistance in 1999, (6) the share of single-parent families in 2000, and

[20]In Appendix Table A.5, we also find no evidence of selection of children from specific cohort ranges, with the exception of daughters born in 1895–1910 or 1921–1930. As we will show, our main findings are driven by sons so we do not view this as a major concern.

(7) the measure of upward income mobility from Chetty et al. (2026).[21]

Intuitively, children's last known ZIP code should be informative about how well they fared over their lives, as access to good neighborhoods is a function of accumulated net resources. Here we provide novel empirical evidence to support this claim using restricted data from the Health and Retirement Study (HRS), a nationally representative longitudinal survey of Americans aged 50 and older. Specifically, we leverage information on HRS respondents' ZIP code of residence at the time of their last interview as well as information on their lifetime earnings from SSA administrative records (see Appendix A.5 for more details). This allows us to examine the relationship between lifetime earnings and late-life ZIP quality. Appendix Figure A.6 provides a visual representation of this relationship by plotting moments of the distribution of lifetime earnings through age 65 by ZIP-level income percentile rank bin.[22] In panel (a), the sample consists of men and their own lifetime earnings are used. In panel (b), the sample consists of women whose spouses also responded to the HRS and the couple's lifetime earnings are used (since many women from the HRS cohorts had low lifetime labor force attachment).

Both figures reveal a strong, positive relationship between the average income in respondents' late-life ZIP code and their lifetime earnings. The percentiles imply that the distribution of lifetime earnings shifts to the right as late-life ZIP income increases. For example, median lifetime earnings for men in the bottom ZIP income bin is $707,000 (in 2016 dollars), rises to 2 million dollars in the 25th bin (which includes the median ZIP code), and reaches 3.7 million dollars in the top bin. The patterns are qualitatively similar for women and other percentiles of the lifetime earnings distribution. In our sample, the correlation between late-life ZIP income and lifetime earnings is 0.28 for men, while for women the corresponding correlation with household lifetime earnings is 0.29. Overall, this suggests that late-life neighborhood characteristics are a valid proxy for long-term socioeconomic status.

Table 4 shows the OLS estimates from equation (3), where the dependent variables are ZIP-level mean AGI, median home values, or the neighborhood quality index (either in percentile ranks, logs or levels). Throughout the analysis, we estimate separate models pooling all children or restricting the sample to sons or daughters. The positive coefficients imply that children whose parents were more exposed to Social Security—i.e., whose fathers were predicted to be eligible for Social Security at an earlier age—lived in better neighborhoods towards the end of their life.[23] Im-

[21] Data on ZIP-level mean AGI comes from the Internal Revenue Service, while data on other ZIP-level indicators comes from the National Historical Geographic Information System (Schroeder et al., 2025). The neighborhood quality index is constructed in three steps: (1) normalize each sub-index to have a mean of 0 and standard deviation of 1, (2) take the average across all sub-indices, and (3) normalize the resulting composite index.

[22] HRS respondents are first assigned a percentile rank depending on where their last known ZIP code of residence falls within the national distribution of ZIP codes according to mean AGI in 2001. The 100 ZIP percentile ranks are then aggregated into 50 bins, each containing two adjacent percentile ranks.

[23] Appendix Figure A.7 shows results based on ranking ZIP codes according to the other sub-indices underlying the

portantly, we find that the these effects are primarily driven by sons, for whom the coefficients are twice as large in magnitude and highly statistically significant. In contrast, consistent with the predictions discussed in Section 2.2, the coefficients for daughters are closer to zero and insignificant at conventional levels. In terms of magnitudes, a 10-year earlier predicted Social Security eligibility for fathers—roughly the difference between being covered in 1935 vs. 1950 for workers born in 1877—leads to a 2-percentile increase in the rank of sons' late-life ZIP code. Alternatively, the same 10-year difference leads to a 2.4-percent increase in ZIP-level income, a 4.3-percent increase in ZIP-level home values, and a 0.054-standard deviation increase in neighborhood quality.

The results are robust to a wide range of alternative specifications. Appendix Figure A.8 shows that the ZIP income rank results for sons and daughters are robust to using alternative measures of exposure to Social Security (a variant of our baseline measure that assumes that 1950-covered workers become eligible in 1954 instead of 1950 to account for the liberalization of coverage provisions for state and local government employees in the early 1950s, and a dollar-based measure capturing expected average monthly Social Security benefits between ages 65 and 84) and including additional controls (fixed effects for fathers' 3-digit industry, fixed effects for fathers' 3-digit occupation-by-Social Security coverage category, controls for household characteristics in 1930, and fixed effects for parents' 1930 county of residence). Occupation-by-coverage fixed effects allow for 1930 occupation to carry different implications depending on the worker's Social Security coverage category. The fact that our estimates are robust to the inclusion of industry fixed effects alleviates a potential concern that our findings reflect differences in industry-based exposure to the Great Depression. The county fixed effects alleviate any concern that our findings might reflect differences in the characteristics of parents' 1930 locations that are correlated with our Social Security eligibility variation. This rules out differential place-based exposure to the Great Depression, among other potential concerns.

Appendix Table A.6 uses alternative samples linking parents to the 1940 Census to provide evidence that the results we see for sons are not due to changes in outcomes for their families over the 1930s, prior to eligibility for Social Security. We find little evidence of differences in wage income, labor force participation, or employment in 1940 for men predicted to be eligible for Social Security earlier. Since education is reported in the 1940 Census, we additionally test for differences in years of education for both fathers and mothers. We find no evidence that mother's education differed across families with different Social Security exposure, and although we find a statistically significant, positive relationship between Social Security exposure and father's education in one sample, it implies that a 10-year difference in predicted eligibility is associated with only a 0.1-year difference in father's education, relative to a mean of 7.4, too small to account for the magnitude

neighborhood quality index. We find that the higher neighborhood quality difference is associated with more-educated and lower-poverty areas, but we see little difference in homeownership rates or rates of upward income mobility.

of the gains we find for sons; in addition, this result does not appear in the other sample.

5.2 Sons' labor market outcomes

The findings in the previous section show that sons whose fathers were likely eligible for Social Security at an earlier age lived in better neighborhoods towards the end of their lives. Although we showed that late-life neighborhood quality is positively correlated with lifetime earnings, the 1973 OCG survey provides more direct evidence that sons had higher earnings in adulthood.

Table 5 displays the estimates from equation (3), where the dependent variables are the occupational income scores associated with sons' first job, the job they held in 1962, and their longest job in 1972 (in logs or levels). Occupational income scores are obtained by first mapping OCG occupation codes into 1950 Census occupation codes and using the race-and-region-specific occupational income scores for men from Collins and Wanamaker (2022). We use 1940-based scores for sons' first job, 1960-based scores for their job in 1962, and 1970-based scores for their job in 1972. Since OCG respondents were asked about their father's employment information when they were aged 16, which might have endogenously responded to the introduction of Social Security for fathers of sons born after 1919, we show separate results restricting to sons born by 1919.

The estimates indicate that fathers' Social Security eligibility led to significant earnings gains that appear in the middle and later periods of their sons' careers. The first two columns show no detectable impact on sons' first jobs, perhaps because parents had not yet received Social Security, or because gains simply appeared later. In contrast, the estimates for sons' occupational income score in 1962 and 1972 are large in magnitude and generally statistically significant. Because outcomes in any given calendar year are at a broad range of ages and effects may differ across the life cycle, we also examine impacts on sons' occupational standing around age 50, which we define as 1962 for sons born in 1907–1917 cohorts and 1972 for sons born after 1917. We find similar results. We broadly summarize these results as implying that a 10-year earlier predicted eligibility for fathers increased sons' occupational income scores in adulthood by around 35 percent.[24]

Note that the roughly 35 percent increase in sons' late-career occupational income scores are not directly comparable to the 2.4 percent increase in ZIP-level income and 4.3 percent increase in ZIP-level home values. The former are individual-level measures, whereas the latter are neighborhood-level measures. To assess whether the two sets of estimates are quantitatively consistent, in Appendix Table A.7 we estimate the cross-sectional relationship between the log of late-career earnings (ages 50–65) and the log of late-life ZIP income or home values using

[24]Although occupational information is missing for around 10 percent of respondents, we find no evidence that this is correlated with our Social Security variation. Appendix Figure A.10 shows that the estimates for sons' log occupational income scores around age 50 are also robust to alternative ways of measuring fathers' exposure to Social Security and the inclusion of additional controls.

the restricted HRS data described in Section 5.1. The estimated gradient for ZIP-average income is about 0.085. Thus, 10-percent higher late-career earnings is associated with only about 0.85-percent higher ZIP-level income, reflecting substantial within-ZIP heterogeneity in resources. So if occupational income score gains map one-for-one into earnings gains, a 35 percent occupational income score gain yields a predicted ZIP income effect of around 3.0 percent, fairly close to our direct estimate of 2.4 percent and within its 95% confidence interval. The analogous calculation for ZIP-level home values yields about 3.2 percent, somewhat below our direct estimate of 4.3 percent but within its 95% confidence interval. We therefore view the earnings and neighborhood estimates as quantitatively consistent once individual-level gains are translated into neighborhood-level averages. See Appendix D for more details on this reconciliation exercise.

One drawback of occupational income scores is that they only capture variation in average earnings across occupation-region-race cells. To provide evidence on earnings that vary at the individual level, we leverage the fact that a subset of the OCG was linked to SSA earnings records from 1937 to 1975. As measures of income, these SSA earnings records have the advantages that they are not self-reported and that we observe them over many years, making them less prone to transitory income fluctuations than income observed at a single point in time. We focus on earnings from 1951 onwards because the data only contains information on aggregate earnings for the 1937–1950 period, which prevents us from converting them into comparable real dollars.

The full period from 1951 through 1975 contains years when some of the cohorts we study are still entering the labor force, and years when other cohorts are near retirement age. We aim to focus on ages at which our cohorts are mostly working full time, which presents a tradeoff between examining a broad range of calendar years and a broad range of cohorts. In panel (a) of Appendix Figure A.9, we show the share of respondents in the OCG that reported four quarters of coverage (QCs) in each calendar year, separately for two early (1907 and 1909), one middle (1919) and two late cohorts (1928 and 1930). At their peak, around 80 percent of each cohort have four QCs; the other 20 percent may have been working and earning full-time in employment not covered by Social Security, or not working full-time. A smaller share of the youngest cohorts had four QCs in the early years as they entered the labor force, and similarly the oldest cohorts had fewer QCs in the later years as they retired. Based on these patterns, we focus on subsets of years in which the cohorts we examine have steady and high rates of Social Security-covered employment: 1957–1967 when examining the 1907–1930 cohorts, and 1951–1967 when examining the 1907–1919 cohorts. As an alternative approach, we also examine earnings from ages 44 through 53, the 10-year range that allows us to include the broadest range of cohorts observed in their 40s and 50s.

The results in Table 6 provide further evidence that fathers' Social Security eligibility increased son's earnings. In columns 1 and 2, we test whether the log of total earnings over the relevant range of calendar years is higher for sons whose fathers had greater Social Security exposure; the second column limits attention to years in which the respondent had four QCs in order to alleviate any concern that those with fewer QCs had non-covered earnings that we do not observe.[25] An important caveat regarding the results on total earnings is that observed earnings are censored at the Social Security taxable maximum, and the degree of censoring is significant. Panel (b) of Appendix Figure A.9 shows that around half to two thirds of 4-QC workers have censored earnings in our data. Because we lack data to impute earnings above the taxable maximum, in column 3 we instead define an indicator for earnings exceeding the Social Security taxable maximum in *all* 4-QC years, which can be viewed as a proxy for having "high" earnings during prime working years. Columns 5, 6, and 7 report analogous results focusing on a balanced panel of ages rather than a balanced panel of years. Because of the censoring issues, we do not focus on magnitudes, but instead on the general pattern and statistical significance of the results. Across all specifications and samples, the results indicate that fathers' Social Security exposure led to a higher probability of having earnings at or above the taxable maximum in all years with 4 QCs. Estimates using the log of total earnings vary in statistical significance, but the magnitudes are fairly stable across samples. Overall, the pattern of results supports the finding that Social Security had a positive impact on sons' earnings in adulthood, consistent with the occupation-based results.

5.3 Children's migration patterns

As we have emphasized, one channel through which Social Security may have benefited children of recipients is migration. We test this hypothesis using the subset of children in our main linked sample that we can observe in the 1950 Census, 10 years after monthly benefits were first issued. Table 7 shows the impact of fathers' Social Security eligibility on children's propensity to live outside their parents' 1930 place of residence (county, state, census division, census region) by 1950.[26] We treat parents' location in 1930, when they are in their 40s and 50s, as a proxy for where they spent their last remaining years and therefore where children would have had to live to co-reside with them or care for them. Consistent with the notion that Social Security relaxed location constraints, we find that sons whose fathers were predicted to be eligible for Social Security at an earlier age were more likely to live in a different state or census division by 1950. A 10-year earlier predicted eligibility for Social Security increased sons' likelihood of moving to a different state by 2 percentage points and their likelihood of moving to a different census division by 2.5 percentage points, relative to baseline means of 23 and 16 percent, respectively. These findings are robust to a wide range of sensitivity checks, as shown in Appendix Figure A.11 for cross-division migration.

[25] In all specifications in this table, we limit to respondents with an average of at least 2 QCs over the relevant range in order to mitigate the fact that we cannot observe earnings in non-covered employment. Columns 4 and 8 show that our Social Security variation is uncorrelated with not satisfying this condition.

[26] Since county boundaries have changed over time, we use the crosswalk from Eckert et al. (2020) to map counties to 2,833 county groups with stable boundaries 1930 and 2010.

The fact that the coefficient for moving to a different county is close to zero and statistically insignificant implies that the cross-state and cross-division migration effects are largely offset by a decline in sons' propensity to move to a different county within the same state, as shown in column 2. In other words, as of 1950, the effect of Social Security seems to have been to prompt some sons to migrate farther away than they otherwise would have. Columns 7 and 8 reinforce this point by showing that sons were more likely to move to places over 500 miles away from their parents' location at the expense of moves within 100 to 500 miles. Since we can infer children's late-life location from their ZIP code in death records, we can also examine their post-1950 migration patterns. Appendix Table A.8 shows analogous results comparing children's location in 1950 to their location at death, and reveals that sons whose fathers were more exposed to Social Security were *less* likely to move to a different state, census division, or census region after 1950. Therefore, Social Security not only induced sons to migrate farther away than they otherwise would have, but also earlier. This pattern suggests that Social Security shifted some sons' migration earlier in life rather than changing whether they ever migrated, perhaps because location constraints are eventually relaxed once parents pass away.

In contrast, we find no evidence of any significant migration response among daughters. This is consistent with our hypothesis that caregiving norms prevented them from out-migrating to the same extent as their brothers. This is also in line with evidence that daughters were more likely than sons to stay back and support their parents in response to their fathers experiencing Prohibition-induced job and earnings losses in the early 20th century (Aizer et al., 2025). The absence of migration effects for daughters provides a rationale for the lack of effects on late-life neighborhood quality, reinforcing the notion that migration was a key determinant of the impact of Social Security on children's long-term outcomes.

We provide two additional pieces of evidence that migration was central to sons' long-term gains. First, we show that Social Security exposure shifts the joint distribution of early long-distance migration and late-life neighborhood outcomes in a manner consistent with migration being an important driver of the gains. Appendix Figure A.12 examines the impact of Social Security on the joint outcome of moving to a different census division by 1950 or not *and* ending up in a given quintile of the ZIP income distribution late in life. To the extent that Social Security raised the probability of moving and ending up in a high-income ZIP code at the expense of not moving and ending up in a low-income ZIP code, this would be consistent with migration being an important mediating channel. This is exactly what we find: Sons whose fathers were more exposed to Social Security were less likely to remain in their parents' census division and end up in the bottom two quintiles of the national ZIP income distribution and more likely to move to a different census division and end up in the top two quintiles of the ZIP income distribution.

Second, we provide suggestive evidence that the way migration led to long-term gains for sons was by facilitating moves to labor markets that were a better match for their skills, consistent with the positive effects on occupational standing and earnings in adulthood we have documented. For this analysis we focus on the subset of children in our linked sample for whom we can observe completed years of education, either in the 1950 Census or via additional links to the 1940 Census from Bailey et al. (2026). Because the education question was only asked to sample-line respondents in 1950 (5 percent of the population), and because many children in our cohorts of interest had yet to complete their education by 1940, this sample only contains around 200,000 children.

Education serves as a proxy of children's skills, and we measure the demand for those skills in each state in 1950 using education-specific median earnings among male workers aged 25 to 54. Intuitively, if the demand for workers with 12 years of education is high in a given state (perhaps due to its industrial or occupational mix), then that should be reflected in the earnings of the median worker with that education level. There is substantial variation in median earnings by education across states, and some states can be a good match for some workers but a poor match for others. This is illustrated in Appendix Figure A.13, which plots earnings "premiums" by state for workers with 8 versus 12 years of education. While there are similarities across both state rankings (e.g., Southern states generally offered the lowest wages), there are some notable differences. For example, Minnesota offered a positive premium for workers with 12 years of education but a negative premium for workers with 8 years of education. Moreover, in Virginia the negative premium was four times larger for workers with 8 years of education than workers with 12 years of education. These patterns hold for other education levels (Appendix Table A.11).

Table 8 examines the impact of Social Security on children's direction of migration as measured by the difference in the education-specific earnings premium in their state of residence in 1950 and the corresponding premium in their parents' 1930 state of residence. The estimate in the first column implies that sons whose fathers were predicted to be eligible for Social Security at an earlier age moved to better-matched locations on average. Quantitatively, a 10-year earlier eligibility is associated with a 14-dollar increase in education-specific median earnings premium, relative to a baseline mean of 15.6 dollars (in 1949 dollars). The second and third columns examine other moments of the earnings distribution and reveal a similar positive effect on the 75th percentile but not the 25th percentile, which suggests that sons moved to places with more upside potential. The last three columns show analogous results based on moments of the overall earnings distribution by state. The lack of corresponding effects highlights that moving to a better-matched location does not necessarily entail moving to a place with higher earnings on average.[27] These results suggest that Social Security allowed sons to move to places where they had better labor market prospects,

[27] In separate analyses, we did not find any evidence that sons moved to more urban or developed areas (as measured by the share of manufacturing), suggesting that migration was primarily based on comparative advantage. This is similar to Nakamura et al. (2022), who find large migration gains for children induced to move by a volcanic eruption in Iceland, despite the fact that the affected location had high average incomes.

improving their career trajectory and enabling them to settle in better neighborhoods late in life.

5.4 Heterogeneity across families and children

Figure 5 shows heterogeneity in the impact of father's Social Security eligibility on children's late-life ZIP income rank and propensity to migrate across census divisions by select characteristics of families and children. These include race, family size, birth order, siblings' sex composition, proxies for family resources, and characteristics of parents' location. These results are obtained by estimating equation (3) separately for different subsets of families or children, which allows the controls to flexibly vary across subpopulations. Note that family size, birth order, and the sex composition of siblings are not directly observable and are approximated by pooling information from the 1920 Census, 1930 Census, and Numident records (see Appendix A.6 for details).

We caution against reading too much into any one of these results for two reasons: Partitioning the sample makes the individual subgroup comparisons imprecise, and the relevant prediction is not which families relied most on family old-age support but whose support was most responsive to Social Security, which is difficult to predict a priori. What is more important is the overall pattern, which is broadly consistent with our mechanism in two respects. First, the heterogeneity in migration effects generally aligns well with the heterogeneity in late-life neighborhood gains: Subgroups with larger migration responses also tend to experience larger neighborhood improvements, consistent with migration being an important channel. It is most pronounced for Black sons, for whom both effects are especially large, consistent with evidence that the returns to migration were particularly high for Black individuals in this time period (Collins and Wanamaker, 2014), given the severe location-specific labor market distortions they faced, including Jim Crow laws and occupational exclusion (Myrdal, 1944; Wright, 1986; Mohammed and Mohnen, 2025).[28]

Second, and more tentatively, the neighborhood gains tend to be larger in subgroups one might expect to rely more on the marginal support Social Security displaced: sons with fewer siblings, sons whose fathers held lower-paying occupations, and sons from states with less generous Old Age Assistance. The pattern is not uniform (homeowners and renters in 1930 were about equally affected), but given how hard it is to predict marginal support a priori, the broad alignment with priors is suggestive evidence of the key mechanism.

5.5 Additional outcomes

Appendix Table A.9 provides suggestive evidence that earlier eligibility increased sons' longevity. Because public Numident records mainly cover deaths that occurred between 1988 and 2007, we

[28]To increase sample sizes and precision, the results for Black sons and daughters are based on a slightly expanded sample that includes links with a precision rate of 90 percent or higher (rather than 97 percent).

follow Hollingsworth et al. (2024) and examine the impact of Social Security on mortality for a subset of cohorts for which we have good coverage of deaths for a common age window: the 1923–1930 birth cohorts for ages 65 to 77.[29] We find that Social Security had a positive and significant impact on sons' conditional age at death, with a 10-year earlier eligibility associated with a 0.47-year increase (but no impact on the probability of dying at ages 65 to 77). Consistent with our other results, we find no significant effects on daughters' long-run mortality. Given the incomplete coverage of Numident death records, it is unclear whether these patterns would hold if we could observe unconditional age at death for a broader set of cohorts, so we view these results as suggestive. Nevertheless, they are consistent with the notion that parents' Social Security eligibility had meaningful effects on sons' long-run socioeconomic status.

The first three columns of Appendix Table A.10 show the impact of fathers' Social Security eligibility on the probability of living with any parent, one's father, or one's mother in 1950. When pooling all children in panel (a), we find that father's earlier Social Security eligibility reduced the probability of co-residing with parents, consistent with our hypothesis that Social Security displaced in-kind transfers between generations. The estimates are of similar magnitude when looking at sons and daughters separately, but not statistically significant. Beyond smaller sample sizes, there are two reasons why co-residence effects may be hard to detect. First, here we are measuring co-residence from children's perspective. Since parents typically live with only one of their children at any given time, this makes it harder to identify changes in co-residence patterns. Second, parents from our cohorts of interest may no longer have been alive by 1950.

The last four columns of Appendix Table A.10 show that we find no evidence of impacts on either sons' or daughters' labor force participation, occupational standing, marital status, or fertility in 1950. The absence of early occupational effects is consistent with our findings for sons' first job in the OCG, suggesting that gains emerged gradually as sons climbed the job ladder in better-matched labor markets.

6 Implications

Our estimates imply substantial, long-lasting financial gains for recipients' children. We estimate sons' late-career earnings gains in three complementary ways (see Appendix E and Appendix Table A.14 for details of all calculations in this section). Translating occupational income score gains one-for-one into earnings gains yields a late-career gain of roughly 35 percent (SE $\approx$ 17 percent). As noted in Section 5.2, applying the cross-sectional relationships between late-career earnings and, respectively, late-life ZIP income and ZIP home prices yields earnings gains of

[29] Appendix Figure A.14 shows that while our Numident records mainly cover deaths between 1988 and 2007, they have good coverage of the entire ZIP AGI rank distribution. Intuitively, this is because longevity is only weakly correlated with late-life neighborhood quality (correlation of 0.09 in our sample).

roughly 28 percent (SE $\approx$ 10 percent) and 47 percent (SE $\approx$ 13 percent). The occupational income score approach is most direct but least precise, while the ZIP-based approaches are more precise but require an additional assumption to be unbiased (the bias likely understates the implied earnings gain; see Appendix D). We use the more conservative and precise ZIP income-based estimate of roughly 28 percent (SE $\approx$ 10 percent) as our baseline. Assuming no earnings gains through 1950 (based on the null effects on occupational income score at first job and in 1950), linear earnings growth to the late-career gain from 1950 to 1962, and constant earnings thereafter (based on the similar occupational income score gains in 1962 and 1972) yields a full-career average gain of 11.9 percent and an expected present value gain in 1950 of $130,000 per son in 2020 dollars, for sons born around 1910 (95% CI: [$45K, $215K]; see Table 9).[30] Because most affected children were married and the daughter estimates are small, we interpret the son estimate as approximately the effect on a married couple of both fathers' earlier eligibility.[31]

These gains were large relative to the associated Social Security benefits. We estimate that, for a father in the 1877 birth cohort, a 10-year difference in earliest eligibility age generated about $47,000 in expected lifetime benefits to the parents.[32] Scaling the per-son gain by the 2.3 sons of the average recipient family, children's total earnings gains were about 6.4 times the associated Social Security benefits under our preferred specification (95% CI = [2.2, 10.6]), and about 11 and 8 times under the alternative specifications (see Table 9). While there is uncertainty over the specific modeling choices in these calculations, children's gains were considerably larger than the associated benefits under a wide range of alternative assumptions (Appendix Table A.15), even with a conservative "first stage" of one (i.e., if differences in predicted eligibility translated one-for-one into differences in actual eligibility). This challenges the conventional view that these major Social Security expansions entailed a large redistribution from younger to older cohorts—the frequently-discussed "windfall to the initial old." The economic incidence of these early expansions was much less concentrated on the "initial old" than the statutory incidence.

The intergenerational gains also change the program's efficiency and fiscal implications. Under reasonable assumptions, these expansions reduced recipients' earnings by around 76 cents per

[30]This assumes that the baseline earnings path was median annual earnings for male four-quarter workers from the Social Security Bulletin Annual Statistical Supplements (Social Security Administration, 1965, 1975), with sons born in 1910, near the middle of the birth-year distribution for sons of fathers in the 1877 cohort. To calculate expected present values, we discount the implied earnings gains to reflect time discounting (3% annual real interest rate, discounted to 1950) and mortality (1910 cohort life table for males from the SSA, assuming sons lived to at least 1930 with certainty to match our sampling frame). The standard errors capture sampling uncertainty in the underlying estimates but not uncertainty from the interpolation and other modeling assumptions.

[31]Although we find no gains for daughters, this does not imply that women did not gain. If resources were fully pooled within households, women married to men would gain as much as their husbands.

[32]See Appendix E for details. Using the 1877 cohort is conservative for our key conclusions: The difference in Social Security benefits and the level of recipients' foregone earnings are near their maxima for this cohort, which tends to increase the government's net marginal cost and decrease children's gains relative to the associated benefits.

dollar of benefits. Our estimates that children's earnings increased by about \$6.4 imply a net increase in family earnings of about \$5.7 per dollar of benefits (95% CI: [\$1.4, \$9.9]). So whereas reducing labor supply is thought to be a key efficiency cost of Social Security (see Coile, 2015, for a review), these major Social Security expansions increased total family earnings substantially.[33] As a result, our point estimates imply that these early expansions more than paid for themselves, with a net cost per dollar of the mechanical benefit increase of $-$0.55: The cost to the government of the benefits and the foregone tax revenue from recipients was more than offset by the increased tax revenue from recipients' children (see Table 9). But the confidence interval ([$-$1.77, \$0.67]) includes modest positive costs and furthermore ignores modeling uncertainty; across the labor supply and policy variations in Appendix Table A.15, the point estimate of the net cost per dollar of the mechanical benefit increase ranges from $-$0.57 to \$0.02. The evidence is thus consistent with but does not conclusively establish self-financing.[34] Were it not for the intergenerational effects, by contrast, we estimate that the net cost per dollar of the mechanical benefit increase would have been about \$1.31.

7 Potential Mechanisms

The weight of the evidence points to migration to better-matched labor markets as the primary mechanism, facilitated by the displacement of family old-age support that required physical proximity. Increased financial transfers from parents to children are unlikely to account for more than a small share of the gains. Children's gains were several times the associated benefits, were driven by higher earnings, and were much larger for sons and for early migrants. These patterns are difficult to reconcile with financial transfers, which tend to be smaller than the associated benefits (e.g., Mukherjee, 2022), to decrease recipients' earnings (due to income effects of demand for leisure), and to be divided relatively equally across children (e.g., Menchik, 1980). We also find no evidence that schooling, veteran status, or unionization drove the gains (Appendix Table A.12).[35] A related

[33]Our estimates capture the partial-equilibrium effect of a son's own parental exposure to Social Security. They could overstate aggregate gains if treated sons displaced others. But added congestion in destinations is offset by reduced congestion in origins, so the reallocation roughly nets out. What matters is whether it improved the allocation of workers across locations, and the evidence that sons moved toward better-matched rather than uniformly higher-wage or more urban labor markets (Section 5.3, Table 8) points to comparative advantage rather than competition for a fixed set of jobs. Although we cannot directly estimate general-equilibrium spillovers, relatively elastic mid-century housing supply (Fetter, 2013; Glaeser, 2013; Lyons et al., 2026) would have limited the impact on house prices.

[34]As discussed in Appendix Section E and shown in Appendix Table A.15, the self-financing result is sensitive to the assumed average marginal tax rate on Social Security-induced earnings changes. Negative fiscal cost would imply an infinite marginal value of public funds (MVPF) *if* recipients and their children benefited from these expansions. But in addition to the uncertainty about the government's cost, it seems likely that at least some recipients or their children were harmed (e.g., a recipient for whom the reduction in family old-age support outweighed the increase in cash income from Social Security).

[35]The absence of schooling effects, which contrasts with past studies on the long-run intergenerational effects of cash transfers in this period (Aizer et al., 2016), reflects that recipients' children were adults when their parents received

possibility is that Social Security relaxed credit constraints. Yet benefits were modest and began only when parents reached age 65, by which point most sons were in their late thirties or forties. As we discuss in Section 7.1, eligibility more plausibly mattered as insurance—a guarantee that parents would have income until death—than as liquidity.

The migration evidence fits the location-constraint mechanism. As shown in Section 5.3, Social Security caused sons to substitute longer moves for shorter ones (Table 7), move to states offering higher education-specific earnings (Table 8), and move less after 1950 (Appendix Table A.8). The pattern of effects on the joint outcome of being in parents' 1930 census division in 1950 and dying in a ZIP code in each ZIP income quintile (Appendix Figure A.12) is consistent with Social Security enabling some sons to leave labor markets to which they were poorly matched. Together, these patterns suggest productive sorting rather than mere churning. Moreover, the heterogeneity patterns are broadly consistent with migration being a key driver of the gains: Migration and neighborhood gains broadly align across subgroups (Figure 5), and Social Security increased migration and late-life neighborhood quality for sons but not daughters.

What share of sons' gains might migration account for? Suppose Social Security meaningfully affected the migration of a share p of sons, each of whom realized a full-career earnings return R from that migration. Migration would then account for a share pR/G of sons' average full-career gain G. For G, we take our baseline full-career gain of 11.9 percent (Section 6). For R, we take Nakamura et al.'s (2022) estimated full-career gain of 82 percent for movers induced by a plausibly exogenous shock.[36] If the affected share were only the observed 2.5 percentage point increase in cross-division migration by 1950 ($p = 0.025$), migration would explain about 17 percent of sons' gains. A "cascade" reading of the migration-bin estimates implies an affected share of 6.3 percent, which would explain about 43 percent.[37] If the share whose migration timing, distance, or desti-

Social Security. While we note in Section 7.2 that World War II is important to the generalizability of our findings, the absence of an effect on veteran status is evidence that it does not affect the internal validity of our estimates.

[36] We use Nakamura et al. (2022) as a benchmark because it is among the few studies to recover a full-career earnings gain for induced movers, and because its migration responses, like ours, appear to reflect sorting on comparative advantage (and its estimated life cycle profile of back-loaded gains is quantitatively similar to ours). The analog closest to our mechanism—young adults induced to migrate to cities by a rural pension in China—yields a gain of 86 log points (about 136 percent) (Gai et al., 2025). Other quasi-experimental and structural work corroborates large returns for induced or marginal movers, though in different settings (e.g., Bryan et al., 2014; Bryan and Morten, 2019; Sarvimäki et al., 2022). Studies that instead address selection among average movers in contexts closer to ours point the same way: Collins and Wanamaker (2014), using linked census records, find that Black men who left the South earned 80–130 percent more than those who stayed, and Ward (2022), comparing migrant and non-migrant brothers, finds substantial but smaller returns. Large returns are unsurprising given the large regional and sectoral gaps in opportunity of the era (Caselli and Coleman, 2001; Boustan, 2016; Ganong and Shoag, 2017).

[37] Under a cascade interpretation, Social Security leads sons to shift their moves "one level up" the distance-bin hierarchy. Social Security caused 0.5pp more sons to leave their 1930 county by 1950. If all moved to a different county in the same state (the next "level"), then to match the 1.5pp decrease in different-county-same-state moves, Social Security must have caused 2pp of sons to switch from different-county-same-state moves to different-state-same-division moves, and so on.

nation was affected was 10 percent, migration would explain roughly 69 percent, and migration would fully account for sons' gains if the affected share was $p^* = G/R \approx 14.5$ percent. Our outcomes only capture migration at a few widely spaced points in time and therefore miss variation in the timing of migration, limiting what we can quantify unambiguously. Nevertheless, these calculations suggest that migration plausibly accounted for an important, and potentially central, share of sons' gains.

Migration was likely complemented by occupational upgrading and the accumulation of job-specific skills. Increasing returns to the utilization of human capital (Rosen, 1983) mean investments in one dimension (such as migration) raise the returns to investments in others (such as occupational choice or on-the-job training). The absence of occupational gains at sons' first job or in 1950, followed by large gains in 1962 and 1972, is consistent with gradual advancement in better-matched labor markets rather than a one-time wage premium upon arrival. We therefore view migration and complementary career investments as closely connected rather than separately identifiable channels.

7.1 Why didn't families do this even without Social Security?

If the gains from migration were so large, why didn't families capture them on their own—relocating the children whose local labor markets were the poorest match, while still ensuring that elderly parents were supported? We hypothesize that indivisibilities and missing markets concentrated old-age-support costs on one or a few children, sometimes causing a child to remain near their parents despite a poor labor market match.

One set of obstacles concerns sharing support costs across children. Support through shared housing or a shared location is largely indivisible so its cost falls heavily on whichever child stays near the parents. Spreading that cost more evenly across siblings, or across families, would have required forms of contracting and insurance that were largely absent. Siblings could not easily commit in advance to compensate the one who remained, and no market existed to insure or trade the idiosyncratic risk of how much support a given parent would ultimately require.

A second set of obstacles arises from the nature of migration and its economic returns. The returns to migration are typically back-loaded (see, e.g., Nakamura et al., 2022), so much of a son's gain would have accrued only after his parents had died—too late, given imperfect capital markets, to support his parents. Those returns were also uncertain and difficult to insure (Munshi and Rosenzweig, 2016; Meghir et al., 2022). Even if migration paid off handsomely on average, it might not pay off enough to support a parent in every state of the world. And even when a son did capture higher earnings, he could not readily commit to compensate his parents for the in-person support he otherwise would have provided (a limited-commitment problem of the kind Bau et al., 2026, document in the context of India).

In this view, Social Security substituted for these missing markets. By providing a guaranteed income stream financed by a broad payroll tax, it spread old-age-support costs across workers and over time. It also provided insurance that parents would receive income until death, reducing the risk that a son who moved would later need to return or send large transfers. Social Security thereby reduced the burden on would-be primary providers of old-age support, lowering the benefits of living near their parents and hence their opportunity cost of migration. Hence, it relaxed, at least for some children, a "location constraint" that would have tied them to their parents' location. That these expansions significantly increased the total family earnings of recipients and their children suggests that the associated efficiency gain may have been substantial.

Two further factors, distinct from missing markets, may also have limited families' ability to capture the gains. First, families may simply have underestimated the gains. Quasi-experimental and structural studies in many settings find that those induced to migrate realize large gains (e.g., Bryan et al., 2014; Nakamura et al., 2022; Lagakos et al., 2023; Gai et al., 2025), potentially beyond what families anticipated. Second, social norms—such as an expectation that at least one child remain near the parents—may have deterred families from exploiting potential gains from trade. Norms that promoted efficiency in the past need not continue to do so when circumstances change (Bau, 2021).

We stress that these explanations are speculative. The source of the friction does not affect our estimates of the earnings and fiscal effects of these expansions, which are invariant to why families did not capture the gains on their own. But it does bear on the welfare effects—substituting for a missing market, correcting mistaken beliefs, and overriding norms or preferences may not have the same welfare consequences—and on the generalizability to other settings.

7.2 Generalizability

Several features of our setting were especially conducive to large intergenerational effects. Government old-age support was initially minimal, while family support was widespread (with high rates of co-residence, which we posit was relatively substitutable with cash benefits), so the expansions we study could displace family support on a scale that subsequent expansions might be unlikely to match. In addition, relative to more recent periods, the cohorts we study came of age at a time of unusually heterogeneous local labor markets and correspondingly high returns to leaving poorly matched places (e.g., Caselli and Coleman, 2001; Collins and Wanamaker, 2014, 2015; Boustan, 2016). The significant migration associated with World War II and the Second Great Migration may have led people to be more responsive to shocks than they otherwise would have been. Our estimates may therefore be larger than those in many other settings. In particular, the favorable fiscal effects we estimate should not be extrapolated to marginal expansions in settings with substantial preexisting coverage, including the present-day US, where the large regional wage gaps that made migration so valuable have since compressed (Ganong and Shoag, 2017).

Even so, the underlying mechanism can operate wherever family support constrains where or how much adult children work and can be displaced by government support. It may be especially relevant to long-term care, much of which continues to be provided by adult children (Barczyk and Kredler, 2018). The short-run version of this mechanism is visible in Coe et al.'s (2023) finding that subsidies for parents' private long-term care insurance reduce children's co-residence and geographic proximity while increasing their full-time employment. Our findings suggest that such immediate adjustments may have important longer-run consequences: By relaxing constraints on children's location and labor supply, old-age support can affect migration and earnings over substantially longer horizons. Evaluations of contemporaneous caregiving and labor-supply responses may therefore miss an important component of these policies' intergenerational effects.

8 Conclusion

A full accounting of Social Security's effects requires considering not only direct recipients but also family members whose support obligations it may change. We study these intergenerational effects by exploiting variation created by Social Security's initially incomplete coverage and subsequent expansions, with a newly-built dataset linking parents to the long-run outcomes of their children. We find that greater parental exposure to Social Security increased sons' geographic mobility, earnings, and late-life neighborhood quality. We find no comparable effects for daughters, consistent with historical and contemporary evidence that daughters disproportionately provided hands-on care that cash benefits could not readily replace.

Several pieces of evidence point to migration to better-matched labor markets as a central mechanism. By displacing certain forms of old-age support that required physical proximity, Social Security appears to have relaxed geographic constraints on sons, allowing them to move to locations offering higher returns to their skills. The resulting gains to recipients' children appear to have been considerably larger than the gains to recipients themselves.

Our findings carry significant implications for the distributional, efficiency, and fiscal characteristics of the early Social Security program. It bears emphasis that several features of our historical setting were especially conducive to large intergenerational effects. At the same time, the underlying mechanism likely operates wherever family support constrains location or labor supply and can be replaced by government support. In particular, it may be an important consideration for contemporary policies such as long-term care, where families still provide the bulk of support, and where estimates of short-run responses may not fully capture long-term effects.

References

Abramitzky, Ran, Leah Boustan, Katherine Eriksson, James Feigenbaum, and Santiago Pérez, "Automated Linking of Historical Data," *Journal of Economic Literature*, 2021, *59* (3), 865–918.

_ , _ , _ , Myera Rashid, and Santiago Pérez, *Census Linking Project: Version 3.0 [dataset]*, Harvard Dataverse, 2025.

Aizer, Anna, Gabrielle Grafton, and Santiago Pérez, "Daughters as Safety Net? Family Response to Parental Employment Shocks: Evidence from Alcohol Prohibition," NBER Working Paper 33346, National Bureau of Economic Research 2025.

_ , Shari Eli, Joseph Ferrie, and Adriana Lleras-Muney, "The Long-Run Impact of Cash Transfers to Poor Families," *American Economic Review*, 2016, *106* (4), 935–971.

Althoff, Lukas and Hugo Reichardt, "Jim Crow and Black Economic Progress After Slavery," *The Quarterly Journal of Economics*, 2024, *139* (4), 2279–2330.

_ , Harriet Brookes Gray, and Hugo Reichardt, "Mothers, Schools, and the Making of American Human Capital Mobility," Working paper 2025.

Altmeyer, Arthur J., "The Desirability of Extending Social Security to Employees of Nonprofit Institutions," *Social Security Bulletin*, 1944, *7* (8).

Ardington, Cally, Anne Case, and Victoria Hosegood, "Labor Supply Responses to Large Social Transfers: Longitudinal Evidence from South Africa," *American Economic Journal: Applied Economics*, 2009, *1* (1), 22–48.

Auerbach, Alan J. and Laurence J. Kotlikoff, *Dynamic Fiscal Policy*, Cambridge, MA: Cambridge University Press, 1987.

Bailey, Martha J., A. R. Shariq Mohammed, and Paul Mohnen, "The Evolution of U.S. Educational Mobility over the 20th Century and the Role of Public Education," Atlanta Fed Working Paper 2026-1 2026.

_ , Connor Cole, and Catherine Massey, "Simple Strategies for Improving Inference with Linked Data: A Case Study of the 1850–1930 IPUMS Linked Representative Historical Samples," *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 2020, *53* (2), 80–93.

_ , Hilary W. Hoynes, Maya Rossin-Slater, and Reed Walker, "Is the Social Safety Net a Long-Term Investment? Large-Scale Evidence from the Food Stamps Program," *Review of Economic Studies*, 2024, *91* (3), 1291–1330.

_ , Peter Z. Lin, A. R. Shariq Mohammed, Paul Mohnen, Jared Murray, Mengying Zhang, and Alexa Prettyman, "The creation of LIFE-M: The Longitudinal, Intergenerational, Family Electronic Micro-Database project," *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 2023, *56* (3), 138–159.

_ , _ , A.R. Shariq Mohammed, Paul Mohnen, Jared Murray, Mengying Zhang, and Alexa Prettyman, *LIFE-M: The Longitudinal, Intergenerational Family Electronic Micro-Database*, Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2022.

Barczyk, Daniel and Matthias Kredler, "Evaluating Long-Term-Care Policy Options, Taking the Family Seriously," *The Review of Economic Studies*, 2018, *85* (2), 766–809.

Barro, Robert J., "Are Government Bonds Net Wealth?," *Journal of Political Economy*, 1974, *82* (6), 1095–1117.

Bau, Natalie, “Can Policy Change Culture? Government Pension Plans and Traditional Kinship Practices,” *American Economic Review*, 2021, *111* (6), 1880–1917.

__ **, Gaurav Khanna, Corinne Low, and Alessandra Voena**, “Traditional Institutions in Modern Times: Dowries as Pensions when Sons Migrate,” *Quarterly Journal of Economics*, 2026, *141* (1), 205–262.

Becker, Gary S., “A Theory of Social Interactions,” *Journal of Political Economy*, 1974, *82* (6), 1063–1093.

Blau, Peter M., Otis Dudley Duncan, David L. Featherman, and Robert M. Hauser, *Occupational Changes in a Generation, 1962 and 1973 (ICPSR 6162)*, Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1994.

Boustan, Leah Platt, *Competition in the Promised Land: Black Migrants in Northern Cities and Labor Markets*, Princeton, NJ: Princeton University Press, 2016.

Bryan, Gharad and Melanie Morten, “The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia,” *Journal of Political Economy*, 2019, *127* (5), 2229–2268.

__ **, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak**, “Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh,” *Econometrica*, 2014, *82* (5), 1671–1748.

Carter, Susan B., Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, Gavin Wright, Price V. Fishback, and Melissa A. Thomasson, *Historical Statistics of the United States Millennial Edition Online*, New York, NY: Cambridge University Press, 2006.

Caselli, Francesco and Wilbur John Coleman II, “The U.S. Structural Transformation and Regional Convergence: A Reinterpretation,” *Journal of Political Economy*, 2001, *109* (3), 584–616.

Chetty, Raj, John N. Friedman, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter, “The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility,” *American Economic Review*, 2026, *116* (1), 1–51.

Clark, Robert L., Lee A. Craig, and John Sabelhaus, *State and Local Retirement Plans in the United States*, Edward Elgar, 2011.

Coe, Norma B., Gopi Shah Goda, and Courtney Harold Van Houtven, “Family spillovers and long-term care insurance,” *Journal of Health Economics*, 2023, *90.*

Coile, Courtney C., “Economic Determinants of Workers’ Retirement Decisions,” *Journal of Economic Surveys*, 2015, *29* (4), 830–853.

Collins, William J. and Marianne H. Wanamaker, “Selection and Economic Gains in the Great Migration of African Americans: New Evidence from Linked Census Data,” *American Economic Journal: Applied Economics*, 2014, *6* (1), 220–252.

__ **and** __ , “The Great Migration in Black and White: New Evidence on the Selection and Sorting of Southern Migrants,” *The Journal of Economic History*, 2015, *75* (4), 947–992.

__ **and** __ , “African American Intergenerational Economic Mobility since 1880,” *American Economic Journal: Applied Economics*, 2022, *14* (3), 84–117.

Costa, Dora L., “Displacing the Family: Union Army Pensions and Elderly Living Arrangements,” *Journal of Political Economy*, 1997, *105* (6), 1269–1292.

__ , *The Evolution of Retirement* National Bureau of Economic Research Books, University of Chicago Press, 1998.

__ , “A House of Her Own: Old Age Assistance and the Living Arrangements of Older Nonmarried Women,” *Journal of Public Economics*, 1999, *72* (1), 39–59.

Coyne, David, Itzik Fadlon, Shanthi P. Ramnath, and Patricia K. Tong, "Household Labor Supply and the Value of Social Security Survivors Benefits," *American Economic Review*, 2024, *114* (5), 1248—-1280.

Diamond, Peter A., "National Debt in a Neoclassical Growth Model," *American Economic Review*, 1965, *55* (5), 1126–1150.

Eckert, Fabian, Andrés Gvirtz, Jack Liang, and Michael Peters, "A Method to Construct Geographical Crosswalks with an Application to US Counties since 1790," NBER Working Paper 26770, National Bureau of Economic Research 2020.

Edmonds, Eric V., Kristin Mammen, and Douglas L. Miller, "Rearranging the Family?: Income Support and Elderly Living Arrangements in a Low-Income Country," *Journal of Human Resources*, 2005, *40* (1), 186–207.

Engelhardt, Gary V., Jonathan Gruber, and Anil Kumar, "Early Social Security Claiming and Old-Age Poverty: Evidence from the Introduction of the Social Security Early Eligibility Age," *Journal of Human Resources*, 2022, *57* (4), 1079–1106.

_ , _ , and Cynthia D. Perry, "Social Security and Elderly Living Arrangements Evidence from the Social Security Notch," *Journal of Human Resources*, 2005, *40* (2), 354–372.

Feldstein, Martin, "Social Security, Induced Retirement, and Aggregate Capital Accumulation," *Journal of Political Economy*, 1974, *82* (5), 905–926.

_ and Jeffrey B. Liebman, "Social Security," in A. J. Auerbach and M. Feldstein, eds., *Handbook of Public Economics*, Vol. 4 of *Handbook of Public Economics*, Elsevier, 2002, chapter 32, pp. 2245–2324.

Fetter, Daniel K., "How Do Mortgage Subsidies Affect Home Ownership? Evidence from the Mid-century GI Bills," *American Economic Journal: Economic Policy*, May 2013, *5* (2), 111–147.

_ and Lee M. Lockwood, "Government Old-Age Support and Labor Supply: Evidence from the Old Age Assistance Program," *American Economic Review*, 2018, *108* (8), 2174–2211.

Fetter, Daniel, Lee Lockwood, and Paul Mohnen, "The Intergenerational Incidence of Government Old-Age Support: Evidence from the Early Social Security Era," in Mohnen, Paul Hong-Lum, *Essays in Labor Economics*, Northwestern University PhD Dissertation 2018.

Fitzpatrick, Maria and Timothy Moore, "The mortality effects of retirement: Evidence from Social Security eligibility at age 62," *Journal of Public Economics*, 2018, *157*, 121–137.

Gai, Qingen, Naijia Guo, Bingjing Li, Qinghua Shi, and Xiaodong Zhu, "Rural Pensions, Labor Reallocation, and Aggregate Income: An Empirical and Quantitative Analysis of China," *Econometrica*, 2025, *93* (5), 1663–1696.

Ganong, Peter and Daniel Shoag, "Why has regional income convergence in the U.S. declined?," *Journal of Urban Economics*, 2017, *102* (C), 76–90.

Geanakoplos, John, Olivia S. Mitchell, and Stephen P. Zeldes, "Social Security's Money's Worth," in Olivia S. Mitchell, Robert J. Myers, and Howard Young, eds., *Prospects for Social Security Reform*, University of Pennsylvania Press, 1998, pp. 79–151.

Gelber, Alexander, Adam Isen, and Jae Song, "The Effect of Pension Income on Elderly Earnings: Evidence from Social Security and Full Population Data," Working paper 2016.

Glaeser, Edward L., "A Nation of Gamblers: Real Estate Speculation and American History," *American Economic Review*, 2013, *103* (3), 1–42.

Goldstein, Joshua and Casey Breen, "Berkeley Unified Numident Mortality Database: Public administrative records for individual-level mortality research," *Demographic Research*, 2022, *47* (5), 111–142.

Goldstein, Joshua R., Monica Alexander, Casey Breen, Andrea Miranda González, Felipe Menares, Maria Osborne, Mallika Snyder, and Ugur Yildirim, *CenSoc Mortality File: Version 3.0*, Berkeley, CA: University of California, 2023.

Gratton, Brian, "The Poverty of Impoverishment Theory: The Economic Well-Being of the Elderly, 1890–1950," *The Journal of Economic History*, 1996, *56* (1), 39–61.

Green, Allison, "Networks and Geographic Mobility: Evidence from World War II Navy Ships," Working paper 2024.

Grigoryeva, Angelina, "Own Gender, Sibling's Gender, Parent's Gender: The Division of Elderly Parent Care among Adult Children," *American Sociological Review*, 2017, *82* (1), 116–146.

Grove, Robert D. and Alice M. Hetzel, "Vital Statistics Rates in the United States 1940–1960," Technical Report, Public Health Service, National Center for Health Statistics 1968.

Gruber, Jonathan, *Public Finance and Public Policy*, 4th ed., Macmillan, 2013.

Hayhow, Edgar C., "Pensions and Life Insurance Plans for Hospital Employees," *Hospitals: the Journal of the American Hospital Association*, 1940, *14* (11), 73–77.

Hollingsworth, Alex, Krzysztof Karbownik, Melissa A. Thomasson, and Anthony Wray, "The Gift of a Lifetime: The Hospital, Modern Medicine, and Mortality," *American Economic Review*, 2024, *114* (7), 2201–2238.

Horowitz, Amy, "Sons and Daughters as Caregivers to Older Parents: Differences in Role Performance and Consequences," *The Gerontologist*, 1985, *25* (6), 612–617.

Huang, Wei and Chuanchuan Zhang, "The Power of Social Pensions: Evidence from China's New Rural Pension Scheme," *American Economic Journal: Applied Economics*, 2021, *13* (2), 179–205.

Jensen, Robert T., "Do private transfers 'displace' the benefits of public transfers? Evidence from South Africa," *Journal of Public Economics*, 2004, *88* (1-2), 89–112.

Kestenbaum, Bert, "State and Local Government Employees Covered Under Social Security, 1977–81," *Social Security Bulletin*, 1982, *45* (12), 11–13.

Ko, Ami, "An Equilibrium Analysis of the Long-Term Care Insurance Market," *Review of Economics Studies*, 2022, *89* (4), 1993–2025.

Krueger, Alan B. and Jörn-Steffen Pischke, "The Effect of Social Security on Labor Supply: A Cohort Analysis of the Notch Generation," *Journal of Labor Economics*, 1992, *10* (4), 412–437.

Lagakos, David, Ahmed Mushfiq Mobarak, and Michael E. Waugh, "The Welfare Effects of Encouraging Rural–Urban Migration," *Econometrica*, 2023, *91* (3), 803–837.

Leimer, Dean R., "Cohort-Sepcific Measures of Lifetime Net Social Security Transfers," Technical Report, Social Security Administration 1994. Office of Research and Statistics Working Paper Number 59.

Lyons, Ronan C., Allison Shertzer, Rowena Gray, and David Agorastos, "The Price of Housing in the United States, 1890–2006," *The Quarterly Journal of Economics*, 2026, *141* (1), 559–603.

Mastrobuoni, Giovanni, "Labor supply effects of the recent social security benefit cuts: Empirical estimates using cohort discontinuities," *Journal of Public Economics*, 2009, *93* (11–12), 1224–1233.

McCamman, Dorothy, "Higher Educational Institutions and Social Security," *Social Security Bulletin*, 1948, *11* (1), 5–11.

_, "Retirement Protection for State and Local Employees: Ten Years of Growth," *Social Security Bulletin*, 1953, *16* (5), 3–10,24.

McGarry, Kathleen and Robert F. Schoeni, "Social Security, Economic Growth, and the Rise in Elderly Widows' Independence in the Twentieth Century," *Demography*, 2000, *37* (2), 221–236.

Meghir, Costas, A Mushfiq Mobarak, Corina Mommaerts, and Melanie Morten, "Migration and Informal Insurance: Evidence from a Randomized Controlled Trial and a Structural Model," *The Review of Economic Studies*, 2022, *89* (1), 452–480.

Menchik, Paul L., "Primogeniture, Equal Sharing, and the U.S. Distribution of Wealth," *The Quarterly Journal of Economics*, 1980, *94* (2), 299–316.

Moffitt, Robert A., "Trends in Social Security Wealth by Cohort," in "Economic Transfers in the United States" NBER Chapters, National Bureau of Economic Research, Inc, 1984, pp. 327–358.

Mohammed, A. R. Shariq and Paul Mohnen, "Black Economic Progress in the Jim Crow South: Evidence from Rosenwald Schools," *Journal of Economic History*, 2025, *85* (1), 72–109.

Mommaerts, Corina, "Long-Term Care Insurance and the Family," *Journal of Political Economy*, 2025, *133* (1), 1–52.

Mukherjee, Anita, "Intergenerational Altruism and Retirement Transfers: Evidence from the Social Security Notch," *Journal of Human Resources*, 2022, *57* (5), 1466–1497.

Munshi, Kaivan and Mark Rosenzweig, "Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap," *American Economic Review*, 2016, *106* (1), 46–98.

Myrdal, Gunnar, *An American Dilemma: The Negro Problem and Modern Democracy*, New York, NY: Harper & Brothers Publishers, 1944.

Nakamura, Emi, Jósef Sigurdsson, and Jón Steinsson, "The Gift of Moving: Intergenerational Consequences of a Mobility Shock," *The Review of Economic Studies*, 2022, *89* (3), 1557–1592.

Parker, Florence E., "Experience Under State Old-Age Pension Acts in 1935," *Monthly Labor Review*, 1936, *34* (4), 811–837.

Rosen, Sherwin, "Specialization and Human Capital," *Journal of Labor Economics*, 1983, *1* (1), 43–49.

Ruggles, Steven, "The Decline of Intergenerational Coresidence in the United States, 1850 to 2000," *American Sociological Review*, 2007, *72* (6), 964–989.

_ , Matt A. Nelson, Matthew Sobek, Catherine A. Fitch, Ronald Goeken, J. David Hacker, Evan Roberts, and J. Robert Warren, *IPUMS Ancestry Full Count Data: Version 4.0 [dataset]*, Minneapolis, MN: IPUMS, 2024.

_ , Nesile Ozder, Catherine A. Fitch, Matthew Sobek, Julia A. Rivera Drew, J. David Hacker, Jonas Helgertz, Cheyenne Lonobile, Matt A. Nelson, Evan Roberts, and John Robert Warren, *IPUMS Multigenerational Longitudinal Panel: Version 2.0 [dataset]*, Minneapolis, MN: IPUMS, 2025.

_ , Sarah Flood, Matthew Sobek, Daniel Backman, Grace Cooper, Julia A. Rivera Drew, Stephanie Richards, Renae Rogers, Jonathan Schroeder, and Kari C.W. Williams, *IPUMS USA: Version 16.0 [dataset]*, Minneapolis, MN: IPUMS, 2025.

Samuelson, Paul A., "An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money," *Journal of Political Economy*, 1958, *66*, 467–467.

Sarvimäki, Matti, Roope Uusitalo, and Markus Jäntti, "Habit Formation and the Misallocation of Labor: Evidence from Forced Migrations," *Journal of the European Economic Association*, 2022, *20* (6), 2497–2539.

Schieber, Sylvester J. and John B. Shoven, *The Real Deal: The History and Future of Social Security*, New Haven, CT: Yale University Press, 1999.

Schorr, Alvin Louis, *Filial responsibility in the modern American family: an evaluation of current practice of filial responsibility in the United States and the relationship to it of social security programs*, Washington, D.C.: U.S. Department of Health, Education, and Welfare, Social Security Administration, Division of Program Research., 1960.

Schroeder, Jonathan, David Van Riper, Steven Manson, Katherine Knowles, Tracy Kugler, Finn Roberts, and Steven Ruggles, *IPUMS National Historical Geographic Information System: Version 20.0 [dataset]*, Minneapolis, MN: IPUMS, 2025.

Shanas, Ethel, *Family Relationships of Older People: Living Arrangements, Health Status, and Family Ties of Those Aged 65 and Over, As Reportd by the Aged, the Persons to Whom They Would Turn in a Health Crisis, and the General Public*, New York, NY: Health Information Foundation, 1961.

Sharp, Harry and Morris Axelrod, "Mutual Aid Among Relatives in an Urban Population," in Ronald et al. Freedman, ed., *Principles of Sociology: A Text with Readings*, New York, NY: Henry Holt and Company, 1956, pp. 433–439.

Social Security Administration, "Social Security Bulletin, Annual Statistical Supplement, 1965," 1965. U.S. Government Printing Office.

_ , "Social Security Bulletin, Annual Statistical Supplement, 1975," 1975. U.S. Government Printing Office.

_ , *Current Population Survey, 1973, and Social Security Records: Exact Match Data (ICPSR 7616)*, Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2005.

_ , *Social Security Longitudinal Earnings Public Use File, 1937-1975 (ICPSR 7617)*, Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2008.

SSA Office of the Chief Actuary, "OASDI Trustees Report," 2013.

Steiner, Peter Otto and Robert Dorfman, *The Economic Status of the Aged*, Berkeley, CA: University of California Press, 1957.

Streib, Gordon F., "Family Patterns in Retirement," *Journal of Social Issues*, 1958, *14* (2), 46–60.

Subcommittee on Social Security, "Background Material on Social Security Coverage of Governmental Employees and Employees of Nonprofit Organizations," in "Coverage and Termination of Coverage of Government and Nonprofit Organization Employees under the Social Security System: Public Hearings before the Subcommittee on Social Security of the Committee on Ways and Means, U.S. House of Representatives, Ninety-Fourth Congress, Second Session," U.S. Government Printing Office, 1976.

Sussman, Marvin B., "The Help Pattern in the Middle Class Family," *American Sociological Review*, 1953, *18* (1), 22–28.

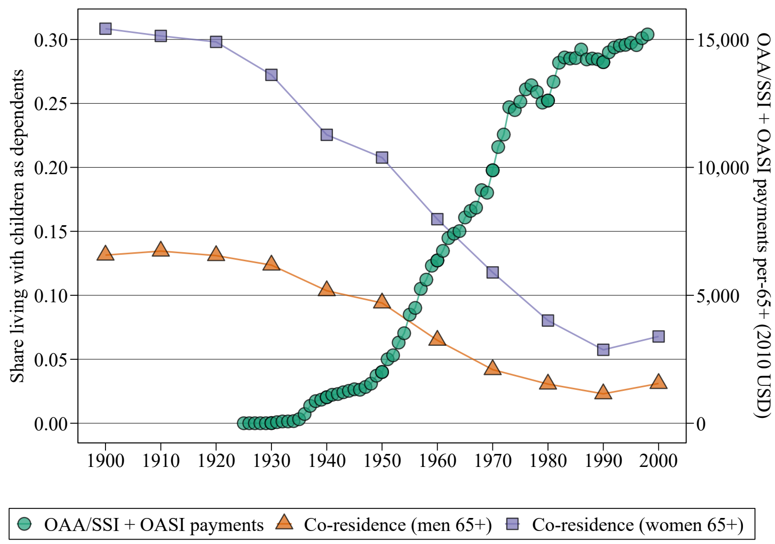
_ , "Relationships of Adult Children with their Parents in the United States," in Ethel Shanas and Gordon F. Streib, eds., *Social Structure and the Family: Generational Relations*, Englewood Cliffs, NJ: Prentice-Hall, Inc., 1965, pp. 62–92.

United States Bureau of the Census, *Retirement Systems for State and Local Government Employees: 1941*, U.S. Department of Commerce, Bureau of the Census, 1943.

Ward, Zachary, "Internal Migration, Education, and Intergenerational Mobility: Evidence from American History," *Journal of Human Resources*, 2022, *57* (6), 1981–2011.

Wright, Gavin, *Old South, New South: Revolutions in the Southern Economy Since the Civil War*, New York, NY: Basic Books, 1986.

Figure 1: Intergenerational co-residence over the 20th century

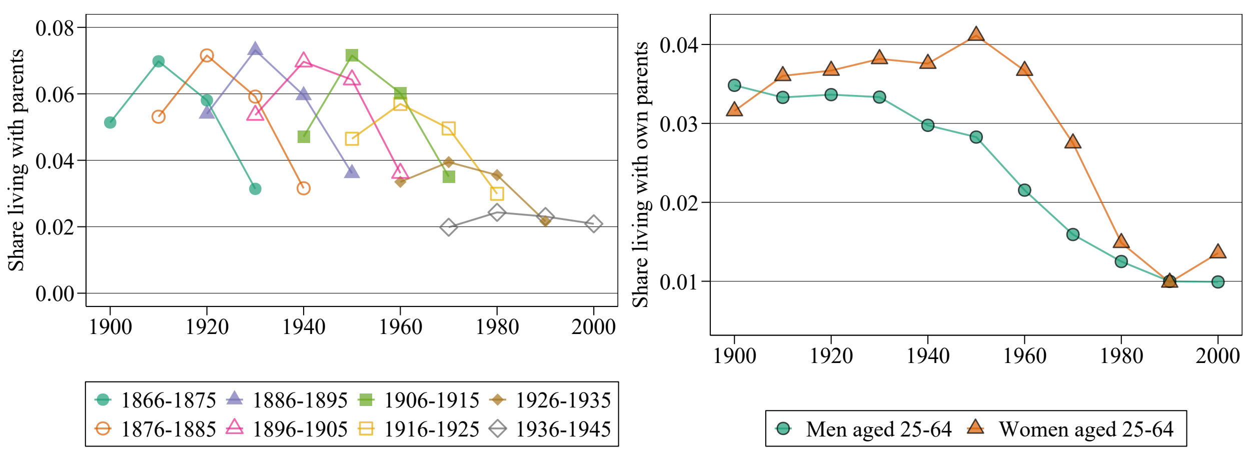


Notes: This figure plots the share of men and women aged 65 or older living with one of their children or children-in-law, not as a household head or spouse of the household head. Also displayed are total payments under Old Age Assistance (OAA) and Old Age and Survivors Insurance (OASI) divided by the 65+ population (in 2010 dollars). OAA payments data come from Parker (1936) for 1925 to 1935 and Series Bf621 of Carter et al. (2006) for 1936 onwards. SSI data come from Series Bf609 of Carter et al. (2006). OASI payments data come from Series BF396 of Carter et al. (2006). Data on co-residence come from Ruggles et al. (2025b).

Figure 2: Co-residence with parents from children's perspective

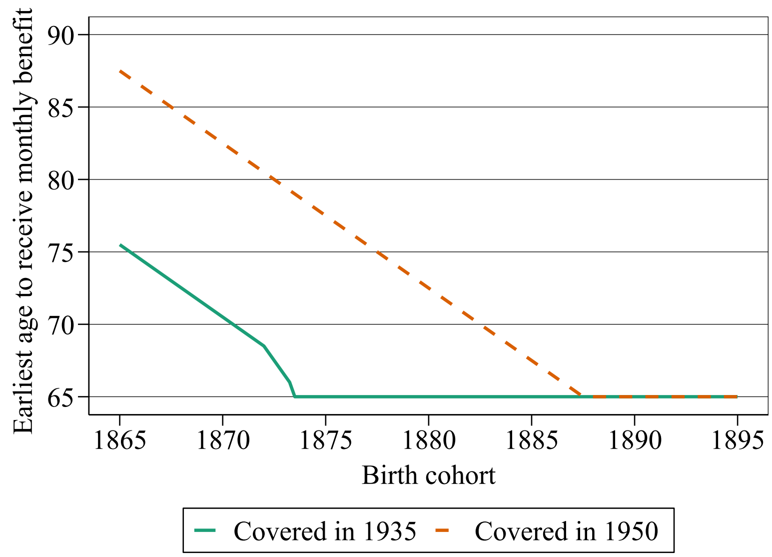
(a) Co-residence with parents, by birth cohort

(b) Co-residence with own parents, by sex



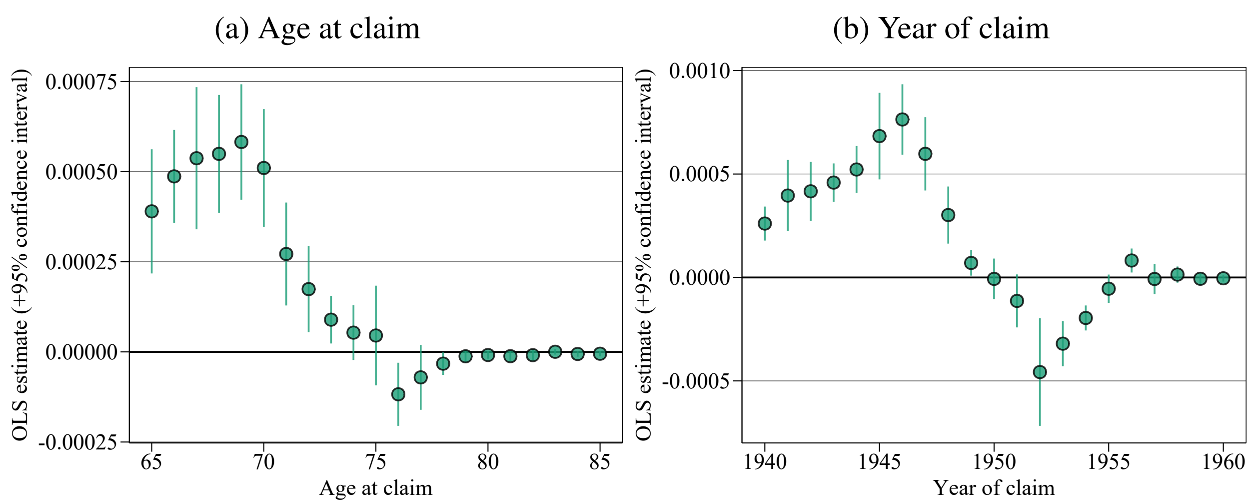
Notes: Panel (a) shows the share of US-born men and women who are household head or spouse of the head and co-reside with a parent or parent-in-law in a specified year, separately by 10-year birth cohort. Each cohort is shown at ages 25–34, 35–44, 45–54, and 55–64. Panel (b) shows the share of US-born men and women aged 25–54 who are household head or spouse of the head and co-reside with their own parents in a specified year. Data come from Ruggles et al. (2025b).

Figure 3: Earliest age to receive monthly Social Security benefits



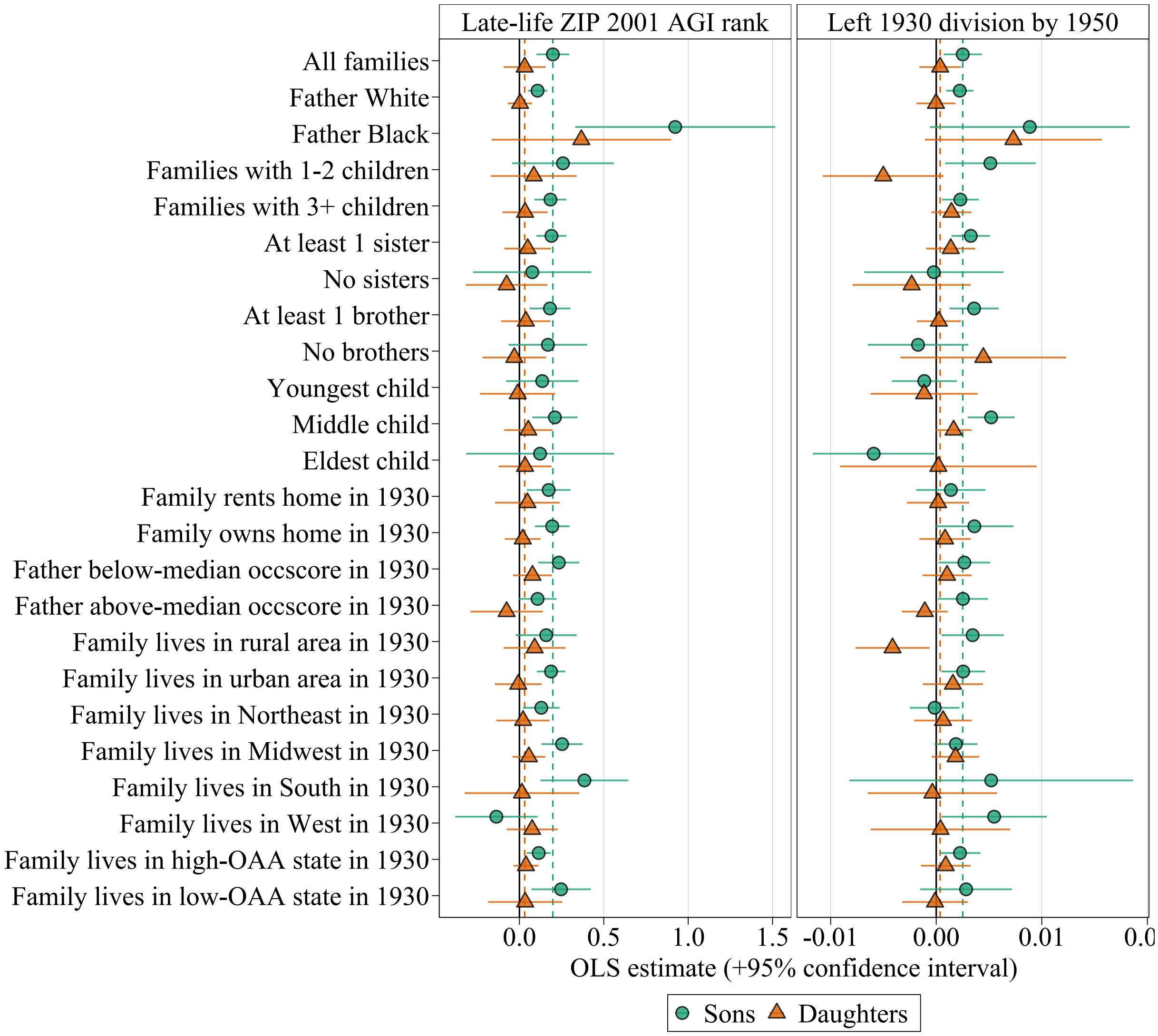
Notes: Minimum ages are based on Social Security Act of 1935, 1939 Amendments, and 1950 Amendments, assuming continuous employment in 1935 or 1950-covered employment. The minimum ages shown are those under the 1950 Amendments.

Figure 4: The impact of exposure to Social Security on age at claim and year of claim



Notes: Panel (a) shows the coefficients on Social Security exposure (i.e., negative of earliest age of Social Security eligibility) from equation (2), where the sample consists of all men in our population of interest in the 1930 Census and the dependent variable is an indicator for appearing in the Numident claims-census linked sample and claiming at a specified age. Panel (b) plots analogous estimates for claiming in a specified year. All regressions include fixed effects for men's 3-digit occupation-by-birth cohort, Social Security coverage category, race, and birthplace. 95% confidence intervals based on robust standard errors clustered by men's 3-digit industry.

Figure 5: The impact of father's exposure to Social Security on children's late-life ZIP rank in terms of 2001 AGI and propensity to live outside their parents' 1930 census division of residence by 1950: Heterogeneity by select characteristics of families and children



Notes: This figure shows the coefficients on father's Social Security exposure (i.e., negative of earliest age of Social Security eligibility) from equation (3), where the sample consists of children in our Numident death-census linked sample and the outcome is either the percentile rank of a child's late-life ZIP code in terms of mean AGI (left panel) or an indicator for living outside parents' 1930 census division of residence by 1950 (right panel). Black sample based on 90 percent precision links. Sample restricted to families with at least two children (born in different years) for sibling sex composition (birth order) heterogeneity cuts. High and low-OAA states defined as states with above or below-median OAA payments per person-65+ in December 1939 (Fetter and Lockwood, 2018). Colors/shapes indicate separate estimates for sons and daughters. All regressions include fixed effects for children's sex-by-birth cohort and fathers' 3-digit occupation-by-birth cohort, Social Security coverage category, race, and birthplace. Regressions for sibling sex composition heterogeneity cuts additionally include fixed effects for the number of children in the family. Observations weighted using inverse propensity score weights that re-weight the sample to match the population of interest in 1930. 95% confidence intervals based on robust standard errors clustered by fathers' 3-digit industry.

Table 1: The impact of exposure to Social Security on subsequent claiming behavior

	Dependent variable:	
	Linked to claims data	Age at claim
	(1)	(2)
Social Security exposure	0.0034***	-0.1143***
	(0.0005)	(0.0162)
Mean of dep. variable	0.093	67.5
N	3,920,008	367,732

Notes: This table shows the coefficients on Social Security exposure (i.e., negative of earliest age of Social Security eligibility) from equation (2), where the sample consists of all men in our population of interest in the 1930 Census in column 1 and the subset of those men appearing in our Numident claims-census linked sample in column 2. In column 1, the outcome is an indicator for appearing in the Numident claims-census linked sample. In column 2, the outcome is age at the time of claiming. All regressions include fixed effects for men's 3-digit occupation-by-birth cohort, Social Security coverage category, race, and birthplace. In column 2, observations are weighted using inverse propensity score weights that re-weight the sample to match the population of interest in 1930. Robust standard errors in parentheses, clustered by men's 3-digit industry. *** 1%, ** 5%, * 10% significance.

Table 2: Selection of 1930 households and children into Numident death-census linked sample

	Dependent variable: Any/number of children linked					
	All children		Sons		Daughters	
	Any	Number	Any	Number	Any	Number
	(1)	(2)	(3)	(4)	(5)	(6)
Social Security exposure	-0.0001	0.0003	0.0001	0.0004	-0.0002	-0.0001
	(0.0003)	(0.0007)	(0.0003)	(0.0005)	(0.0002)	(0.0003)
Mean of dep. variable	0.219	0.403	0.152	0.215	0.135	0.189
N	3,920,008	3,920,008	3,920,008	3,920,008	3,920,008	3,920,008

Notes: This table shows the coefficients on Social Security exposure (i.e., negative of earliest age of Social Security eligibility) from equation (2), where the sample consists of all men in our population of interest in the 1930 Census and the outcome is either an indicator for any children/sons/daughters appearing in our Numident death-census linked sample or a discrete variable for the number of children/sons/daughters appearing in our linked sample. All regressions include fixed effects for men's 3-digit occupation-by-birth cohort, Social Security coverage category, race, and birthplace. Robust standard errors in parentheses, clustered by men's 3-digit industry. *** 1%, ** 5%, * 10% significance.

Table 3: Placebo checks: 1930 Census characteristics

	Dependent variable: Household characteristic in 1930								
	Homeowner	Home value (log)	Own a radio	Live in urban area	Live on farm	Number of children	Dad literate	Mom literate	Mom in labor force
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel (a): Population									
Social Security exposure	-0.0009	-0.0017	-0.0002	0.0002	0.0002	0.0020	-0.0004	-0.0002	-0.0001
	(0.0006)	(0.0014)	(0.0009)	(0.0006)	(0.0002)	(0.0023)	(0.0003)	(0.0003)	(0.0003)
Mean of dep. variable	0.49	2,750	0.47	0.74	0.05	2.2	0.94	0.93	0.11
N	3,912,603	1,761,693	3,920,008	3,920,008	3,920,008	3,920,008	3,920,008	3,707,183	3,707,185
Panel (b): Families in Numident death-census linked sample									
Social Security exposure	-0.0004	-0.0015	0.0006	0.0003	0.0004	0.0032	-0.0002	-0.0002	-0.0001
	(0.0008)	(0.0018)	(0.0013)	(0.0008)	(0.0004)	(0.0030)	(0.0004)	(0.0004)	(0.0008)
Mean of dep. variable	0.49	2,760	0.48	0.74	0.05	3.8	0.94	0.94	0.09
N	857,934	434,707	858,612	858,612	858,612	858,612	858,612	839,750	839,750

Notes: This table shows the coefficients on Social Security exposure (i.e., negative of earliest age of Social Security eligibility) from equation (2), where the outcome is a specified household characteristic (indicated in the column titles). In panel (a), the sample consists of all men in our population of interest in the 1930 Census while in panel (b) the sample only includes the subset of men who appear in our Numident death-census linked sample. All regressions include fixed effects for men's 3-digit occupation-by-birth cohort, Social Security coverage category, race, and birthplace. In column 2, mean 1930 home values in contemporary dollars are shown. In panel (b), observations are weighted using inverse propensity score weights that re-weight the sample to match the population of interest in 1930. Robust standard errors in parentheses, clustered by men's 3-digit industry. *** 1%, ** 5%, * 10% significance.

Table 4: The impact of father's exposure to Social Security on children's late-life ZIP quality

	Dependent variable: Death ZIP characteristics					
	Mean AGI (2001)		Median home values (2000)		ZIP quality index	
	Rank	Log	Rank	Log	Rank	SD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): All children						
Social Security exposure	0.102*	0.0009	0.119**	0.0022*	0.105*	0.0036
	(0.053)	(0.0007)	(0.050)	(0.0011)	(0.058)	(0.0026)
Mean of dep. variable	64.0	47,937	68.8	144,113	54.4	0.16
N	1,303,454	1,303,422	1,270,763	1,269,551	1,267,789	1,267,789
Panel (b): Sons						
Social Security exposure	0.198***	0.0024***	0.199***	0.0043***	0.194***	0.0054**
	(0.049)	(0.0008)	(0.044)	(0.0012)	(0.046)	(0.0024)
Mean of dep. variable	63.8	47,639	68.7	143,128	54.3	0.15
N	655,886	655,862	638,754	638,195	637,235	637,235
Panel (c): Daughters						
Social Security exposure	0.031	0.0001	0.051	0.0005	0.062	0.0034
	(0.064)	(0.0009)	(0.063)	(0.0015)	(0.073)	(0.0028)
Mean of dep. variable	63.6	47,509	68.4	142,359	53.7	0.13
N	647,568	647,560	632,009	631,356	630,554	630,554

Notes: This table shows the coefficients on father's Social Security exposure (i.e., negative of earliest age of Social Security eligibility) from equation (3), where the sample consists of children in our Numident death-census linked sample and the outcome is the quality a child's late-life ZIP code in terms of mean AGI, median home values or the neighborhood quality index (either in percentile ranks or in logs). Panels show separate estimates for all children, sons and daughters. All regressions include fixed effects for children's sex-by-birth cohort and fathers' 3-digit occupation-by-birth cohort, Social Security coverage category, race, and birthplace. In columns 2 and 4, means of ZIP-level 2001 AGI and 2000 home values in contemporary dollars are shown. Observations weighted using inverse propensity score weights that re-weight the sample to match the population of interest in 1930. Robust standard errors in parentheses, clustered by fathers' 3-digit industry. *** 1%, ** 5%, * 10% significance.

Table 5: The impact of father's exposure to Social Security on sons' occupational standing in adulthood

	Dependent variable: Occupational income score											
	First occupation			1962 occupation			1972 occupation			Occupation around age 50		
	Log	Level	Missing	Log	Level	Missing	Log	Level	Missing	Log	Level	Missing
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel (a): Sons born in 1907–1930												
Social Security exposure	-0.003	-33	-0.032*	0.040***	378***	-0.006	0.024**	253**	-0.011	0.051***	489***	0.000
	(0.023)	(82)	(0.017)	(0.015)	(128)	(0.024)	(0.012)	(119)	(0.016)	(0.014)	(127)	(0.02)
Mean of dep. variable	8.07	3,553	0.07	8.93	8,154	0.10	9.19	10,526	0.09	9.07	9,489	0.07
N	2,816	2,816	3,037	2,747	2,747	3,037	2,754	2,754	3,037	2,817	2,817	3,037
Panel (b): Sons born in 1907–1919												
Social Security exposure	-0.003	-19	-0.034	0.033*	332*	-0.017	0.021	210	-0.019	0.038*	362**	-0.004
	(0.02)	(67)	(0.022)	(0.02)	(172)	(0.028)	(0.015)	(167)	(0.023)	(0.02)	(180)	(0.023)
Mean of dep. variable	8.04	3,476	0.07	8.93	8,185	0.10	9.17	10,419	0.13	8.97	8,606	0.10
N	1,627	1,627	1,754	1,577	1,577	1,754	1,524	1,524	1,754	1,587	1,587	1,754

Notes: This table shows the coefficients on father's Social Security exposure (i.e., negative of earliest age of Social Security eligibility) from equation (3), where the sample consists of sons in our 1973 OCG sample and the outcome is the occupational income score associated with sons' first job, job in 1962, or longest job in 1972 (in logs or in 1972 dollars), or an indicator for having missing occupational information. To measure occupational standing at these three points in time, we respectively use the 1940, 1960, and 1970-based sex-by-race-by-region-specific occupational income scores from Collins and Wanamaker (2022). Panels show separate estimates for sons born in 1907–1930 and sons born in 1907–1919. The last three columns show analogous outcomes for sons' occupational standing around age 50, defined as 1962 for sons born in 1907–1917 and 1972 for sons born after 1917. All regressions include fixed effects for sons' birth cohort and race, and fathers' 3-digit occupation-by-birth cohort, Social Security coverage category, and birthplace. Observations weighted using OCG sampling weights. Robust standard errors in parentheses, clustered by fathers' 3-digit industry. *** 1%, ** 5%, * 10% significance.

Table 6: The impact of father's exposure to Social Security on sons' earnings in Social Security covered employment

	Dependent variable:							
	Earnings in 1957–1967 (1907–1930 cohorts) or in 1951–1967 (1907–1919 cohorts)				Earnings at ages 44–53 (1907–1922 cohorts)			
	Total earnings (log)	Total earnings in 4-QC years (log)	Earnings ⩾ taxable max in all 4-QC years	Mean QC < 2	Total earnings (log)	Total earnings in 4-QC years (log)	Earnings ⩾ taxable max in all 4-QC years	Mean QC < 2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel (a): Sons born in 1907–1930								
Social Security exposure	0.058**	0.066**	0.052**	-0.007	0.040	0.046	0.061*	0.001
	(0.025)	(0.027)	(0.025)	(0.018)	(0.035)	(0.036)	(0.036)	(0.02)
Mean of dep. variable	67,847	66,069	0.46	0.18	65,283	63,567	0.46	0.19
N	2,290	2,290	2,290	2,818	1,674	1,674	1,674	2,086
Panel (b): Sons born in 1907–1919								
Social Security exposure	0.070	0.078	0.063**	0.001	0.044	0.051	0.066*	0.002
	(0.047)	(0.048)	(0.03)	(0.022)	(0.044)	(0.046)	(0.036)	(0.023)
Mean of dep. variable	96,497	93,846	0.37	0.18	61,991	60,304	0.47	0.18
N	1,336	1,336	1,336	1,625	1,324	1,324	1,324	1,625

Notes: This table shows the coefficients on father's Social Security exposure (i.e., negative of earliest age of Social Security eligibility) from equation (3), where the sample consists of the subset of sons in our 1973 OCG sample linked to SSA earnings records. In columns 1 to 4 the outcomes pertain to earnings in 1957–1967 (1907–1930 cohorts) or 1951–1967 (1907–1919 cohorts). In columns 5 to 8, the outcomes pertain to earnings at ages 44–53 (1907–1922 cohorts only). In columns 1 and 5, the outcome is the log of total earnings. In columns 2 and 6, the outcome is the log of total earnings in all years in which the respondent had 4 quarters of coverage. In columns 3 and 7, the outcome is an indicator for earnings exceeding the Social Security taxable maximum in all years in which the respondent had 4 quarters of coverage (see Appendix Figure A.9 for maximum amounts by year). For all outcomes, we restrict the sample to respondents with an average of 2 or more quarters of coverage over the relevant time frame. In columns 4 and 8, the outcome is an indicator for not satisfying this condition. Means in columns 1, 2, 5 and 6 are for corresponding outcomes in levels (in 1972 dollars). Panels show separate estimates for sons born in 1907–1930 and sons born in 1907–1919. All regressions include fixed effects for sons' birth cohort and race, and fathers' 3-digit occupation-by-birth cohort, Social Security coverage category, and birthplace. Observations weighted using OCG sampling weights. Robust standard errors in parentheses, clustered by fathers' 3-digit industry. *** 1%, ** 5%, * 10% significance.

Table 7: The impact of father's exposure to Social Security on children's propensity to live outside their parents' 1930 place of residence by 1950

	Dependent variable: Children's 1950 location vs. parents' 1930 location							
	Different location					Distance between counties		
	County	County (same state)	State	Census division	Census region	> 0 and $\leqslant 100$ miles	> 100 and $\leqslant 500$ miles	> 500 miles
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel (a): All children								
Social Security exposure	-0.0002	-0.0012*	0.0010	0.0013*	0.0006	0.0000	-0.0007	0.0005
	(0.0010)	(0.0007)	(0.0007)	(0.0007)	(0.0006)	(0.0007)	(0.0009)	(0.0005)
Mean of dep. variable	0.43	0.20	0.23	0.16	0.12	0.19	0.14	0.10
N	601,276	601,276	601,321	601,321	601,321	601,276	601,276	601,276
Panel (b): Sons								
Social Security exposure	0.0005	-0.0015*	0.0020**	0.0025***	0.0013	0.0000	-0.0013	0.0018*
	(0.0017)	(0.0009)	(0.0010)	(0.0009)	(0.0011)	(0.0008)	(0.0011)	(0.0010)
Mean of dep. variable	0.42	0.19	0.23	0.16	0.13	0.18	0.13	0.11
N	309,657	309,657	309,682	309,682	309,682	309,657	309,657	309,657
Panel (c): Daughters								
Social Security exposure	-0.0010	-0.0008	-0.0002	0.0004	0.0000	-0.0003	-0.0001	-0.0006
	(0.0010)	(0.0010)	(0.0009)	(0.0010)	(0.0007)	(0.0011)	(0.0011)	(0.0006)
Mean of dep. variable	0.43	0.21	0.22	0.15	0.11	0.19	0.14	0.10
N	291,619	291,619	291,639	291,639	291,639	291,619	291,619	291,619

Notes: This table shows the coefficients on father's Social Security exposure (i.e., negative of earliest age of Social Security eligibility) from equation (3), where the sample consists of children in our Numident death-census linked sample who are additionally linked to the 1950 Census and the outcome is an indicator for living outside parents' 1930 county/state/division/region of residence or a specified number of miles away from their 1930 county of residence by 1950. A handful of counties cannot be mapped to a county group with stable borders between 1930 and 2010, resulting in slight differences in sample sizes across columns. Panels show separate estimates for all children, sons and daughters. All regressions include fixed effects for children's sex-by-birth cohort and fathers' 3-digit occupation-by-birth cohort, Social Security coverage category, race, and birthplace. Observations weighted using inverse propensity score weights that re-weight the sample to match the population of interest in 1930. Robust standard errors in parentheses, clustered by fathers' 3-digit industry. *** 1%, ** 5%, * 10% significance.

Table 8: The impact of father's exposure to Social Security on children's direction of migration

	Dependent variable: Earnings in 1950 in children's 1950 state vs. parents' 1930 state					
	Education-specific earnings			Overall earnings		
	Median	25th percentile	75th percentile	Median	25th percentile	75th percentile
	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): All children with information on completed years of schooling from 1940 or 1950						
Social Security exposure	0.13	-0.65	0.64	-0.82	-0.76*	-0.32
	(0.55)	(0.43)	(0.57)	(0.69)	(0.45)	(0.66)
Mean of dep. variable	15.8	-0.6	20.6	17.7	2.2	22.2
N	196,275	196,275	196,275	196,275	196,275	196,275
Panel (b): Sons with information on completed years of schooling from 1940 or 1950						
Social Security exposure	1.39**	-0.29	1.95***	-0.17	-0.66	-0.12
	(0.61)	(0.47)	(0.52)	(1.31)	(0.54)	(1.01)
Mean of dep. variable	15.6	-1.6	21.5	16.0	0.8	21.6
N	94,382	94,382	94,382	94,382	94,382	94,382
Panel (c): Daughters with information on completed years of schooling from 1940 or 1950						
Social Security exposure	-1.08	-0.91*	-0.56	-1.32	-0.85	-0.39
	(0.82)	(0.52)	(0.79)	(1.12)	(0.59)	(0.90)
Mean of dep. variable	15.8	0.1	19.6	19.0	3.3	22.8
N	101,893	101,893	101,893	101,893	101,893	101,893

Notes: This table shows the coefficients on father's Social Security exposure (i.e., negative of earliest age of Social Security eligibility) from equation (3), where the sample consists of children in our Numident death-census linked sample who are additionally linked to the 1950 Census and for which we can observe completed years of schooling either via additional links to the 1940 Census (respondents aged 25 or older and not attending school) or in the 1950 Census (sample-line respondents aged 25 or older and not attending school). In columns 1 to 3, the outcomes are differences in moments of the 1950 education-specific earnings distribution in children's 1950 state of residence versus their parents' 1930 state of residence (median, 25th percentile, and 75th percentile). In columns 4 to 6, the outcomes are analogous but for moments of the overall earnings distribution by state in 1950. Moments of the earnings distribution by state (and education) are calculated using (sample-line) men aged 25–54 in the 1950 Census. All amounts are expressed in 1949 dollars. Panels show separate estimates for all children, sons and daughters. All regressions include fixed effects for children's sex-by-birth cohort and fathers' 3-digit occupation-by-birth cohort, Social Security coverage category, race, and birthplace. Observations weighted using inverse propensity score weights that re-weight the sample to match the population of interest in 1930. Robust standard errors in parentheses, clustered by fathers' 3-digit industry. *** 1%, ** 5%, * 10% significance.

Table 9: Derived effects on lifetime earnings, family gains, and fiscal implications

	Source of late-career earnings-gain estimate		
	ZIP income (baseline) (1)	ZIP home prices (2)	OCG occscore (3)
Estimate	2.4%	4.3%	35.0%
	[0.8%, 4.0%]	[1.9%, 6.7%]	[1.7%, 68.3%]
Implied individual earnings gain			
Late career (1962 onward)	28.2%	46.7%	35.0%
	[9.7%, 46.8%]	[20.7%, 72.8%]	[1.7%, 68.3%]
Full-career average	11.9%	19.7%	14.8%
	[4.1%, 19.8%]	[8.7%, 30.7%]	[0.7%, 28.8%]
EPV per son	$130K	$215K	$161K
	[$45K, $215K]	[$95K, $335K]	[$8K, $314K]
Implied family gains (per recipient family) and fiscal implications			
EPV of children's earnings gain	$299K	$494K	$370K
	[$102K, $495K]	[$219K, $770K]	[$18K, $723K]
$\div$ ΔEPV(SS)	6.4	10.6	8.0
	[2.2, 10.6]	[4.7, 16.5]	[0.4, 15.5]
Net family earnings $\div$ ΔEPV(SS)	5.7	9.9	7.2
	[1.4, 9.9]	[3.9, 15.8]	[−0.4, 14.8]
Net govt. cost $\div$ mech. effect	−0.55	−1.77	−0.99
	[−1.77, 0.67]	[−3.48, −0.05]	[−3.19, 1.20]

Notes: Amounts in 2020 dollars, present values as of 1950. 95% confidence intervals (in brackets) capture sampling uncertainty but not modeling uncertainty. Late-career earnings: ZIP estimates divide treatment effect by cross-sectional gradient; OCG maps occscore gains one-to-one into earnings gains. Full-career earnings: No gain through 1950 (null effects on occscore at first job and in 1950); linear earnings growth 1950–1962 to late-career gain and constant thereafter (similar occscore gains in 1962 and 1972); sons born in 1910. EPV per son: Baseline earnings path is median annual earnings for male four-quarter workers from Social Security Bulletin Annual Statistical Supplements (Social Security Administration, 1965, 1975); discount to 1950 at 3% annual interest rate; mortality based on SSA 1910 cohort male life table conditional on survival to 1930. EPV of children's earnings: Scale per-son gain by average sons per recipient father (2.3). ΔEPV(SS): simulated change in PV of Social Security per recipient family from 10-year difference in earliest eligibility age ($\approx$ $47,000). Net family earnings: 1877 father cohort; 25% reduction in recipients' labor force participation after eligibility at 65; "first stage" of 1 (differences in predicted eligibility translate one-for-one into differences in actual eligibility, a conservative upper bound). Net government cost: Additional Social Security benefits plus forgone tax revenue from recipients' earlier retirement less extra revenue from recipients' children; net tax rates are 18.8% for recipients and 25.3% for sons (see Appendix Section E.1). Mechanical effect (increase in benefits holding behavior fixed) differs from ΔEPV(SS) by the "behavioral effect" from fathers' behavioral responses to Social Security. See Appendix Section E for more details and Appendix Table A.15 for robustness tests.

Online Appendix for

Long-Run Intergenerational Effects of Social Security

By Daniel K. Fetter, Lee M. Lockwood, and Paul Mohnen

Contents of Online Appendix

A Data appendix

A.1 Linking families in SS-5 records to households in the 1920 and 1930 Censuses

The first step involves reconstituting families in the SS-5 data by linking siblings together primarily based on parent names (including mothers' maiden names).[1] The sibling linkages are then used to create a crosswalk between SSNs and family identifiers. In practice, we use the crosswalk from Bailey et al. (2026). These links were made using the LIFE-M supervised ML approach described in this paper, except that linking siblings together is an instance of one-to-many linking (since people can have multiple siblings). As a result, trainers were allowed to select multiple links, and their decisions were modeled using a random forest model treating each pair of records as a separate linking decision. In the final crosswalk, 35 percent of individuals can be linked to at least one other sibling in the SS-5 data (see Bailey et al. (2026) for more details).

The next step involves linking reconstituted families in SS-5 records to households in the 1930 Census. Although we are mainly interested in families in which the father was born in 1875–1888, since SS-5 records lack information on parents' year of birth we link all families with at least one child aged under 25 at the time of the census and later restrict the sample based on fathers' age in the census. Note that families can consist of a single SS-5 record if no sibling was found in the previous step. To generate the set of candidate census households for each family, we block on fathers' last name initial and the union of children's state/continent of birth.[2] Blocking on the union of children's place of birth implies that the only requirement we impose is that at least one of the children's place of birth in candidate families matches one of the children's place of birth in the family under consideration. Within each block, we select the top 20 candidate households based on a Jaro-Winkler string similarity score between parent names.

Next, we generate training data for a random sample of 500 families using the process described in Section 3. In particular, we display the following information to trainers: fathers' full name, mothers' first and middle name, and children's first and middle name, age in 1930, and state/country of birth (we also display the modal race among children). We only display children that would have been alive at the time of the census, and mask children in the census if there is little chance that they match any of the children in the family under consideration (based on age and birthplace information). We also only display the eight youngest children to avoid overcrowding the computer screen during the training process.

We are able to manually link 61.9 percent of families in the training data to a household in the 1930 Census. To train the two-stage model from Bailey et al. (2023) to mimic the hand-linking process, we first split the training data into a "training set" containing 70 percent of the training data and a "test set" containing the remaining 30 percent. The training set is used to iteratively train the algorithm by varying model features until we are satisfied with its performance, while the test set is used to assess the out-of-sample performance of the model once it is finalized. The final set of features used to train the model are shown in Table A.13. Following standard practice, probabilities are estimated using 10-fold cross-validation to avoid overfitting. Panel (a) in Figure A.15 shows the resulting precision-recall curve. The model is able to achieve a recall rate of 85.6 percent at 97

[1] Place of birth, year of birth, and race information is also used to influence linking decisions, as siblings are more likely to be born in the same state/country within a few years of each other and have matching race information.

[2] Blocking on a certain characteristics is common practice in historical record linking since it reduces the number of pairs for which string similarity scores must be computed.

percent precision in the training set (match rate of 53 percent). The corresponding numbers in the test set are 78.5 percent recall at 96.1 percent precision (match rate of 50.7 percent).[3]

Using the trained model and cross-validated probability threshold, we scale-up linking to the full sample of SS-5 families. This yields 7 million links, for a match rate of 33.9 percent.[4] Among those links, there are cases in which the same census household is linked to two (or more) families in the SS-5 data. While some of these cases are true conflicts (reflecting our model's 3 percent error rate, for example), a significant share of them are not. The reason is that parents can naturally appear on multiple SS-5 records for their different children and we are not able to make all possible links between these records (we can manually link only a subset of all true links due to missing or imperfectly reported names, and the model can only replicate a subset of those links at 97 percent precision). Although we might not be able to link two SS-5 records together solely based on parent name information, we might still be able to correctly link them to the same census household since the latter additionally exploits information on children. As a result, we end up with apparent conflicts that simply reflect the incompleteness of our family identifiers.

Rather than throwing away all these conflicts, we manually examined a random sample of 200 cases, and either chose one link over the others, concluded all links were correct, or chose none of them if we were not confident. We then used this labeled data to assess the performance of different conflict resolution rules in terms of their ability to retain all correct links and drop all incorrect links.[5] The rule we ultimately settled on is able to retain 80 percent of correct links and drop 75 percent of incorrect links in the random sample of cases we manually inspected. The final number of SS-5 families linked to a household in the 1930 Census at 97 percent precision after resolving these conflicts is 6.9 million, for a final match rate of 33.6 percent. Using an analogous procedure and recycling the training data and model, we are able to link 4.1 million SS-5 families to a household in the 1920 Census (match rate of 30.1 percent). This number of links to 1920 is lower because fewer children are born by 1920 in the SS-5 data.

A.2 Linking households in the 1930 Census to households in the 1920 Census

The process of linking households in the 1930 Census to households in the 1920 Census is very similar to the process used to link SS-5 families to households in the 1920 and 1930 Censuses. One slight difference is that the censuses additionally contain information on parents' age and place of birth. Therefore, when generating sets of potential candidates, we block on fathers' state/continent of birth and a +/-5-year window around fathers' age in 1920 (in addition to fathers' last name initial). We also display fathers' age in 1920, mothers' age in 1920, and mothers' place of birth during the training process. We are able to manually link 52.6 percent of households headed by a man born in 1860–1905 in the 1930 Census to a household in the 1920 Census in our training data (500 cases). We used a similar set of features as in Table A.13 to train the model, which is able to achieve a recall rate of 90 percent at 97 percent precision in the training set, for a match rate

[3] Note that in this instance the training set only contains 150 cases. With more training data the discrepancy between the in-sample and out-of-sample performance statistics would be smaller.

[4] This match rate is lower than the match rates in our training data because we dropped the requirement that SS-5 families have at least one child aged under 25 at the time of the census when scaling up. This added some extra links, but artificially reduced the match rate by expanding the denominator to include families with a much lower probability of being linked.

[5] The two extreme rules are to keep all conflicts or drop all conflicts, which exactly achieve one objective but not the other. The optimal rule lies somewhere in between.

of 47.3 percent (see panel (b) in Figure A.15 for the precision-recall curve). The corresponding numbers in the test set are 90.6 percent recall at 100 percent precision (match rate of 51.3 percent).

Note that after scaling up the model it is possible for a household in the 1920 Census to be linked to multiple households in the 1930 Census. But in contrast to when we link SS-5 families to households in the censuses, these are true conflicts since parents can only appear in one household in any given census. To resolve those conflicts, we used model-based match probabilities to prioritize higher-quality links and dropped all remaining conflicts. The final number of households headed by a man born in 1860–1905 in the 1930 Census linked to a household in the 1920 Census at 97 percent precision is 7.6 million (match rate of 47.3 percent).

Note that while most of our final linked sample consists of links between SS-5 records and the 1920 or 1930 Census combined with links between the 1920 and 1930 Censuses, around 80,000 observations are based on links between SS-5 records and both the 1920 and 1930 Censuses *without* links between the 1920 and 1930 Censuses.

A.3 Linking men and women in SS-5 records to the 1950 Census

To link male and female children in SS-5 records to themselves in the 1950 Census, we followed the process described in Bailey et al. (2026) to link SS-5 records to the 1940 Census (including recycling their training data and model). While we are only interested in a subset of children for our analysis, we simply linked all children who would have been alive in 1950 and imposed our sample restrictions after combining all links. Below we provide an overview of how we adapted the process from Bailey et al. (2026) for our purposes, but more details can be found in their paper.

For male SS-5 records, we search for potential candidates in the 1950 Census that match on sex, first or last name initial, state/continent of birth, and a +/-3 year window around the person's age in 1950. Taking the union of blocks based on first or last name initials implies that potentials candidates only need to satisfy one of those two conditions, which is a weaker requirement. We then compute Jaro-Winkler name similarity scores and retain the top 25 candidates. The original training data for linking men in SS-5 records to the 1940 Census was generated using a similar process to the one described in Section 3. It was based on a random draw of 1,500 records and the following information was displayed to trainers: first, middle and last names, age in 1940, country of birth (if born outside the US), and race.[6] The manual match rate in this training data was 42.6 percent, and the recall rate achieved by the model in the training set was 68.6 percent at 97 percent precision (match rate of 29.2 percent). The corresponding numbers in the test set were 69.9 percent recall at 95.3 percent precision (match rate of 28.5 percent). We used this model and the cross-validated probability threshold to generate 97-percent precision links to the 1950 Census, ultimately linking 6.5 million men in SS-5 records for a final match rate of 33 percent.[7]

The process of linking female SS-5 records is analogous, with a few key differences. First, multiple versions of each female record are generated, one with women's birth name as their last name and others using all possible married names as their last name (most women have at most

[6]Race and country of birth were displayed as optional pieces of information to take into account since they are not always accurately reported.

[7]Similar to linking households across censuses, it is possible for the same census individual to be linked to two different SSNs, which are true conflicts since a person can only be associated with one SSN. To resolve those cases, we again prioritized higher-quality links and discarded the remaining conflicts.

one married name, but some have multiple married names reflecting re-marriages).[8] Second, we additionally block on marital status when generating potential candidates. Women linked using their birth name are linked to single women in the census, while women linked using their married name are linked to non-single women in the census (married, divorced, or widowed). The original training data for linking women to the 1940 Census was based on a random sample of 500 SSNs (corresponding to around 1,500 name variants). Unlike other links, it is possible for female SSNs to be linked to different people in the census (if their birth and married names happen to be common names, for example). For women that had multiple links in the training data, those cases were manually resolved by either choosing one link and discarding the others or discarding all links if there was too much ambiguity. The SSN-level match rate in this training data was 52 percent and the SSN-level recall rate achieved by the model in the training set was 70 percent at 96.6 percent precision (match rate of 36.4 percent). The corresponding numbers in the test set were 68.3 percent recall at 95.3 percent precision (match rate of 34.4 percent). Again, we used this trained model and the cross-validated probability threshold to generate 97-percent precision links to the 1950 Census for women (here we simply discarded all links when there was a within-SSN conflict). The final number of links for women was 5.7 million, for a final match rate of 33.4 percent.

A.4 Linking men in claims records to the 1930 Census

Claims records for men born between 1865 and 1899 are linked to the 1930 Census using individual-level information. Given that our objective is to explore claiming behavior for retirement benefits, we focus on the subset of life claims and drop death claims.[9] We separately link claims records with non-missing birthplace information (40 percent of records) and claims records with missing birthplace information (60 percent of records). Claims records with non-missing birthplace information are linked using name, year of birth, and birthplace information. To generate sets of potential candidates, we block on state/continent of birth, last name initial, and a +/-3-year window around age in 1930. We then compute name similarity scores and select the top 20 candidates. We manually linked a random sample of 1,000 cases, displaying name, age, and country of birth information (when applicable). We also displayed the state associated with the first 3 digits of a person's SSN and the state of residence of potential candidates in the 1930 Census. Until 1972, the first 3 digits of an SSN were tied to the field office in which the application was filed. Since 1930 is relatively close in time to the period in which these cohorts would have applied for an SSN, this optional piece of information is useful when choosing between multiple plausible candidates. We are able to link 52.8 percent of claims records with non-missing birthplace information to the 1930 Census, and our model achieves a recall rate of 66.5 percent at 97 percent precision in the training set for a match rate of 35.4 percent (see panel (c) in Figure A.15 for the precision-recall curve). The corresponding numbers in the test set are 67.5 percent recall at 97.3 percent precision (match rate of 37.7 percent).

We use an analogous procedure to link claims records with missing birthplace information, except that we cannot block on state/continent of birth, and instead block on the first two characters of the last name for computational reasons (in addition to the +/-3-year age window). This results

[8]In practice, we use every unique last name variant across all SS-5 entries for the same SSN, which can include name misspellings that were later corrected. This generally leads to duplicate links that can be dropped at the end.

[9]Life claims capture retirement and disability benefits. Death claims capture survivor and death benefits. Around 77 percent of records in the claims data are life claims.

in more ambiguity in the set of potential candidates. Consequently, we are only able to manually link 33.9 percent of claims records with missing birthplace information to the 1930 Census. The model recall rate at 97 percent precision is 67.1 percent, for a match rate of 22.7 percent (see panel (d) in Figure A.15 for the precision-recall curve). The corresponding numbers in the test set are 64.5 percent recall at 98.4 percent precision (match rate of 20.3 percent).

After scaling up linking to the full set of claims records and resolving conflicts in which the same census individual was linked to two different claims records, we end up with 1.3 million links, for a final match rate of 26.2 percent.

When documenting a "first stage" on retirement claiming behavior in Section 4.3, we define retirement claims as life claims that were made at age 65 and older, and with no "date of onset" indicated (to exclude disability claims). Additionally, we exclude claims made with Social Security Numbers whose first 3 digits were between 700 and 728, which were used for Railroad Retirement Benefits. Finally, we exclude claims made in 1965 or 1966. The Numident claims data contains an unusually large number of claims in these years, likely because it is disproportionately likely to include transitional benefits and "Prouty" benefits, which were available to individuals from certain older birth cohorts who had not accrued enough quarters of coverage to qualify for Social Security benefits. Since we are unable to identify specific records as transitional or Prouty benefits, we simply exclude those years.

We also use links to claims records to augment our main linked sample (connecting children to their families in the 1920 and 1930 Censuses). More specifically, we link men born between 1895 and 1920 in claims records to themselves in the 1920 Census using the same process (keeping all types of claims since we are not interested in the claims records themselves). This yields 1.2 million links, for a match rate of 13.5 percent. This match rate is lower because the share of claims records with non-missing birthplace information in this sample is lower (16 percent). We then combine these links with linkages between households in the 1920 and 1930 Censuses that we have previously described. In the end, this results in around 150,000 additional links for sons that could not be made via SS-5 records (primarily because some SSNs are contained in claims records but not SS-5 records).

A.5 Details on HRS sample construction

To document the statistical relationship between late-life neighborhood quality and measures of lifetime and late-career earnings, we leverage data from the Health and Retirement Study, a nationally representative longitudinal survey of Americans aged 50 and older that has taken place every two years since 1992.[10] More specifically, we combine publicly available data on HRS respondents with restricted-access information on: (1) their ZIP code of residence at the time of the surveys (Health and Retirement Study, 2023), and (2) estimates of their lifetime and late-career earnings based on linkages to administrative records from the SSA (Health and Retirement Study, 2021).

The sample construction proceeds in two steps. Starting with the HRS Cross-Wave Tracker File covering all waves between 1992 and 2018 (Health and Retirement Study, 2024), we first identify respondents' last known spouse in the HRS files (if available), and merge in estimates of lifetime earnings for respondents and their spouses. This allows us to generate measures of

[10]The HRS (Health and Retirement Study) is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan.

lifetime earnings at both the individual and household level (household lifetime earnings are likely a better measure for women in the HRS cohorts, many of whom had low lifetime labor force attachment). Lifetime earnings are based on administrative earnings records provided by the SSA, which in combination with imputation and projection methods are used to construct estimates of cumulative lifetime earnings through age 65.[11] We can also generate estimates of late-career earnings by subtracting lifetime earnings through age 50 from lifetime earnings through age 65 (i.e., cumulative earnings between ages 50 and 65). In a second step, we merge in wave-level ZIP code of residence information. This allows us to observe the quality of respondents' last known ZIP code of residence (in the last wave they were interviewed), as measured by mean income or other characteristics. The final sample contains around 22,200 HRS respondents aged 65+ at the time of their last interview, around 55 percent of whom are women.

A.6 Measuring family size, sibling sex composition, and birth order

Since the 1920 and 1930 Censuses and Numident records only contain a subset of all children ever born in a given family, and because there are no identifiers that we can use to identify unique children (i.e., the same child can appear across multiple records), characteristics like family size, sibling sex composition, and birth order must be approximated using what we can observe in the data.[12] Here we describe our approach.

To obtain estimates of family size (i.e., the total number of children in a family), we add together: (1) the number of sons and daughters that appear in the household in 1920, (2) the number of sons and daughters who were born after 1920 that appear in the household in 1930, and (3) the number of sons and daughters who were born after 1930 that appear in our linked sample (i.e., in Numident records). In some families, the unconditional total number of sons and daughters in our linked sample (i.e., regardless of when they are born) can strictly exceed the corresponding total number of sons and daughters obtained by adding (1), (2) and (3). This can happen for a variety of reasons, including measurement error in self-reported sex and year of birth information, the fact that some children born before 1930 can appear in Numident records but not in the censuses (e.g., older children who left the household prior to 1920), but also linking errors. In those cases, we use the higher number as our final estimate of the total number of sons and daughters. Those numbers are combined to obtain an estimate of the total number of children, and we also use them to determine the sex composition of siblings. For example, a daughter has at least 1 brother if the number of sons in the family is greater than 0, and she has at least 1 sister if the number of daughters is greater than 1. In the heterogeneity cuts by sibling sex composition, we focus on families with at least two children and control for family size to try to avoid confounding differences in sex composition with differences in family size.

Rather than trying to determine the exact birth order of a given child in a family, we only try to determine whether that child is the youngest child, the eldest child, or a middle child. To that end,

[11] For example, if a respondent only has matched SSA earnings records through age 62, those earnings are projected forward to generate lifetime earnings through age 65. For respondents without matched SSA earnings records, their cumulative earnings at HRS entry and subsequent annual earnings are imputed based on observable characteristics using a nearest neighbor algorithm. For more details, see Fang (2021).

[12] For the most part, children can be uniquely identified within Numident records and within each decennial census (apart from rare cases where the same person was erroneously assigned two SSNs or accidentally listed twice during the census enumeration process). The challenge lies in identifying unique children across records.

we compute the minimum and maximum year of birth among all children we observe across the 1920 Census, 1930 Census, and Numident records. We then define a child to be a youngest child if his or her year of birth in the Numident data is equal to the maximum year of birth, an eldest child if it is equal to the minimum, and a middle child if it is strictly between the minimum and maximum. Note that eldest and youngest children are not necessarily unique due to multiple births (e.g., twins). Also, it is possible for the minimum and maximum year of birth to coincide (e.g., only-children). In the heterogeneity cuts by birth order, we focus on families with at least two children born in different years to avoid having the same children simultaneously being counted as eldest and youngest children.

B Classification of employment into OASI coverage categories

B.1 Identifying Social Security coverage categories

Our assignment of Social Security coverage year to an individual is based on his reported occupation, industry, and "class of worker" (which includes information on whether a worker was self-employed, a wage or salary earner, or an unpaid family worker). Although our analysis uses only some of the Social Security coverage categories, we classify all combinations of occupation, industry, and class of worker and describe our approach here. We do not offer an exhaustive description of all categories, but cover the largest ones. Our description and classification is based primarily on Nelson (1985), supplemented with information from Marquis (1955) and Social Security Administration (1980), as well as the Social Security Act and subsequent amendments. Additional useful references include Leibowitz (1950) on the 1950 Amendments, Schottland (1956) on the 1956 Amendments, and McCamman (1956) on state and local government workers under the 1954 Amendments.

The original Social Security Act in 1935 included most wage and salary workers in its definition of covered employment. It excluded self-employment, as well as certain types of wage and salary work, the largest of which were agricultural wage work; domestic work in private homes; and work for nonprofit organizations, state and local government, the federal government, and railroads.

Since our analysis compares only wage and salary workers, we first describe the extension of Social Security coverage to the categories of wage and salary workers that were excluded from coverage in the 1935 Act. Regular agricultural wage work and domestic work were both covered under the 1950 Amendments, and the requirements for regularity were relaxed under the 1954 and 1956 Amendments. Employees of nonprofit organizations had been excluded from coverage under the 1935 Act because of the tax-exempt status of nonprofits (McCamman, 1948), but (with the exception of ministers) could be covered on an elective basis starting with the 1950 Amendments. The decision to elect coverage was made by both employers and employees: if the organization waived its exemption from Social Security taxes and employees voted to be covered by Social Security, all new employees and existing employees who voted for coverage would be covered, while existing employees who voted against coverage would not be covered. Although not all organizations elected coverage immediately upon its becoming available, take-up was reasonably rapid: as of 1951, 56% of employees of nonprofit organizations were covered under Social Security, rising to 76% by 1960 (Subcommittee on Social Security, 1976, p. 27). State and local government employees were initially excluded from coverage because of perceived constitutional limitations on levying taxes on state and local governments (see, e.g. Clark et al., 2011), but beginning with the 1950 Amendments, states could voluntarily enter into agreements for Social Security to cover employees who were not covered under an existing retirement system. The 1954 Amendments extended coverage to state and local employees already under a retirement system, with the exception of firefighters and police officers, if the state elected to make coverage available and employees voted in favor of coverage. Amendments in 1956 further liberalized the provisions under which state and local government employees could be covered by Social Security.[13] About 11 percent of all state and local government employees were covered by Social Security at the beginning of 1951 (the first year in which they could be), and by 1958 the majority of state and local government employees (54 percent) were covered by Social Security (Kestenbaum, 1982).

Other significant types of work excluded from coverage under the 1935 Act, but that do not play a role in our analysis, are self-employment and work for railroads or the federal government. The 1950 Amendments covered most nonfarm self-employment, but excluded self-employed farmers and farm managers, as well as certain professional groups (accountants, architects, engineers, funeral directors, lawyers, medical professionals, and veterinarians). Most self-employed farmers and farm managers were covered in 1954, and additional farmers were covered under liberalized provisions in 1956. Professional self-employment, except dentists, medical professionals, and lawyers, were covered under the 1954 Amendments; except for self-employed doctors, the professional groups excluded in the 1954 Amendments were covered under the 1956 Amendments. Self-employed doctors were covered under the 1965 Amendments. Railroad workers with at least ten years of service were largely covered under the Railroad Retirement Act of 1937, a system parallel and fairly similar to Social Security; those with less than ten years of service were covered by Social Security beginning in 1951. Many employees of the federal government were covered by existing retirement programs in 1935, and therefore excluded from Social Security. The 1950 Amendments extended coverage to a large number of federal employees, mostly part-time or temporary employees, who were not under a Federal retirement system. The 1983 Amendments extended coverage to all civilian employees of the Federal government hired in 1984 onwards. Military employment was covered under the 1956 Amendments (veterans also received noncontributory wage credits for military service between 1940 and 1956).

Our main coding of Social Security coverage is based on the 1950 Census classifications for occupation and industry, which IPUMS uses as a common coding across years (Ruggles et al., 2025b), as well as "class of worker" (which includes information on whether a worker was self-employed, a wage or salary earner, or an unpaid family worker). The foregoing discussion noted that there was a regularity of employment condition for coverage in some types of work, such as agricultural wage work and domestic work. A single Census provides information at only one point in time and, in the case of the 1930 Census, does not provide sufficient information on regularity of employment to map into the regularity of employment tests. Hence, our classification implicitly assumes that, for example, a worker who reports his occupation as an agricultural wage worker is employed on a sufficiently regular basis that he would be covered by Social Security in the first instance (1950 for agricultural wage workers).

We assign 1935 coverage to any wage and salary worker who, based on industry and occupation, was not in one of the original excluded categories. The exclusions were primarily on the basis

[13]Firefighters and police were treated separately, and do not play a role in our analysis. The 1956 Amendments allowed firefighters and police to be covered in certain states, and the 1967 Amendments allowed firefighters to be covered in all states.

of industry, although there are a few small categories we exclude on the basis of occupation (such as fishermen and oystermen, or officers and crew members of ships). We assign 1950 coverage to the categories of employment covered on a compulsory basis: agricultural wage workers, domestic workers, and the non-farm self-employed (except for the professional groups excluded from coverage in 1950). We also assign 1950 coverage to types of employment that were covered on an elective basis beginning in 1950: nonprofits and state and local government. We treat the following industry codes as implying employment in either nonprofits or state and local government: hospitals, educational institutions, welfare and religious organizations, nonprofit membership organizations, and state and local public administration. Of these, hospitals were one case with a nontrivial number of for-profit entities: based on American Hospital Association (AHA) data, about 30 percent of AHA hospitals were were private in 1935.[14] However, private hospitals tended to be smaller than government or nonprofit hospitals: they held only about 6 percent of hospital beds in 1935, suggesting that they represented only a small share of hospital employment. In the main analysis, we do not attempt to separate out for-profit hospitals, but their presence should tend to attenuate the first stage—which nevertheless is strong—and the size of the reduced form effects. We are aware of no simple way in the Census data of distinguishing nonprofit from government employment in hospitals or education, so we treat them in a parallel way, with elective coverage beginning in 1950. Finally, we categorize police and firefighters separately due to their special coverage under the Social Security Act and exclude them from our analysis. They would not be part of the basis for comparisons anyway, since there would be no variation in Social Security coverage within the same occupation.

If a worker's industry, occupation, or self-employment information is missing or not classified into informative codes, we classify their Social Security coverage status as "unclear" and exclude them from our analysis. Similarly, in some cases a worker's reported industry and occupation may imply different dates of coverage (such as a worker whose occupation is "housekeeper, private household" and industry is "Security and commodity brokerage and investment companies."). These cases we also classify as unclear. We also classify as unclear any combination of occupation and industry that was impermissible according to the rules used to classify the 1950 Census returns, as specified in the *Alphabetical Index of Occupations and Industries* (U.S. Bureau of the Census, 1950). Otherwise, we do not exclude, or attempt to correct, apparently inconsistent occupation and industry responses that nevertheless have the same implication for dates of coverage.

To apply our Social Security coverage categories in the OCG sample, which reports occupation and industry according to 1970 Census codes, we first map those codes into 1950 codes, and then apply the Social Security coverage classification described above. Fortunately, despite the changes between 1950 and later years, we still assign a unique Social Security coverage code to each combination of occupation, industry, and self-employment status in the 1970 codes.

B.2 Validation of the coverage definition

In addition to the first stage analysis presented in the paper, in Table A.3 we verify our coverage definition against the 1940 Census, where 5 percent of the population was asked whether they had obtained a Social Security Number (SSN). One qualification to the classification described above, not relevant to our main analysis but relevant when using employment information from 1940, is

[14]The AHA data appear in the August issues of *Hospitals: The Journal of the American Hospital Association*. We thank Heidi Williams for generously sharing a digitized version of the data.

that we exclude public emergency workers. Emergency work was not covered by Social Security (see Marquis, 1955), and we do not know if these workers had worked in covered employment previously. The vast majority of those whose 1940 employment is classified as covered in 1935 had obtained an SSN by 1940, while those whose employment is classified as not covered until later mostly had not obtained an SSN by 1940. Those whose 1940 employment we do not classify as covered, but who did report having an SSN, may have been employed in a private firm in a largely nonprofit industry; they may have formerly held covered employment; they may have been workers whose employment varied seasonally; or they may have simply registered for an SSN on their own, since coverage of one's employment was not required to get one.

C Pension coverage in the comparison group

Because comparisons are restricted to those within occupation and using variation that arises from exclusions of certain industries from coverage under the original Social Security Act, the comparison group in our empirical analysis is made up largely of certain categories of domestic workers, employees of nonprofits, and employees of state and local governments. To aid interpretation of our main results, it is useful to characterize the prevalence and characteristics of retirement plans for these types of workers. Although we are aware of no systematic information on "retirement plans" for domestic workers (formal or informal), there is enough information on retirement plans for nonprofit and state and local workers to characterize their pension coverage, at least broadly.

A large share of State and local government employees were covered by retirement systems prior to their inclusion in the Social Security system, but within that group, the occupations that form the basis of our main empirical comparisons were relatively less likely to be covered by a retirement system. At the beginning of 1942, when the Census Bureau carried out the first systematic national survey, 46 percent of all State and local government employees were covered by a retirement system; the likelihood of coverage was greater for employees of school systems (58 percent) than for non-school employees (38 percent) (United States Bureau of the Census, 1943, Table 2). This aggregate figure conceals wide disparities across states: in California, New York, Ohio, and Pennsylvania, more than 85 percent of permanent, full time employees were covered, while in a handful of states (Alabama, Idaho, Iowa, Mississippi, Missouri, Nebraska, North Carolina, Oklahoma, Oregon, South Carolina, South Dakota, and Wyoming) less than ten percent of permanent, full-time employees were covered (McCamman, 1943). Coverage under State and local retirement systems grew over the course of the 1940s. In the next nationwide survey, in 1952, the share of State and local government employees covered by a State or a local retirement plan had increased to 67 percent, with a further 9.7 percent covered under Social Security; 74 percent of school employees and 62 percent of non-school employees were covered by State or local systems, with an additional 4.9 percent and 13 percent, respectively, covered by Social Security (United States Bureau of the Census, 1953, Table 2).

Beyond broad categories, there is no systematic information on which types of employees were covered in which localities or states, but the available evidence suggests that the public employees that form part of the basis of our comparison group were, compared to other public employees, less likely to be covered by retirement systems. The State and local employees with the highest rates of coverage by State and local systems were police, firefighters, and teachers (McCamman, 1943; Clark et al., 2003, 2011). Police and firefighters in particular had a long history of coverage and are thought to have had close to universal coverage, typically through localities rather than States, by the end of the 1920s; teachers also had widespread, though far from universal, coverage under either State or local plans (Clark et al., 2003, 2011). We exclude police and firefighters from our analysis entirely, and teachers are used for the comparison only to the extent that they appear in covered industries (private, for-profit entities) as well as uncovered ones (public or nonprofit organizations).

On the other hand, particularly as of the early 1940s, less than a third of states had general plans for nonschool employees (United States Bureau of the Census, 1943; McCamman, 1953). Altmeyer (1945) noted that within retirement systems for school department employees, it was also often the case as of the mid-1940s that only teachers were eligible, not custodial or nonteaching employees. Consistent with this claim, a 1946 National Education Association study that described characteristics of retirement plans for school employees in 45 states indicated that all 45 states had plans that covered instructional staff of public school systems, but only 15 of these had plans in which custodial employees of school districts could have membership, and 27 had plans in which clerical employees were covered (National Education Association, 1946, p. 18). Retirement plans for public employees grew significantly over the 1940s, which likely reduced this gap to some extent; McCamman (1953) reports that 32 states had state-wide systems for nonschool city employees as of 1950, 21 of which had been introduced since the 1942 survey. Nevertheless, it appears that many of the public-sector workers who form the basis of our comparison group (such as custodial employees of school districts) were not as likely to be covered by retirement systems as were public employees as a whole.

The nature of benefits under State and local retirement plans varied widely, but there are a few generalizations that are relevant to considering how existing retirement coverage in the control group would affect the interpretation of our results. First, public retirement systems tended to have less generous survivor's insurance than Social Security. Most relevant to our analysis, many of these plans had no provision for widow's benefits, or required that retirees accept a lower benefit during their lifetime in order to continue a benefit to a survivor if the retiree died first (McCamman, 1953). This is particularly relevant to our analysis in that co-residence, and perhaps other types of transfers as well, were especially common among widows. Second, benefits under public retirement systems tended to favor lower-paid workers to a lesser extent that Social Security did; they also required longer periods of service than Social Security, and unlike Social Security after the 1939 Amendments, heavily weighted benefits in favor of those with longer periods of service (McCamman, 1953).

Retirement plans for employees of nonprofit organizations existed by 1940, but appear to have been less common than for public employees. Altmeyer (1944) noted that there were voluntary plans made available by professional associations as of the early 1940s, but take-up tended to be low. Based on a survey carried out by the American Hospital Association (AHA), Hayhow (1940) estimated that about 8 percent of nonprofit hospitals in the United States and Canada had pension plans. In 1946, the AHA introduced its own pension plan, which all member nonprofit hospitals were eligible to join (Bugbee, 1946). By early 1949, 120 hospitals and allied organizations had joined (American Hospital Association, 1949); this take-up was relatively modest compared to the 3,160 nonprofit institutions at the time. Like many public retirement systems, coverage under the AHA plan differed from Social Security; of particular note is that securing continuing payments to a retiree's widow after his death required election of a smaller benefit during one's retirement.

Private institutions of higher education comprise another major group of nonprofit organizations. In the 1940s, about one-third of the approximately 1,700 institutions of higher learning were public, and two-thirds were private; of those two-thirds, more than 90 percent were nonprofit organizations (Murray and Smith, 1940; McCamman, 1948). McCamman (1948) reported that about 30 percent of private institutions (or 338 total) had retirement plans, 78 percent of which were through contracts with the Teachers Insurance and Annuity Association (TIAA). Faculty in these institutions are not part of our empirical comparisons, since they do not appear in both covered and uncovered industries (and because of the rarity of for-profit higher educational institutions, we classify all institutions as public or nonprofit); instead, the employees most relevant to our comparisons are maintenance and other nonprofessional employees. For the most part, contracts that private institutions held with TIAA did not include these types of employees (McCamman, 1948).[15] Of the smaller share that were not covered through TIAA contracts, some were likely affiliated with religious organizations; it is worth noting that Church pension funds frequently covered all employees (Murray and Smith, 1940).

D Reconciling late-career earnings and late-life neighborhood estimates

Our estimates imply that a 10-year earlier predicted Social Security eligibility for fathers increased ZIP-level average income and median home values near the end of sons' lives by about 2 to 4 percent, while increasing sons' late-career occupational income scores by about 35 percent—an order of magnitude difference. In this section, we show that these two sets of estimates are quantitatively consistent once one accounts for the fact that they are not directly comparable: The former are neighborhood-level measures, whereas the latter are individual-level measures. The analysis has two main ingredients. The first is the cross-sectional relationship between late-life neighborhood-level characteristics and individual late-career earnings, which we estimate using the HRS. The second is a simple model of neighborhood choice, which we use to characterize and quantify the potential biases in the mappings from one set of estimates to the other. The broad consistency of the estimates suggests that our more precise ZIP-level estimates can be used as an additional approach to quantifying the late-career earnings gains from Social Security.

D.1 Cross-sectional relationship between late-career earnings and late-life neighborhood

We estimate the cross-sectional relationships between individual late-career earnings and late-life neighborhood characteristics using the HRS sample described in Section 5.1 and Appendix A.5. Specifically, among male HRS respondents we estimate

$$\log(\text{ZIP characteristic}_i) = \alpha + \beta \log(\text{earnings}_i) + \varepsilon_i, \qquad (4)$$

where ZIP characteristic_i is a ZIP-level characteristic of individual i's last known ZIP code of residence (mean adjusted gross income in 2001 or the median home price in 2000) and earnings_i is a measure of individual i's late-career earnings from SSA administrative records (defined as cumulative earnings between ages 50 and 65). The results are in Table A.7. For ZIP income, we estimate $\hat{\beta} = 0.085$ (SE: 0.003), implying that a 10 percent increase in individual earnings

[15] Similar to the other types of retirement systems described here, at the time of retirement, TIAA annuity contracts offered the choice between an annuity that ceased at the retiree's death and one that made smaller payments that continued until his widow's death, if the retiree died first.

is associated with about a 0.85 percent increase in ZIP-level income. For ZIP home prices, we estimate $\hat{\beta}_{\text{h}} = 0.092$ (SE: 0.005).[16]

Mapping the estimated earnings gains into implied ZIP gains Based on the estimated earnings gains and the cross-sectional relationships between late-career earnings and late-life neighborhood characteristics, what effects on neighborhood characteristics would one expect?[17] Using our estimate of the cross-sectional gradient $\hat{\beta}$ from equation (4), the estimated earnings gain from Social Security of roughly 35 percent implies a ZIP-income gain of about 3.0 percent:

$$\text{Implied } \Delta \log(\text{ZIP income}) = \hat{\beta} \times \Delta \log(\text{earnings}) \approx 0.085 \times 0.35 \approx 0.03. \tag{5}$$

This is somewhat above our direct estimate of 2.4 percent but well within its 95 percent confidence interval [0.8%, 4.0%]. The analogous calculation for ZIP-level home values yields a predicted gain of about 3.2 percent (0.092 $\times$ 0.35), somewhat below our direct estimate of 4.3 percent but well within its 95 percent confidence interval [1.9%, 6.7%].

Mapping the estimated ZIP gains into implied earnings gains Using our estimate of the cross-sectional gradient $\hat{\beta}$ from equation (4), the estimated ZIP-income gain from Social Security of roughly 2.4 percent implies an earnings gain of about 28 percent:

$$\text{Implied } \Delta \log(\text{earnings}) = \frac{\Delta \log(\text{ZIP income})}{\hat{\beta}} = \frac{0.024}{0.085} \approx 0.28. \tag{6}$$

This is within the range of our direct estimates of 21–51 percent (SEs from 12–20 percent) and close to the central tendency of about 35 percent. Using the delta method, this estimate has a standard error of about 9.5 percent and a 95 percent confidence interval of roughly [9.7%, 46.8%].[18]

The analogous calculation for ZIP-level home values yields a predicted earnings gain of about 47 percent (0.043/0.092) (SE $\approx$ 13%), also within the range of our direct estimates of 21–51 percent (SEs from 12–20%). It is also consistent with the ZIP-income-based estimate given the statistical precision of each.

The ZIP-income-based estimate of the late-career earnings gain from Social Security is both more precise and more conservative for our conclusions than the OCG-based estimates and the

[16]The modest magnitudes of these gradients reflects the substantial heterogeneity in individual resources within ZIP codes. A ZIP code contains residents spanning a wide range of individual incomes, so even large differences in individual earnings are associated with only modest differences in the average income of one's neighborhood. One might wonder about doing the projection the other way around, by regressing log individual earnings on log ZIP income. This "reverse regression" is not the right tool for the reconciliation: OLS projections are asymmetric, and the reverse gradient β_{R} is not the inverse of the "forward" gradient β unless $R^2 = 1$. In our data $R^2 \approx 0.07$: ZIP income explains only about 7 percent of the cross-sectional variation in individual late-career earnings, reflecting the substantial within-ZIP heterogeneity just noted.

[17]Throughout this section, we treat occupational income score gains as proxies for individual earnings gains. The occupational income score is actually the sex-by-race-by-region-specific average income in the individual's occupation from Collins and Wanamaker (2022), not individual earnings. To the extent that the occupational income score itself is an imprecise proxy for individual earnings, these estimates carry additional uncertainty beyond their standard errors.

[18]$\text{SE}(g) = |g|\sqrt{(\text{SE}_\delta/\hat{\delta})^2 + (\text{SE}_\beta/\hat{\beta})^2} = 0.282 \times \sqrt{(0.008/0.024)^2 + (0.003/0.085)^2} \approx 0.095$, where δ is the causal effect of Social Security on ZIP income and $g = \delta/\beta$ is the implied change in log earnings. This calculation treats the ZIP effect and HRS gradient as independent, since they come from different samples.

ZIP-home-price-based estimate. For these reasons, we use the ZIP-income-based estimate as our main estimate. But the other estimates provide useful corroborating evidence and, if anything, suggest that the ZIP-income-based estimate may be conservative.

Interpretation We interpret these comparisons as suggestive consistency checks on the key estimates. While the predicted effects are reasonably close to the direct estimates in each case, imprecision in the estimates—and so in the downstream predicted effects as well—limits the power of these consistency checks. So although the relatively close matches are consistent with the necessary conditions for unbiasedness of these mappings being satisfied (discussed below in Appendix D.2), they do not constitute strong evidence. At the same time, the comparisons reveal that the order-of-magnitude gap between the individual-earnings effects and the neighborhood-level effects is what one would expect given prevailing patterns of neighborhood sorting. In other words, this analysis suggests that the estimates of sons' gains in terms of late-career earnings and late-life neighborhood characteristics are quantitatively consistent.

D.2 Characterizing the potential biases in these mappings in a simple model

This subsection uses a simple structural model to derive conditions under which the mappings from ZIP gains to earnings gains and from earnings gains to ZIP gains are unbiased, and to characterize and quantify the potential biases when these conditions are violated. The analysis suggests that the key mapping from our relatively-precisely-estimated ZIP income gain to the implied late-career earnings gain likely understates the true earnings gain, and is unlikely to overstate the true gain by more than about 8 percent.

Model Let z and N denote a son's post-tax lifetime earnings and lifetime non-earnings net resources, respectively, both in expected present value dollars.[19] N aggregates net transfers from parents, governments, and other sources, as well as the net effect of moving costs and economies of scale from shared housing. Let ZIP denote a desirable characteristic of a son's late-life ZIP code. Suppose neighborhood sorting is governed by the structural demand equation

$$\mathrm{ZIP}_i = \alpha_D + \beta_D \mathrm{z}_i + \gamma_D \mathrm{N}_i + \varepsilon_{Di}, \tag{7}$$

where ε_D is mean independent of (z, N) (i.e., $E[\varepsilon_D \mid \mathrm{z}, \mathrm{N}] = 0$) and unaffected by father's Social Security, and $\beta_D, \gamma_D \geq 0$ (more resources increase demand for desirable ZIP characteristics).

The demand equation implies the following relationship between the causal effects of father's Social Security on ZIP characteristics, earnings, and non-earnings net resources:

$$\Delta\mathrm{ZIP} = \beta_D \Delta\mathrm{z} + \gamma_D \Delta\mathrm{N}, \tag{8}$$

where $\Delta\mathrm{X}$ is the causal effect of father's Social Security on variable X.

[19] Defining z and N this way is useful for constructing an informative bound, since it makes net transfers from the income tax and transfer system part of z, which we can measure, instead of part of N, which we cannot. It is not important that the z here match the z from our empirical work (estimating the causal effect $\Delta\mathrm{z}$ and the cross-sectional gradient β); redefining the z in our empirical work to be post-tax earnings as opposed to pre-tax earnings would simply rescale $\Delta\mathrm{z}$, β, and β_D in such a way as to leave $\beta\Delta\mathrm{z}$ and $\beta_\mathrm{D}\Delta\mathrm{z}$ unchanged.

Mappings from earnings effects to ZIP effects, and from ZIP effects to earnings effects Consider the mappings that we used above—from earnings effects to ZIP effects and from ZIP effects to earnings effects—through the lens of this simple model. The levels (as opposed to logs) versions of the mappings given by equations (5) and (6) above are:

$$\text{Inferred } \Delta \text{ZIP} = \beta \times \Delta \text{z} \tag{9}$$

$$\text{Inferred } \Delta \text{z} = \frac{1}{\beta} \times \Delta \text{ZIP}, \tag{10}$$

where β is the slope of the regression of ZIP characteristics on earnings in the cross-section (the levels version of equation (4)). We present the model in levels as opposed to logs for clarity; for the modest percentage changes relevant here, the dollar bound below carries over to a bound on the bias in $\Delta \log \text{z}$ of about the same relative magnitude (in fact, slightly smaller, given the concavity of the log over these changes).

Bias formulas Solving equation (8) for Δz gives the causal effect of father's Social Security on earnings as a function of the causal effects of father's Social Security on ZIP characteristics and on non-earnings net resources:

$$\Delta \text{z} = \frac{1}{\beta_D} \Delta \text{ZIP} - \frac{\gamma_D}{\beta_D} \Delta \text{N}. \tag{11}$$

Relative to this true relationship between the causal effects of father's Social Security, the mapping of ΔZIP to the inferred Δz (equation (10)) (i) uses β in place of β_D and (ii) omits the ΔN term.

Under what circumstances will the mapping in equation (10) yield the true causal effect of father's Social Security on earnings given by equation (11)? First, solve for the slope of the cross-sectional regression of ZIP on z in terms of the structural demand parameters and the joint distribution of earnings and non-earnings net resources:

$$\beta \equiv \frac{dE(\text{ZIP}|\text{z})}{d\text{z}} = \beta_D + \gamma_D \frac{dE(\text{N}|\text{z})}{d\text{z}}, \tag{12}$$

where $\frac{dE(\text{N}|\text{z})}{d\text{z}}$ is the slope of the conditional mean of N with respect to z in the cross-section. This uses the mean independence of ε_D, which ensures $dE(\varepsilon_D|\text{z})/d\text{z} = 0$. This is the levels analog of the log-log regression in equation (4).

Next, use equations (9), (11), and (12) to solve for the bias:

$$\begin{aligned}
\text{Bias in the } \Delta\text{ZIP to } \Delta\text{z mapping} &\equiv \text{Inferred } \Delta\text{z} - \text{Actual } \Delta\text{z} \\
&= \frac{1}{\beta} \times \Delta\text{ZIP} - \left(\frac{1}{\beta_D}\Delta\text{ZIP} - \frac{\gamma_D}{\beta_D}\Delta\text{N} \right) \\
&= \frac{\gamma_D}{\beta_D}\Delta\text{N} - \frac{\beta - \beta_D}{\beta_D \beta}\Delta\text{ZIP} \\
&= \frac{\gamma_D}{\beta} \left(\underbrace{\Delta\text{N}}_{\substack{\text{actual SS} \\ \text{effect on N}}} - \underbrace{\frac{dE(\text{N}|\text{z})}{d\text{z}}\Delta\text{z}}_{\substack{\text{cross-section-} \\ \text{implied effect on N}}} \right). \qquad (13)
\end{aligned}$$

The bias is the product of (i) the marginal propensity to sort on N relative to z, $\frac{\gamma_D}{\beta}$, and (ii) the gap between father's Social Security's actual causal effect on N and the effect on N that would make father's Social Security's causal effects on N and z match the slope of the cross-sectional relationship between N and z, given father's Social Security's causal effect on z.

The bias is zero if *either* of the following sufficient conditions hold: (i) N does not affect neighborhood choice ($\gamma_D = 0$) or (ii) father's Social Security's causal effects on N and z align with the cross-sectional relationship between these variables:

$$\Delta\text{N} = \frac{dE(\text{N}|\text{z})}{d\text{z}} \times \Delta\text{z}. \qquad (14)$$

We call this the *on-gradient condition*: father's Social Security's causal effects on compliers lie along the cross-sectional slope; father's Social Security shifts compliers' earnings z and non-earnings net resources N along the cross-sectional gradient of N with respect to z. This is a sufficient but not necessary condition for the mapping from ZIP characteristics to earnings to be unbiased. An even stronger condition that would imply the on-gradient condition (and so guarantee unbiasedness of the mapping) is that compliers sort across neighborhoods conditional on earnings in the same way as the broader population, and that father's Social Security shifts them along this common gradient.

Parallel steps can be used to characterize the bias in the mapping in the other direction, from earnings to ZIP characteristics: the gap between the inferred effect on ZIP characteristics (equation (9)) and the true effect (equation (8)):

$$\begin{aligned}
\text{Bias in the } \Delta\text{z to } \Delta\text{ZIP mapping} &\equiv \text{Inferred } \Delta\text{ZIP} - \text{Actual } \Delta\text{ZIP} \\
&= \gamma_D \left(\frac{dE(\text{N}|\text{z})}{d\text{z}}\Delta\text{z} - \Delta\text{N} \right). \qquad (15)
\end{aligned}$$

The bias in the mapping from earnings to ZIP gains (equation (15)) is proportional to and opposite-signed from the bias in the mapping from ZIP to earnings gains (equation (13)):

$$\text{Bias in the } \Delta\text{z to } \Delta\text{ZIP mapping} = -\beta \times \text{Bias in the } \Delta\text{ZIP to } \Delta\text{z mapping}. \qquad (16)$$

Both mappings therefore share the same sufficient conditions for unbiasedness: The bias is zero if *either* (i) N does not affect sorting ($\gamma_D = 0$) or (ii) father's Social Security's causal effects on N and z align with the cross-sectional relationship between these variables (the on-gradient condition).

One way the on-gradient condition could fail is if (i) the cross-sectional ZIP-on-earnings gradient among compliers differs from that among the broader population (β) and (ii) father's Social Security shifts compliers along their own gradient rather than β. One potential reason this could happen is that compliers had unusually back-loaded earnings profiles, which may have led them to sort into worse neighborhoods than others with similar late-career earnings. This would make the cross-sectional ZIP-on-earnings gradient among compliers smaller than β, so dividing the ZIP estimate by β would lead the key mapping from ZIP gains to earnings gains to understate the implied earnings gain (i.e., to be conservative on the key conclusions).

Bounding the bias That the predicted effects are reasonably close to the direct estimates in each case, as discussed in Appendix D.1, is suggestive that at least one of the sufficient conditions for unbiasedness approximately holds. But the imprecision of the estimates means these reasonably close matches do not alone constitute strong evidence of unbiasedness. So in this section, we use economic logic and our estimates to bound the potential bias of the key ZIP-to-earnings mapping. We argue that the bias likely works against our conclusion of large earnings gains, i.e., that our ZIP-to-earnings conversion likely understates the true earnings gain, and we derive a conservative upper bound on the possible overstatement of the earnings gain.

Recall the bias in the ZIP-to-earnings mapping (equation (13)):

$$\text{Bias in the } \Delta\text{ZIP to } \Delta\text{z mapping} = \frac{\gamma_D}{\beta}\left(\Delta\mathrm{N} - \frac{dE(\mathrm{N}|\mathrm{z})}{d\mathrm{z}}\Delta\mathrm{z}\right). \tag{13}$$

It seems plausible that the term in parentheses is negative, that is, that father's Social Security increased earnings relative to non-earnings net resources more than the cross-sectional relationship between these variables. The rationale is that father's Social Security caused sufficiently large earnings gains that the associated effect on non-earnings net resources seems unlikely to be large enough to match the cross-sectional gradient between these variables. In that case, the bias works against our key results: The ZIP-to-earnings mapping understates the true earnings gain.

But in the absence of direct estimates of the causal effect of father's Social Security on non-earnings net resources, a safer approach is to use a more conservative bound that does not require any assumptions about the relative causal effects of father's Social Security on earnings versus non-earnings net resources. One such approach is the following. Under the natural assumption of a positive cross-sectional relationship between non-earnings net resources and post-tax earnings ($dE(\mathrm{N}|\mathrm{z})/d\mathrm{z} \geq 0$)—since a key component of N is net transfers from parents, and better-off families both earn more and leave larger transfers—an upper bound on the bias is:

$$\begin{aligned}\text{Bias in the } \Delta\text{ZIP to } \Delta\text{z mapping} &\leq \frac{\gamma_D}{\beta} \times \Delta\mathrm{N} \\ &\leq \frac{\gamma_D}{\beta_D} \times \Delta\mathrm{N}. \end{aligned} \tag{17}$$

The first line follows because it drops the second term in the bias equation (equation (13)), which is non-positive. The second line follows from equation (12) (the slope of the cross-sectional conditional mean $E(\mathrm{ZIP}|\mathrm{z})$ in terms of the demand function parameters), the assumption that $dE(\mathrm{N}|\mathrm{z})/dz \geq 0$, and the assumption that $\Delta\mathrm{N} \geq 0$. The first of these assumptions implies

$$\beta = \beta_D + \gamma_D \frac{dE(\mathrm{N}|\mathrm{z})}{dz} \geq \beta_D, \tag{18}$$

so $\gamma_D/\beta \leq \gamma_D/\beta_D$. Multiplying this inequality by $\Delta\mathrm{N}$ preserves its direction given the assumption that $\Delta\mathrm{N} \geq 0$.[20]

Under the natural benchmark assumption of equal marginal propensities to consume neighborhood characteristics out of lifetime earnings and lifetime non-earnings net resources ($\beta_D = \gamma_D$, i.e., neighborhood demand depends only on total lifetime resources, $(\mathrm{z} + \mathrm{N})$, not their composition), the bound becomes:

$$\begin{aligned} \text{Bias in the } \Delta\text{ZIP to } \Delta\text{z mapping} &\leq \frac{\gamma_D}{\beta_D} \times \Delta\mathrm{N} \\ &\leq \Delta\mathrm{N}. \end{aligned} \tag{19}$$

This bound, which is in expected present value dollars, depends solely on the causal effect of father's Social Security on non-earnings net resources. It is independent of the unknown parameters (of the demand function, β_D and γ_D, and of the cross-section, $dE(\mathrm{N}|\mathrm{z})/dz$).

Numerical bound Although we lack the data to estimate the causal effect of father's Social Security on non-earnings net resources, we can use economic logic to bound the likely effect. Given that earnings are defined as after-tax earnings, the key component of non-earnings net resources is likely net transfers from one's parents.[21] Likely a loose upper bound on $\Delta\mathrm{N}$ comes from assuming 100 percent pass-through of father's Social Security benefits to greater net transfers from recipients to their children. If Social Security raised parents' net downstream transfers by the full amount of the benefits received (a marginal propensity to transfer out of Social Security of one), and the additional transfers were split equally across recipients' children, then with our estimated expected present value of Social Security benefits of \$47,000 and an average of 4.6 children per father, $\Delta\mathrm{N}$ = \$10,200 per son.[22]

In this case, equation (19) implies that the mapping from ZIP gains to earnings gains would overstate the expected present value of sons' lifetime earnings gains by no more than about \$10,200 per son. Relative to the full-career earnings gain implied by the mapping (about \$130,000 per son;

[20] The case $\Delta\mathrm{N} \geq 0$ is the relevant one for bounding the possible *overstatement* of the earnings gain. (If instead $\Delta\mathrm{N} < 0$, the first line above would be negative, so the mapping would understate the earnings gain and already be conservative for our conclusions.) The numerical bound below assumes $\Delta\mathrm{N} \geq 0$.

[21] Recall that non-earnings net resources also includes the net effect of moving costs and economies of scale from shared housing. Given that father's Social Security increased migration by sons, it tended to reduce their non-earnings net resources through these two channels, by increasing their present value moving costs and decreasing their housing returns to scale. So these factors tend to make the key ZIP-to-earnings conversion understate the true earnings gain.

[22] This likely overstates the change in net downstream financial transfers significantly. In addition to assuming a 100 percent marginal propensity to transfer out of Social Security benefits, it ignores the possibility of crowd-out of other transfers (e.g., from private charity or means-tested government transfers) and of transfers to parties other than one's children (e.g., to charities or other family members). Moreover, not only does this calculation likely overstate the bias from the first term in the bias formula (equation (13)), it also ignores the second term in the bias formula, which further overstates the bias.

see Table 9), even this upper bound implies a relative bias of just 8 percent. Because the expected present value of the lifetime earnings gain is proportional to the late-career gain g, this relative bound is the same whether the gain is measured in present value dollars or as a proportional gain; the common scaling cancels in the ratio of the bias to the gain. So the inferred late-career gain of about 28 percent (equation (6)) is likewise overstated by no more than about 8 percent, implying a true gain of at least roughly 26 percent.[23]

Discussion This analysis considers only financial determinants of neighborhood choice. In reality, father's Social Security might also affect neighborhood choice through non-financial channels, for example by affecting the propensity of recipients' children to sort to different neighborhoods late in life given their higher likelihood of moving earlier in life. While it is difficult to know exactly how such factors might bias these mappings, the broad consistency of the different results—which depend on a variety of different data and assumptions—is suggestive that the different analyses are capturing different aspects of the same overall story without too much bias.

E Details of the implications calculations (Section 6)

Section 6 reports several quantities derived from the estimated late-life ZIP effects and occupational income score effects: the late-career average earnings gain, the full-career average earnings gain, the expected present value (EPV) of lifetime earnings gains, the ratio of children's gains to recipients' Social Security benefits, and fiscal effects. Here we describe these calculations, including how sampling uncertainty in the key estimates propagates to the "downstream" results. Table A.14 collects the key assumptions, their baseline values, and their sources. All monetary amounts are in 2020 dollars, and EPVs are as of 1950.

Late-career earnings gain, g We estimate g in three different ways. The occupational income score-based estimate uses our estimates of the effects of Social Security on occupational income scores in 1962 and 1972 using OCG data. The key assumption is that occupational income score gains translate one-for-one into earnings gains. We use a "central tendency estimate" from the four underlying estimates of occupational income score gains (for each of two years and each of two sample definitions; see Table 5) of $\hat{g} = 35\%$ (SE $= 17\%$). The ZIP income-based estimate divides the estimated effect of Social Security on the log of late-life ZIP mean income (0.024, SE 0.008; see Table 4) by the estimated slope from a cross-sectional regression of the log of ZIP income on the log of late-career earnings in the HRS (0.085, SE 0.003; see Table A.7). See Appendix D for details and the underlying theory. This yields $\hat{g} \approx 28.2\%$ (SE $\approx 9.5\%$; see Appendix D). The ZIP home price-based estimate is analogous to the ZIP income-based estimate. It divides the estimated effect of Social Security on ZIP home prices (0.043, SE 0.012; Table 4) by the estimated slope from a cross-sectional regression of the log of ZIP home prices on the log of late-career earnings using the HRS (0.092, SE 0.005; Table A.7). This yields $\hat{g} \approx 46.7\%$ (SE $\approx 13.3\%$; see Appendix D). We use the ZIP income-based estimate as the baseline since it is more conservative and more precise than the others, but we report results for the other two as well (see Table 9).

[23] On the log scale the bound is slightly tighter—about 7 rather than 8 percent—because the log is concave. The correction is small because the change is small.

Full-career earnings gain and EPV per son We translate the late-career gain into a full-career average gain by linearly interpolating the following earnings-gain profile of a son born in 1910 over his working life (ages 20–65): no gain at the first job (1930) or in 1950, the full gain g in 1962 and 1972, and the gain held at g thereafter. This profile mirrors the occupational income score estimates, which show no effect on first job or job in 1950 and large effects in 1962 and 1972 (Table 5). The baseline earnings path on top of which these gains accrue is median annual earnings of male four-quarter workers from the Social Security Bulletin Annual Statistical Supplements (Social Security Administration, 1965, 1975), converted to 2020 dollars. Annual gains are discounted to 1950 at a 3% real rate and weighted by survival from the SSA 1910 male cohort life table, conditional on surviving to 1930 (to match our sampling frame). For the baseline estimate, $\hat{g} = 28.2\%$, this implies a full-career average gain of 11.9% and an EPV gain of about $130,000 per son. The implied ratio of the late-career to the full-career average gain is about 2.4, close to the late-to-full-career ratio in Nakamura et al. (2022) of about 2.5.

Collective gains to recipients' children per recipient, ΔEPV$(\mathrm{z_k})$ Per-son gains are scaled to the recipient-family level by the average number of sons per recipient family, 2.3. The baseline per-son EPV of $130,000 thus implies a collective gain of about $299,000 per family. We assume that gains to daughters from their father's Social Security benefits are zero, since the estimates for daughters tend to be small.[24]

Social Security benefits and recipients' earnings Social Security benefits and recipients' earnings come from simulating the life cycle of a married couple in which the husband is in the 1877 birth cohort, the wife is three years younger (1880 cohort), and the husband earned the median wage in covered employment and did not change coverage type during his career. Benefits are based on the statutory primary-insurance-amount (PIA) formulas in force each year, with the couple receiving 150% of the worker's PIA while both are alive and a lone survivor receiving 75%.[25] Expected present values discount future financial flows using a discount rate of 3% and survival probabilities from Human Mortality Database cohort tables, conditional on the couple being alive in 1930 (ages 53 and 50). We index the two Social Security regimes by the calendar year in which the worker's employment first contributes to eligibility (the "coverage-start year"). The pre-reform regime is 1950-covered employment, which first counts toward eligibility in 1951, and the post-reform regime is 1935-covered employment, which first counts in 1937. These coverage-start years are 14 calendar years apart, but they should not be confused with benefit-claiming dates. Under the age-65 rule, the 1877 worker's *earliest* claiming age is 65 (in 1942) under the post-reform regime and 75 (in 1952) under the pre-reform regime—since 1950-covered work first counts in 1951 and

[24] As we discuss in Section 6, this does not imply that recipients' daughters gained little from Social Security. For a husband-and-wife couple, the effect of both of their fathers having earlier eligibility for Social Security is the sum of our estimates for sons and daughters. Given that the vast majority of the children of fathers affected by our instrument are married, we interpret our estimates as showing that for husband-and-wife married couples, the Social Security of the husband's father is a more important driver of their migration and living standards than the Social Security of the wife's father. This is in keeping with the idea that sons and daughters-in-law play a different role in family old-age support arrangements than daughters and sons-in-law do, with the role of sons and daughters-in-law being more substitutable with Social Security benefits, perhaps because such couples tend to provide more resource-intensive support whereas daughters and sons-in-law couples provide more time- or care-intensive support.

[25] The survivor share later rose to 100% by 1973, which we examine in a robustness test (Table A.15).

the six required quarters are accrued only by mid-1952—so the reform amounts to a 10-year-earlier earliest eligibility age for this cohort. Let $b(e, p)$ denote the EPV of benefits under coverage regime $e \in \{1937, 1951\}$ and retirement-behavior profile p. Recipients' retirement behavior under 1937 eligibility (profile "37") follows the labor force participation-by-age profile observed in 1930, reduced by 25% after the earliest claiming age under 1937 eligibility. Behavior under 1951 eligibility (profile "51") is a blend of the 1930 and "37" profiles, weighted by the share of the no-Social-Security-to-1937 benefit increase already delivered by 1951 eligibility. The 25% reduction is our baseline labor-supply assumption, but we test robustness to alternatives (Table A.15). The simulation yields three quantities:

- The total change in Social Security benefits caused by the reform: $\Delta\text{EPV(SS)} = b(1937, 37) - b(1951, 51) \approx$ \$46,500. This reflects both the "mechanical effect" of the change in the budget constraint and the "behavioral effect" from recipients' behavioral responses of earlier retirement due to Social Security. We use this as the denominator of the ratio of children's gains to recipients' Social Security benefits to be conservative; using the mechanical effect alone would increase the ratio of children's gains to recipients' benefits. We also use gross benefits as opposed to net benefits (net of Social Security payroll taxes), to be conservative. Social Security taxes were modest in the early years of the program, as mentioned in Section 2.

- The "mechanical effect" of the reform, holding behavior fixed at the pre-reform profile, $\text{ME} = b(1937, 51) - b(1951, 51) \approx$ \$40,600.

- Recipients' earnings response to the reform from earlier retirement, $\Delta\text{EPV}(\mathrm{z_r}) = \text{EPV}(\mathrm{z_r}(37)) - \text{EPV}(\mathrm{z_r}(51)) \approx -$\$35,000.

The gap between the total and the mechanical change in Social Security benefits, $\Delta\text{EPV(SS)} - \text{ME} = b(1937, 37) - b(1937, 51) \approx$ \$5,900 is the "behavioral effect," the extra benefits recipients collect because earlier eligibility leads them to claim earlier. The sum of the behavioral effect and recipients' earnings response, $\Delta\text{EPV}(\mathrm{z_r})$, is the recipient-related portion of the fiscal externality discussed below. We assume a "first stage" of $\phi = 1$, treating a 10-year difference in *predicted* earliest eligibility as a 10-year difference in *actual* eligibility. Because the reduced-form child gain is divided by ϕ to express it per actually-treated recipient, $\phi = 1$ gives a conservative (lower) bound on the structural child gain, and hence on the key downstream results (including child gains relative to Social Security benefits and the degree of self-financing).[26]

Fiscal externalities, government cost, and MVPF The recipient- and child-side fiscal externalities are

$$\text{FE}_r = -\overline{\tau}_r \, \Delta\text{EPV}(\mathrm{z_r}) + \big(\Delta\text{EPV(SS)} - \text{ME}\big) \approx \$12{,}600, \tag{20}$$

$$\text{FE}_k = -\overline{\tau}_k \, \frac{\Delta\text{EPV}(\mathrm{z_k})}{\phi} \approx -\$75{,}500, \tag{21}$$

[26]We stress that this conservatism pertains to these downstream per-recipient quantities, not to a structural reinterpretation of the reduced form. Our reduced-form estimates are effects of a 10-year difference in *predicted* eligibility, and we do not interpret them as the structural effect of a 10-year difference in *actual* eligibility, which—for a true first stage below one—would be larger.

where $\Delta\text{EPV}(\mathrm{z_k}) \approx$ \$299,000 is the pretax collective child gain. $\text{FE}_r > 0$ is the net cost to the government from forgone tax revenue due to recipients' lower earnings and from the additional Social Security benefits due to recipients' earlier claiming. $\text{FE}_k < 0$ is the net saving to the government from recipients' children paying more in taxes due to their higher earnings. We estimate the average net tax rates on recipients' earnings losses ($\overline{\tau}_r$) and sons' earnings gains ($\overline{\tau}_k$) based on historical income, payroll, consumption, property, and unemployment insurance taxes, as well as future OASI benefits. Our main estimates are 18.8 percent for recipients and 25.3 percent for sons. See Section E.1 for details. The government's net marginal cost is

$$\text{GMC} = \text{ME} + \text{FE}_r + \text{FE}_k \approx -\$22{,}300, \tag{22}$$

so the net cost per dollar of the mechanical benefit increase is $\text{GMC}/\text{ME} \approx -0.55$.[27] *If* all recipients and their children benefited from the reform (i.e., had non-negative willingness to pay for it), then the marginal value of public funds (MVPF)—the ratio of the willingness to pay for the policy by all affected parties to the government's marginal cost, WTP/GMC—would be ∞: The reform would benefit recipients and their children at negative cost to the government.[28] If we ignored the fiscal externality from recipients' children's greater earnings, $\text{FE}_k = 0$, then the net cost per dollar of the mechanical benefit increase (without the intergenerational effect) would be $\text{GMC}' = (\text{ME} + \text{FE}_r)/\text{ME} \approx 1.31$. That is, if not for the effects of Social Security on recipients' children, each dollar of the mechanical increase in Social Security benefits would have cost the government about \$1.31.

Baseline values and uncertainty The chain of calculations is affine in g, so each downstream quantity is evaluated at the point estimate and at $g \pm 1.96 \times \text{SE}$ to obtain its 95% confidence interval. Of the components of the key quantities, only the components pertaining to recipients' children's gains depend on g; the recipient-related quantities (e.g., ΔEPV(SS), ME, FE_r, and recipients' earnings) are held fixed and carry no sampling uncertainty here.

Robustness Table A.15 shows tests of the robustness of the results to changes in several key assumptions, including the labor supply responses of recipients and their children to Social Security; policy parameters, including the average marginal tax rate on recipients' and their children's earnings responses to Social Security, the level of survivors benefits, and the level of primary benefits; and sons' earnings responses. These are "stress tests" that aim to determine how far different parameters must be changed from their baseline levels before altering the key conclusions; they are not aiming to be plausible specifications (most are not). The conclusion that recipients' children's earnings gains were considerably larger than the associated difference in Social Security benefits is highly robust across all specifications, ranging from 4.3 to 6.5 times the associated difference

[27] Several of our assumptions work against the finding of negative cost to the government. For example, GMC is right around its maximum for the 1877 recipient cohort, so using this cohort for these calculations works against finding a low or negative GMC. Likely most important is our assumption of a "first stage" of one, which likely significantly overstates the true first stage and so the cost components of GMC (ME and FE_r) relative to the savings component (FE_k).

[28] This follows the convention of assigning an MVPF of infinity to a reform that benefits affected parties while saving the government money. As discussed in Section 6, however, it is likely that at least some recipients or recipients' children were harmed by the reform, due to the induced changes in family transfers.

in benefits. The conclusion that these Social Security expansions increased net family earnings is also extremely robust, with the increase in net family earnings ranging from 3.8 to 5.7 times the associated difference in Social Security benefits.

The conclusion that these Social Security expansions paid for themselves is somewhat less robust, given the conservative assumption of a first stage of one (i.e., differences in predicted eligibility translate one-for-one into differences in actual eligibility), with point estimates of the net cost to the government per dollar of mechanical effect ranging from $-\$0.57$ to \$0.02. The three specifications that cause the point estimate of the net cost to be largest (least negative or most positive) are: 50 percent reduction in recipient labor force participation due to Social Security (versus a baseline of 25 percent), in which case the net cost is $-\$0.08$; sons' estimated earnings response scaled by 0.75, in which case the net cost is also $-\$0.08$; and Social Security benefits scaled up by 50 percent, in which case the net cost is \$0.02. Despite these specifications being fairly extreme, they do not cause the net cost to be very large but roughly zero. So the self-financing result is fairly robust to changes in modeling assumptions, albeit with considerable statistical uncertainty.

Caveats The reported confidence intervals capture only sampling uncertainty in the estimated treatment effects and, for the ZIP-based estimates, the HRS gradient. They do not reflect additional uncertainty from modeling assumptions such as the interpolation scheme, the baseline earnings path, the benefits simulation, and the occupational-income-score-to-earnings mapping, so the true uncertainty is larger. The conclusion that these Social Security expansions increased total family earnings (i.e., that children's earnings gains exceeded recipients' earnings losses) is highly robust, holding across the 95% confidence interval even under the most conservative specification. But the conclusion that these Social Security expansions paid for themselves (negative net cost to the government) does not hold at the most conservative bound of the most conservative specifications.

E.1 Historical average net tax rates on earnings responses

The fiscal calculations require the change in government net revenue associated with recipients' earnings losses and their sons' earnings gains. Because the two responses occur under different tax regimes, we estimate a separate average net tax rate for each. Our preferred estimates are 18.8 percent for recipients and 25.3 percent for sons (see Table A.16). These estimates include taxes on income, payroll, unemployment insurance, consumption, and property, and, for sons, subtract the associated increase in future OASI benefits.

Tax liability changes and representative households Our object is the change in total net tax liability per dollar of the earnings change, not the statutory tax rate on the last dollar of earnings. Let W_t denote median annual earnings of male four-quarter workers in year t, $F_t(z;x)$ federal income tax liability at earnings z for a household with characteristics x, $P_t(z)$ combined employee and employer OASDI and Hospital Insurance payroll taxes, $U_t(z)$ federal and state unemployment insurance payroll taxes, a_t the national average effective state and local personal income tax rate, and s_t the national average marginal state income tax rate. For recipients, the relevant response is an extensive margin change from working at the median wage to retirement. The annual earnings-side tax rate is therefore

$$\tau_{rt}^{E} = \frac{[F_t(W_t; x_r) - F_t(0; x_r)] + [P_t(W_t) - P_t(0)] + [U_t(W_t) - U_t(0)]}{W_t} + a_t. \tag{23}$$

We use an average effective state tax rate here (rather than an average marginal state tax rate) because the earnings change spans the full interval from zero to W_t. We construct a_t as state and local personal income tax receipts divided by personal income excluding transfers in BEA NIPA Tables 3.4 and 2.1 (Bureau of Economic Analysis, 2026). This is better aligned with the extensive margin than a marginal rate, although it remains an aggregate national proxy rather than the exact liability of the representative household. The representative recipient household is the one in our benefit simulation: a married couple filing jointly, with a husband born in 1877, a wife born in 1880, one male earner, no dependents, wage income only, and the standard deduction. We account for both spouses' additional age exemptions beginning in 1948.

For sons, let g_t denote the proportional earnings gain implied by the profile described above: zero in 1950, increasing linearly to $\hat{g} = 28.2$ percent in 1962, and constant thereafter. The annual earnings-side rate before accounting for future OASI benefits is

$$\tau_{kt}^{E} = \frac{[F_t((1+g_t)W_t; x_k) - F_t(W_t; x_k)] + [P_t((1+g_t)W_t) - P_t(W_t)]}{g_t W_t} + \frac{[U_t((1+g_t)W_t) - U_t(W_t)]}{g_t W_t} + s_t, \tag{24}$$

which uses the national average marginal state income tax rate, s_t, since this is an intensive margin change above baseline earnings W_t. The baseline son household is married filing jointly, with a husband born in 1910, a wife three years younger, one male earner, two dependents, wage income only, and the standard deduction.[29]

We aggregate annual rates using the present value of the corresponding annual earnings response as weights:

$$\overline{\tau}_h = \frac{\sum_t \omega_{ht}\tau_{ht}}{\sum_t \omega_{ht}}, \qquad h \in \{r, k\}, \tag{25}$$

where ω_{ht} is the absolute present value earnings change in year t. The weights use the same survival probabilities, CPI conversion, and 3 percent real discount rate that we use throughout the paper.

The recipient weights come from the Social Security benefits simulation discussed above. That simulation computes expected earnings under the post-reform retirement profile (1937 eligibility) and the counterfactual pre-reform profile (1951 eligibility); their difference is the recipient earnings response. For each year, the tax calculation applies the corresponding difference in labor force participation to median earnings, survival, CPI, and discounting before aggregating. Thus the

[29] Assuming that sons' families have two dependents is a conservative assumption that tends to understate sons' net tax rate and so overstate the net cost of their parents' Social Security. The son is born in 1910 and experiences earnings gains from 1951–1975 (age 41–65), with most of the gains in his mid-to-late 50s and early 60s (due to the back-loaded nature of the gains). By that time, most of sons' children would have moved out and no longer been dependent. The effect is minor in any case: Dependent exemptions are mostly inframarginal for an intensive margin rate, since they shift taxable income down in parallel at both W and $(1+g)W$, so the average net tax rate on earnings changes is unchanged unless the increment crosses a bracket boundary.

annual tax weights sum to the recipient earnings change.

Federal income, payroll, state income, and unemployment insurance taxes For 1942–59, we reconstruct federal liability from the complete annual bracket schedules, personal and age exemptions, and nonitemizer deductions, verified against annual Internal Revenue Service forms and instructions (Urban-Brookings Tax Policy Center, 2023; Tax Foundation, 2025; Internal Revenue Service, 2026). The reconstruction accounts for the 1942–43 earned-income exclusion from the normal tax base, the 1943 Victory Tax, the standard deduction beginning in 1944, postwar across-the-board liability reductions, married couple income splitting beginning in 1948, and additional age exemptions. For 1960–75, we calculate federal liability using NBER TAXSIM 35, which captures temporary surtaxes, low-income allowances, and credits (Feenberg and Coutts, 1993; National Bureau of Economic Research, 2026).

We apply the statutory employee and employer OASDI rates, and the employee and employer Hospital Insurance rates beginning in 1966, to covered earnings up to each year's contribution and benefit base (Social Security Administration, 2026b,a). Including both shares is appropriate for the fiscal externality, which measures government revenue per dollar of reported wage earnings. Applying the annual ceiling is important: Median earnings exceed the ceiling in several years, so the son's incremental OASDI/HI payroll rate is zero in those years. In 1974–75, by contrast, the entire incremental earnings range is below the ceiling and the combined employee and employer OASDI/HI payroll rate on that increment is 11.7 percent.

For sons, we use the national income-weighted average marginal state income tax rate constructed by Barro and Redlick (2011). Their series combines historical state tax codes with estimated state income distributions and accounts for federal-state deductibility interactions. The state component adds 1.9 percentage points to the son's rate before benefit accrual. The BEA average effective rate described above adds 0.3 points to the recipient rate. Because our benchmark households are nonitemizers, federal deductibility does not otherwise alter our household-level federal calculations.

We add state UI contributions on the recipient extensive margin using the Department of Labor's annual contributions-to-total-wages series, plus net Federal Unemployment Tax Act (FUTA) contributions under the historical rate and wage base (U.S. Department of Labor, 2026, 2018; Whittaker, 2016). This adds 1.7 percentage points to the recipient earnings-side rate. The sons' earnings increments are above the federal UI wage base in every positively weighted year. Historical state wage-base data also show that their baseline earnings exceeded the state base in nearly all relevant state-year cells, so we round their UI add-on to 0.0 percent. Applying the aggregate state contribution rate per dollar of total wages instead would add 1.0 percentage point and raise the sons' net tax rate to 26.3 percent. We regard that as an upper proxy rather than the main estimate because it assigns to a high-wage incremental dollar the tax rate collected on all wages, including the first dollars below each state's taxable maximum. Experience rating adds further uncertainty, but it does not make the aggregate average rate a better marginal rate for this earnings range.

Future Social Security benefits Higher covered earnings increase sons' future OASI benefits, so taxes alone overstate the net fiscal rate. For a man born in 1910 and retiring in 1975, we cap earnings at each year's contribution and benefit base, use the highest 19 of the 24 earnings years from 1951–74 (the applicable five dropout years), and apply the 1975 average-monthly-wage PIA formula (Social Security Administration, 2023). We value the resulting real benefit increment using the same mortality and discounting conventions as the earnings gain. The present value of the worker's added benefits is 2.1 percent of the present value of his earnings gain:

$$\overline{\tau}_k = \overline{\tau}_k^E - \frac{\Delta \text{EPV}(B_k)}{\Delta \text{EPV}(z_k)} = 20.6\% - 2.1\% = 18.5\%. \tag{26}$$

This 2.1 percent is a lifetime fiscal benefit accrual rate, not the usual annual replacement rate at retirement. It is small principally because only 13.5 percent of the present value earnings gain increases covered earnings; most of the gain is above the Social Security taxable maximum. The five dropout years and discounting of benefits received after age 65 reduce the offset further. As an upper bound, we assume that a nonworking wife receives an incremental spouse benefit while both spouses are alive and a survivor benefit thereafter. This raises the benefit offset to 4.9 percentage points. It is an upper bound because a wife with her own earnings record may receive little or no incremental spouse benefit.

We do not subtract recipients' additional Social Security benefits from their tax rate. The recipient-side fiscal externality already includes the additional benefits induced by earlier retirement and claiming; subtracting them again through the tax rate would double count that cost.

Consumption and property taxes Our preferred specification assumes that all cash remaining after income and employee payroll taxes is eventually consumed and, in the absence of evidence on timing, applies the tax contemporaneously. We use the annual national-accounts consumption tax rate from McDaniel (2007), which includes sales and excise taxes, customs duties, taxes on specific services, and the household portion of recurrent property taxes. For 1942–49, we extend the series using analogous tax receipts and personal consumption expenditures in the BEA National Income and Product Accounts, spliced to McDaniel's 1950 level (Bureau of Economic Analysis, 2026). If c_t is the broad tax rate relative to pre-tax consumption and q_{ht} is the income and employee payroll tax rate on cash earnings, the added revenue per dollar of earnings is $(1 - q_{ht})c_t/(1 + c_t)$. This adds 6.6 percentage points for recipients and 6.8 points for sons.

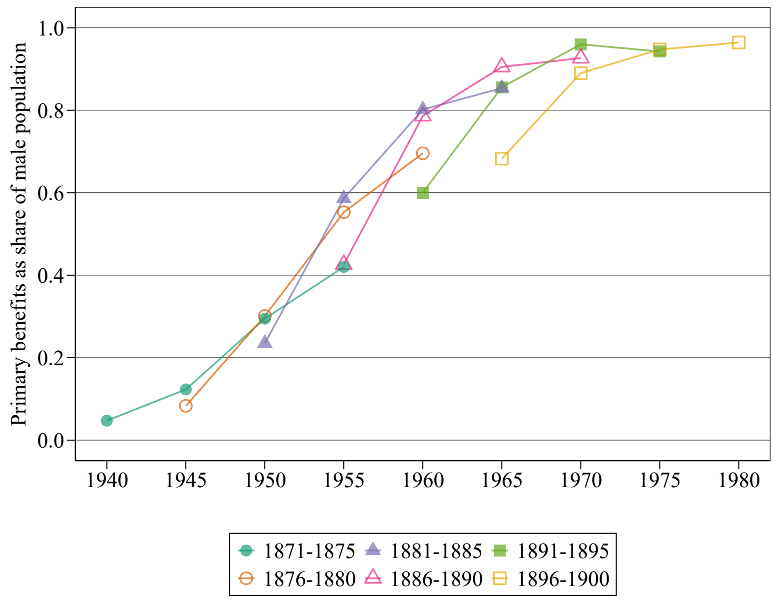
This "full consumption" treatment is uncertain but best captures the desired full, net fiscal effect. Saving changes the timing, but not necessarily the ultimate taxability, of the earnings gain. Immediate consumption ignores discounting during the delay but also applies the generally lower historical tax rates rather than later rates, so the net direction of the timing approximation is ambiguous. Bequests defer consumption to heirs but, like saving, bear on its timing more than its ultimate taxability. Reasons taxable consumption could be below total consumption include that some spending escapes these indirect taxes, such as tax-favored and charitable spending. We test robustness to excluding consumption taxation entirely (see Table A.15).

Results Table A.16 summarizes the results, including the contributions of each component to the overall net tax rates on the earnings changes of recipients and their sons. For recipients, the preferred net tax rate is 18.8 percent, with about 12.3 percent from taxes on earnings—federal income and OASDI/HI payroll taxes, state and local income taxes, and unemployment insurance taxes—and about 6.6 percent from broad consumption and property taxes. For recipients' sons, the preferred net tax rate is 25.3 percent, with about 18.5 percent from taxes on earnings (20.6 percent gross taxes before netting out 2.1 percent for OASI benefit accrual) and about 6.8 percent from broad consumption and property taxes.

Robustness Excluding the broad consumption and property taxes—the most speculative component—leaves a net tax rate of 12.3 percent for recipients and 18.5 percent for sons. Replacing the worker-only OASI benefit offset with the family (spouse/survivor) upper bound lowers the son's net tax rate from 25.3 to 22.5 percent. We test the robustness of the results to both of these changes; see Table A.15. Household filing assumptions matter more for sons than for recipients: Before the OASI offset and consumption taxes, the son earnings-side rate is 22.1 percent for a single filer and 23.2 percent for a married household in which the wife earns 50 percent of the male median wage, compared to 20.6 percent in the married one-earner baseline. The largest remaining approximations are the use of national rather than household-specific state and consumption tax rates and the timing and incidence of consumption. Local wage taxes and income-tested transfer phaseouts are omitted, but the effect is likely to be small and to work against our main conclusions. For a near-median married male earner, and before the 1975 introduction of the Earned Income Tax Credit, baseline means-tested transfers are close to zero. Because any phaseout would raise the net tax rate, omitting it makes the son estimate conservative as an input to the net government cost.

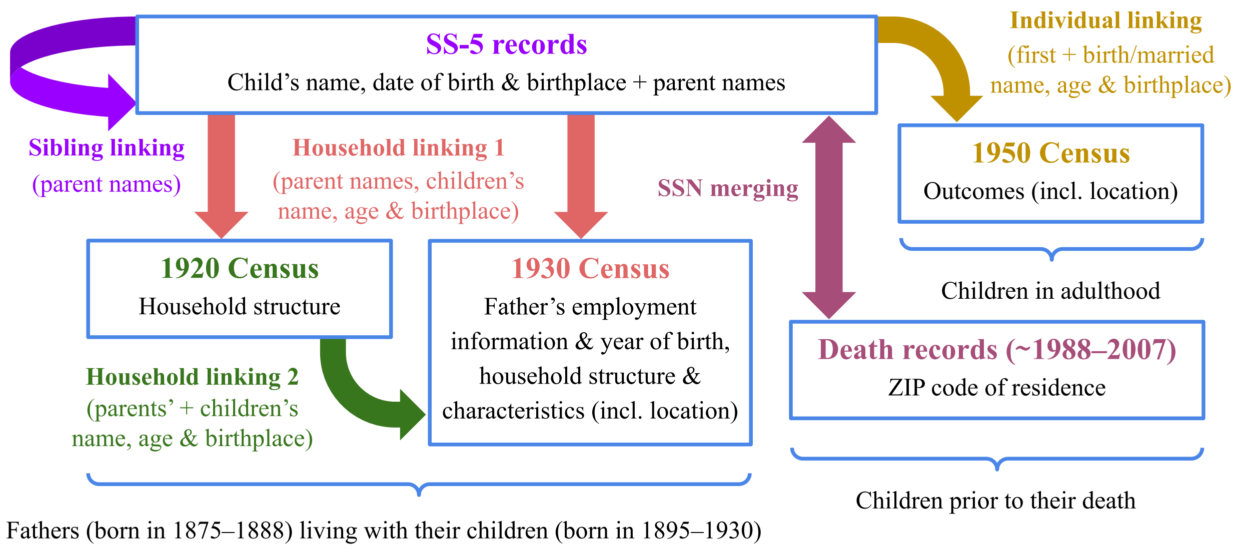
F Appendix Figures and Tables

Figure A.1: Receipt of retirement benefits by birth cohort and year



Notes: This figure shows share of men receiving retirement benefits in a specified year, separately by 5-year birth cohort. Each cohort is shown at ages 65–69, 70–74, 75–79, and 80–84. Data on payments are from the *Social Security Bulletin*, data on male population by age and year are from United States Bureau of the Census (2004).

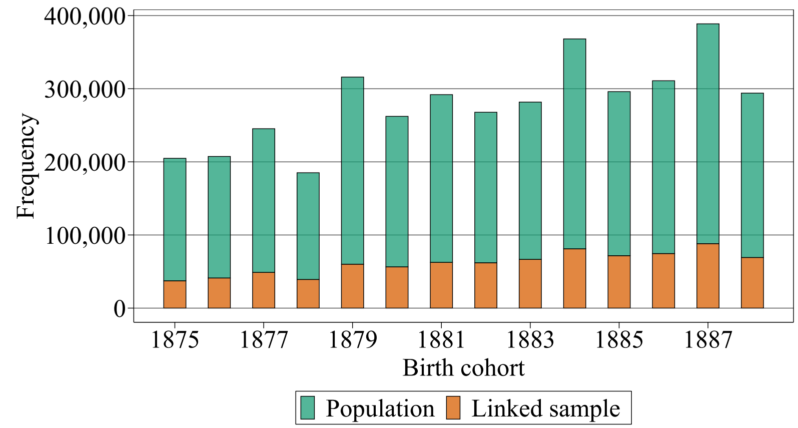
Figure A.2: Overview of Numident death-census linked sample construction



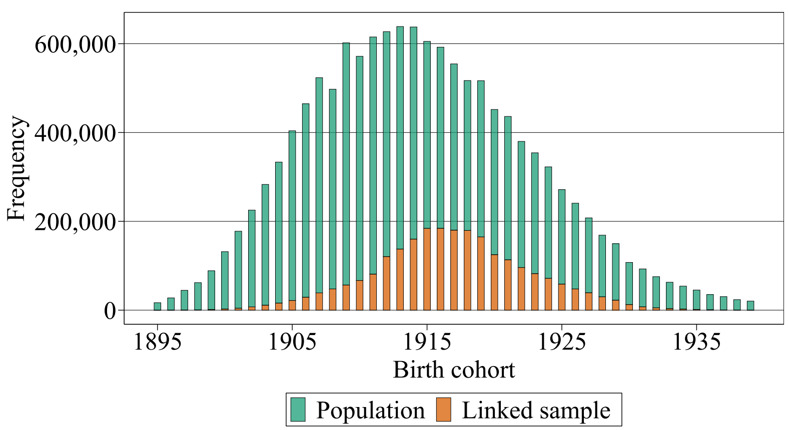
Notes: The key variables of interest from each dataset are indicated in the boxes. The variables used for each type of linkage are indicated in parentheses.

Figure A.3: Coverage of men in our population of interest and their children in Numident death-census linked sample

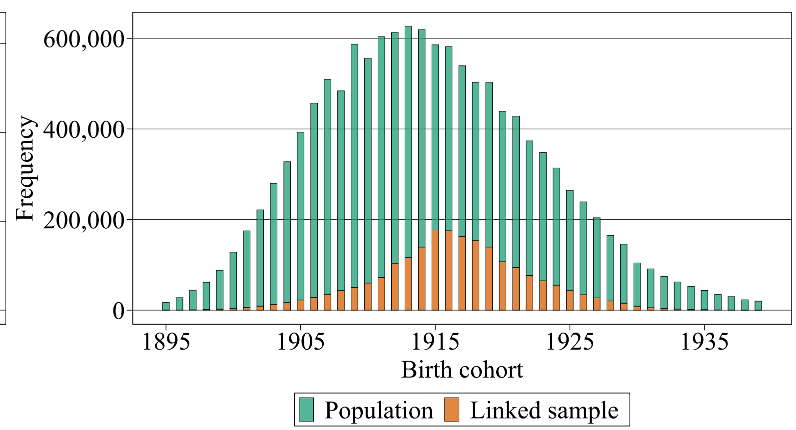
(a) Men born in 1875–1888 covered by Social Security in 1935 or 1950 based on 1930 employment



(b) Sons of men born in 1875–1888

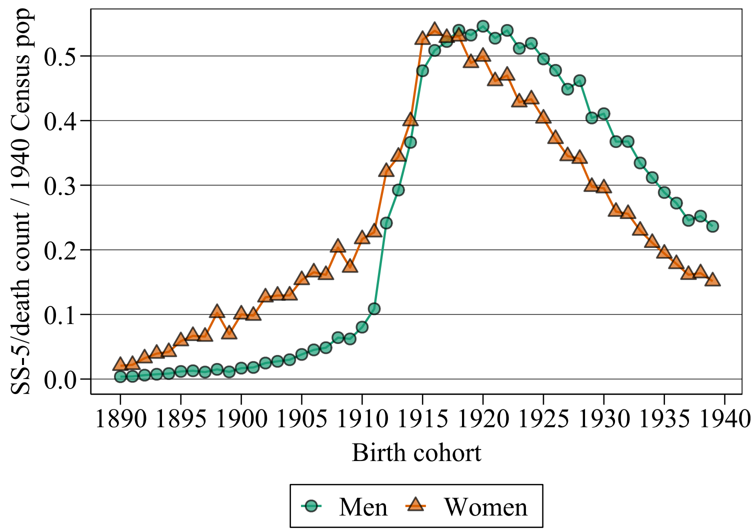


(c) Daughters of men born in 1875–1888



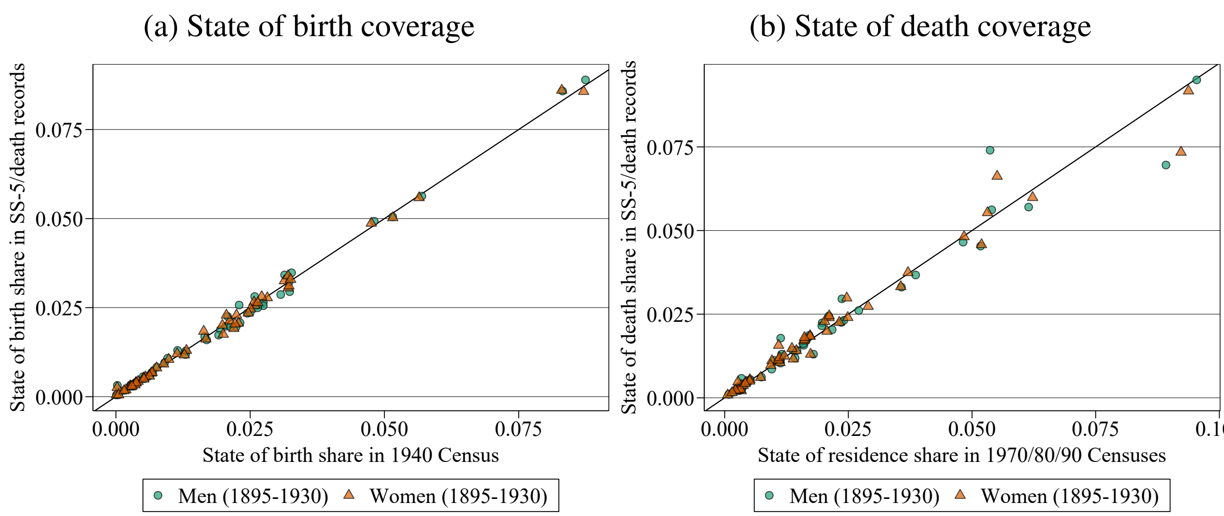
Notes: Population in panel (a): men born in 1875–1888 in the 1930 Census who were covered by Social Security in 1935 or 1950 based on their employment information (excluding the self-employed). Linked sample in panel (a): fathers in Numident death-census linked sample satisfying analogous restrictions. Population in panels (b)–(c): sons/daughters born in 1895–1909/1910–1919/1920–1929/1930–1939 in the 1910/1920/1930/1940 Censuses whose fathers were born in 1875–1888. Linked samples in panels (b)–(c): sons/daughters born in 1895–1939 in Numident death-census linked sample whose fathers were born in 1875–1888.

Figure A.4: Cohort coverage of Numident records by sex



Notes: This figure plots the implied cohort coverage rate of SSNs that can be found in Numident SS-5 and death records, defined as the number of such SSNs by cohort divided by the corresponding number of individuals in the 1940 Census, separately by gender. It therefore captures the share of individuals who survived until 1940 who appear in Numident records.

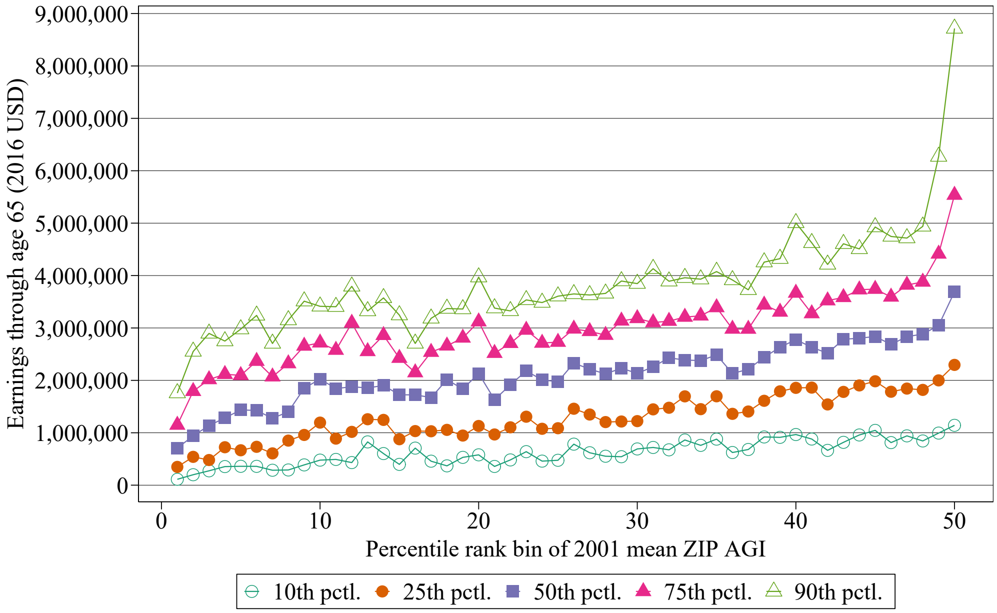
Figure A.5: State of birth and state of death coverage of Numident records by sex



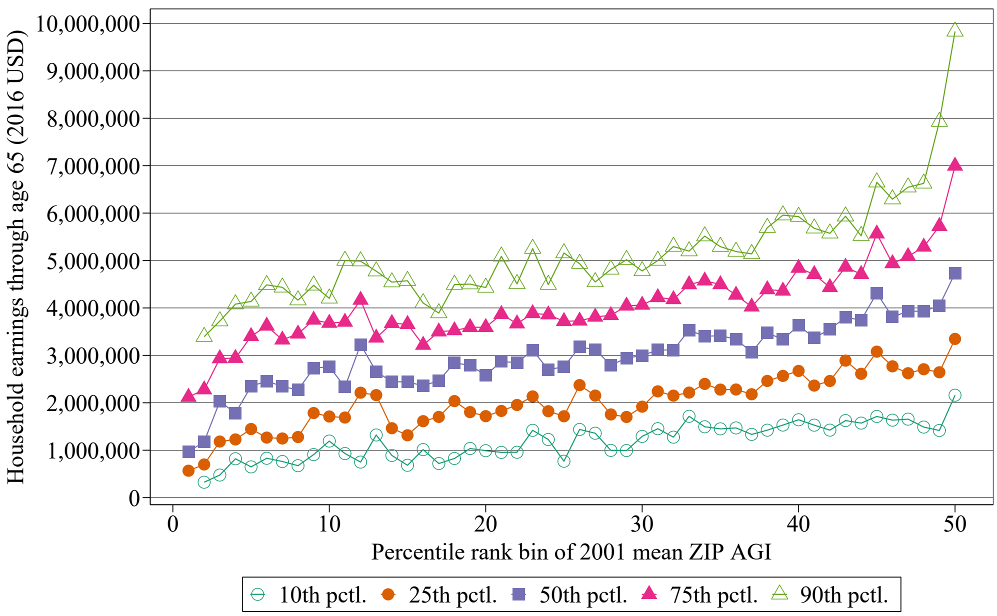
Notes: Panel (a) plots state of birth shares among SSNs that can be found in Numident SS-5 and death records against the corresponding shares in the 1940 Census for individuals born between 1895 and 1930, separately by gender (US-born individuals only). For the same set of cohorts and SSNs, panel (b) plots state of death shares in Numident death records against state of residence shares in the 1970 Census (for 1895–1910 cohorts), 1980 Census (for 1911–1920 cohorts), and 1990 Census (for 1921–1930 cohorts), separately by gender.

Figure A.6: Moments of lifetime earnings distribution by ZIP-level income percentile rank bin

(a) Men, own lifetime earnings



(b) Women, household lifetime earnings

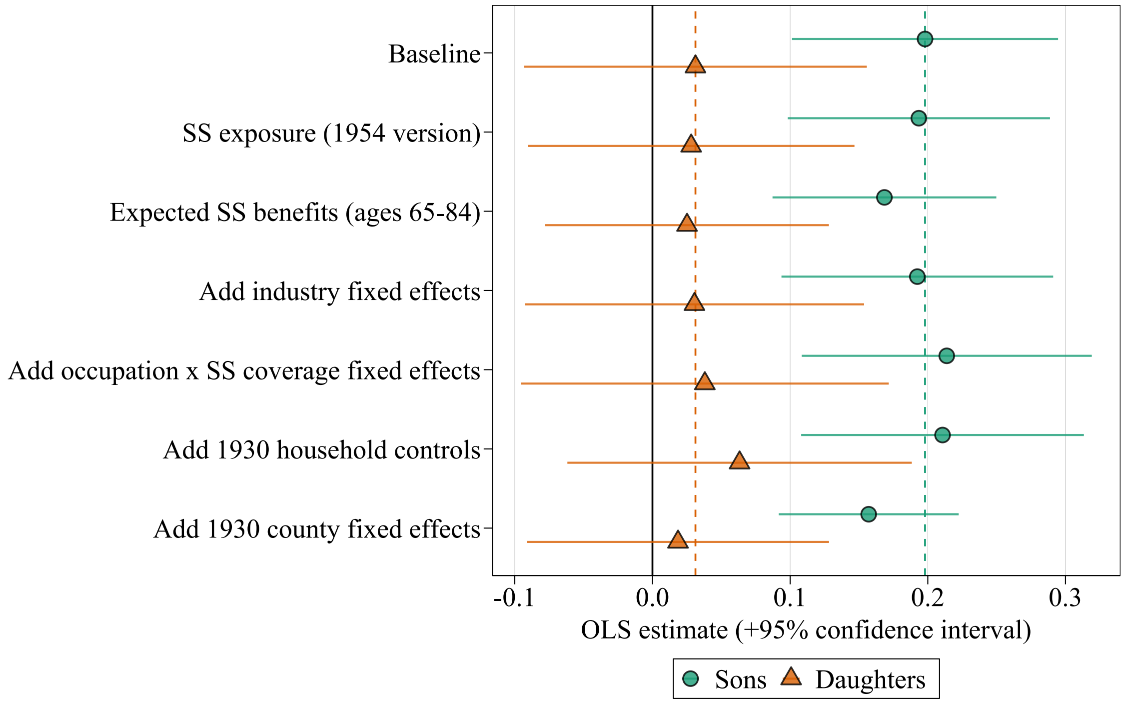


Notes: In this figure, the sample consists of HRS respondents aged 65 or older when they were last interviewed and for which we have ZIP of residence and SSA earnings information (see Appendix A.5 for details). Panel (a) plots moments of the distribution of lifetime earnings through age 65 for men by the rank of their last known ZIP code of residence, where ZIP codes are ranked according to mean AGI in 2001 and categorized into 50 bins each containing two adjacent percentile ranks (e.g., the 1st bin includes the 1st and 2nd percentiles, etc.). Panel (b) analogously plots moments of the household lifetime earnings distribution for women with spouses, where household earnings are defined as the sum of respondents' lifetime earnings through age 65 and that of their spouses.

Figure A.7: The impact of father's exposure to Social Security on children's late-life neighborhood quality in terms of various socioeconomic indicators

Notes: This figure shows the coefficients on father's Social Security exposure (i.e., negative of earliest age of Social Security eligibility) from equation (3), where the sample consists of children in our Numident death-census linked sample and the outcome is the percentile rank of children's late-life ZIP code in terms of various socioeconomic indicators. Signs of the coefficients for the ZIP-level poverty rate, ZIP-level share of households receiving public assistance, and ZIP-level share of single-parent families have been inverted so that a positive coefficient means father's Social Security exposure leads children to live in better neighborhoods late in life. Colors/shapes indicate separate estimates for all children, sons and daughters. All regressions include fixed effects for children's sex-by-birth cohort and fathers' 3-digit occupation-by-birth cohort, Social Security coverage category, race, and birthplace. Observations weighted using inverse propensity score weights that re-weight the sample to match the population of interest in 1930. 95% confidence intervals based on robust standard errors clustered by fathers' 3-digit industry.

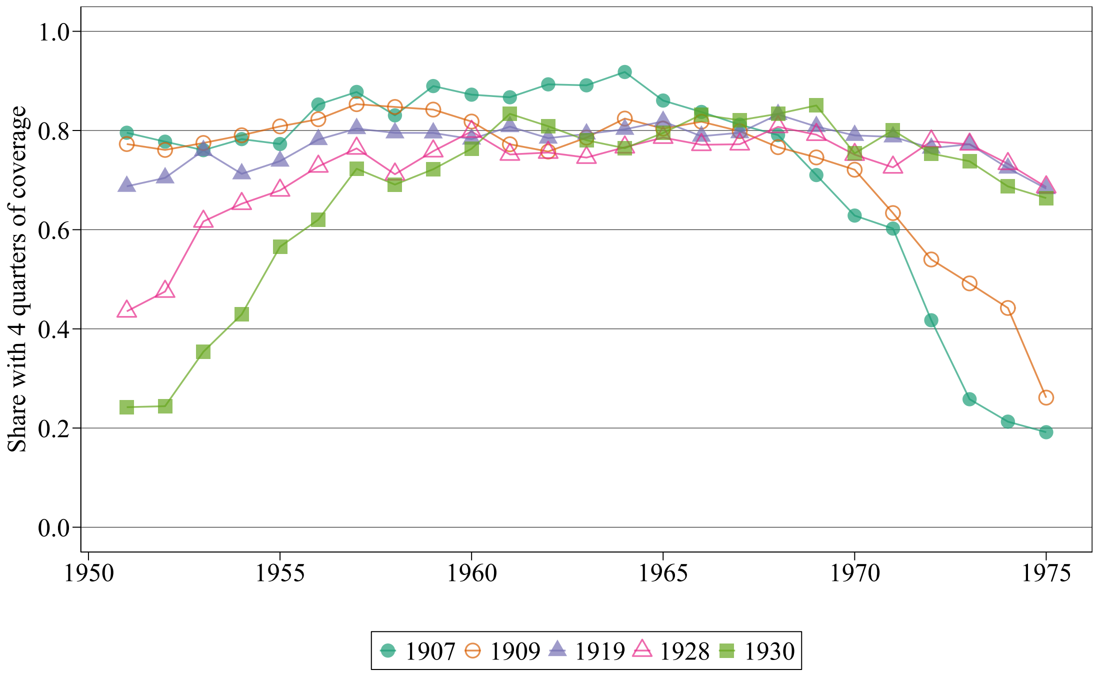
Figure A.8: The impact of father's exposure to Social Security on children's late-life ZIP rank in terms of 2001 AGI: Robustness checks



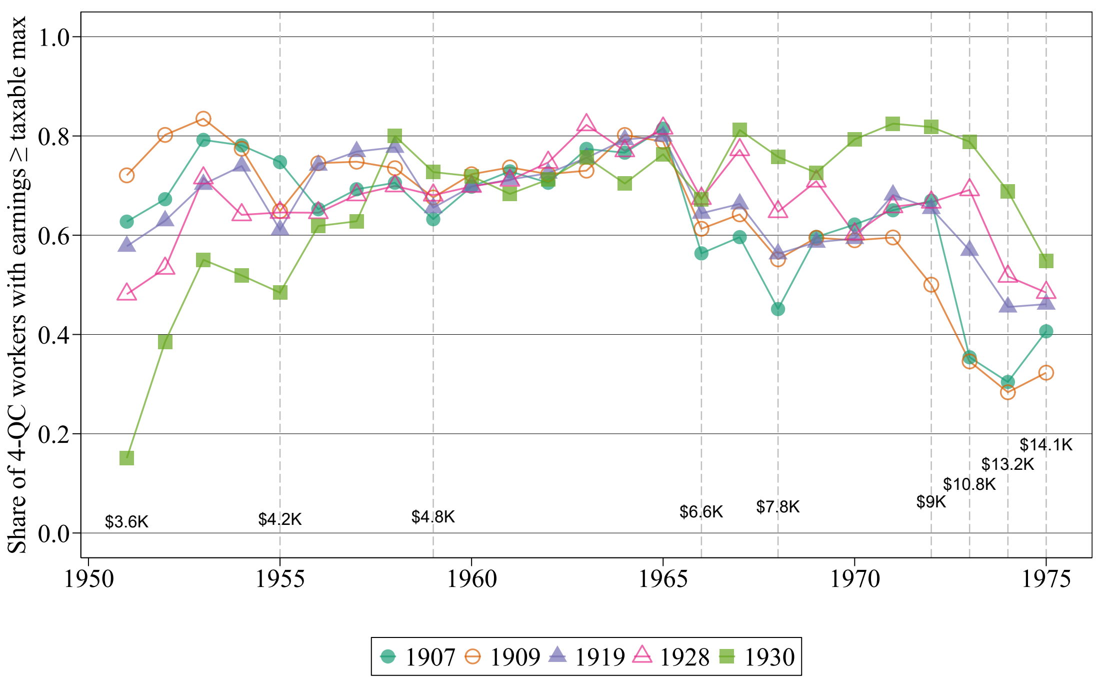
Notes: The top row in this figure shows the baseline coefficients from Table 4 (column 1). The next two rows show estimates using alternative measures of Social Security exposure (variant of the baseline measure that assumes eligibility starts in 1954 instead of 1950 for 1950-covered workers, or expected average monthly benefits between ages 65 and 84 based on the earnings history of a median earner, in tens of 2000 dollars). The bottom four rows show estimates under variants of the baseline specification that include various additional controls (3-digit industry fixed effects, 3-digit occupation-by-Social Security coverage category fixed effects, fixed effects for all outcomes in Table 3 except home values which is included as a linear control, or fixed effects for parents' 1930 county of residence). Colors/shapes indicate separate estimates for sons and daughters. All regressions include fixed effects for children's sex-by-birth cohort and fathers' 3-digit occupation-by-birth cohort, Social Security coverage category, race, and birthplace. Observations weighted using inverse propensity score weights that re-weight the sample to match the population of interest in 1930. 95% confidence intervals based on robust standard errors clustered by fathers' 3-digit industry.

Figure A.9: Full-time covered employment and top-coding of SSA earnings data in OCG sample

(a) Share with 4 quarters of coverage (QC) by year for select cohorts

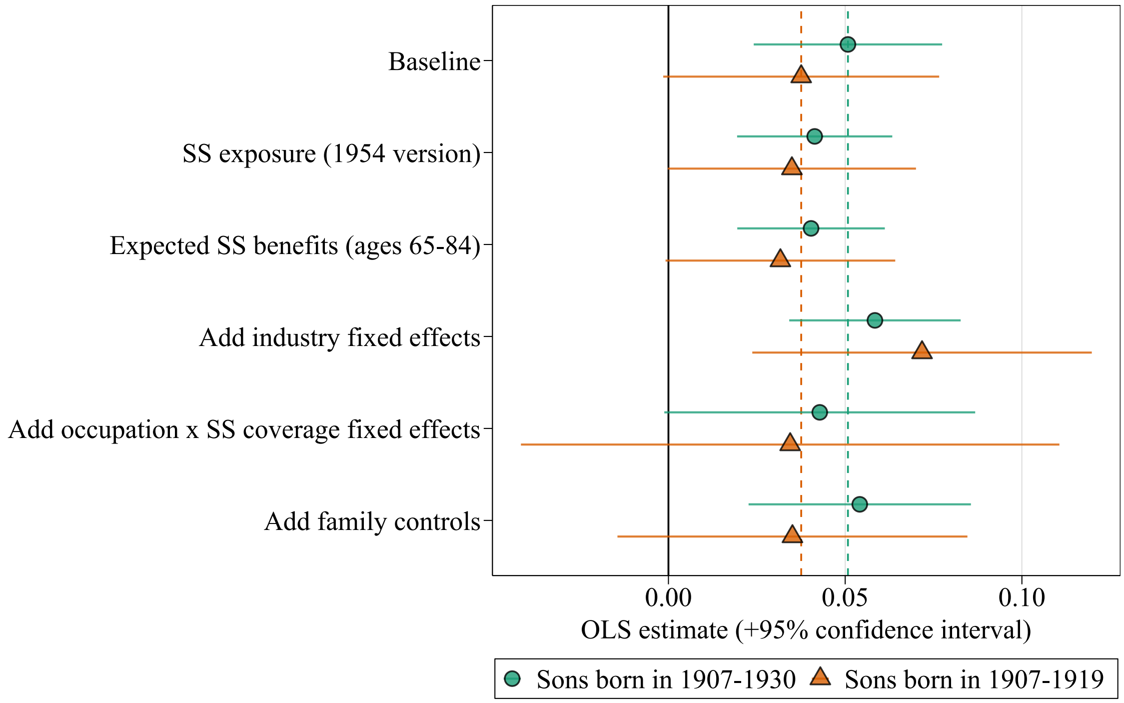


(b) Share of 4-QC workers with earnings exceeding the Social Security taxable maximum by year for select cohorts



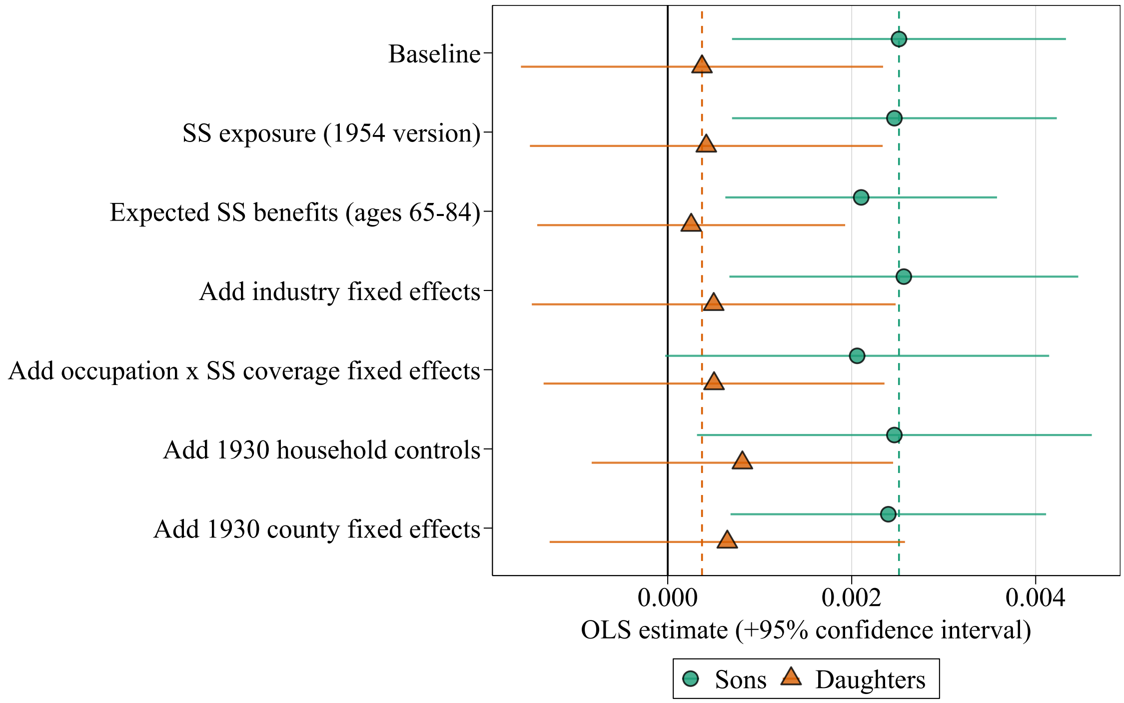
Notes: Panel (a) plots the share of respondents with 4 quarters of coverage by year for select birth cohorts. Panel (b) plots the share of workers with 4 quarters of coverage with earnings exceeding the Social Security taxable maximum by year for select birth cohorts.

Figure A.10: The impact of father's exposure to Social Security on sons' occupational income score around age 50 (log): Robustness checks



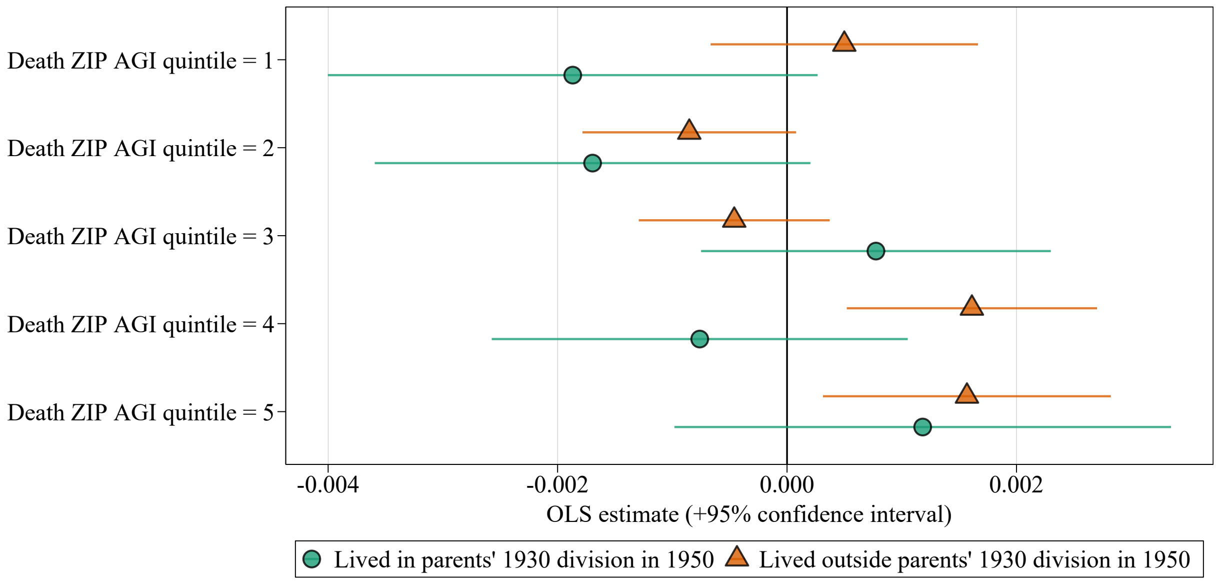
Notes: The top row in this figure shows the baseline coefficients from Table 5 (column 10). The next two rows show estimates using alternative measures of Social Security exposure (variant of the baseline measure that assumes eligibility starts in 1954 instead of 1950 for 1950-covered workers or the expected average monthly benefits between ages 65 and 84, in tens of 2000 dollars). The bottom three rows show estimates under variants of the baseline specification that include various additional controls (3-digit industry fixed effects, 3-digit occupation-by-Social Security coverage category fixed effects, or fixed effects for the number of children, both parents' years of education, and bins for family income at age 16). Colors/shapes indicate separate estimates for sons born in 1907–1930 and sons born in 1907–1919. All regressions include fixed effects for sons' birth cohort and race, and fathers' 3-digit occupation-by-birth cohort, Social Security coverage category, and birthplace. Observations weighted using OCG sampling weights. 95% confidence intervals based on robust standard errors clustered by fathers' 3-digit industry.

Figure A.11: The impact of father's exposure to Social Security on children's propensity to live outside their parents' 1930 census division of residence in 1950: Robustness checks



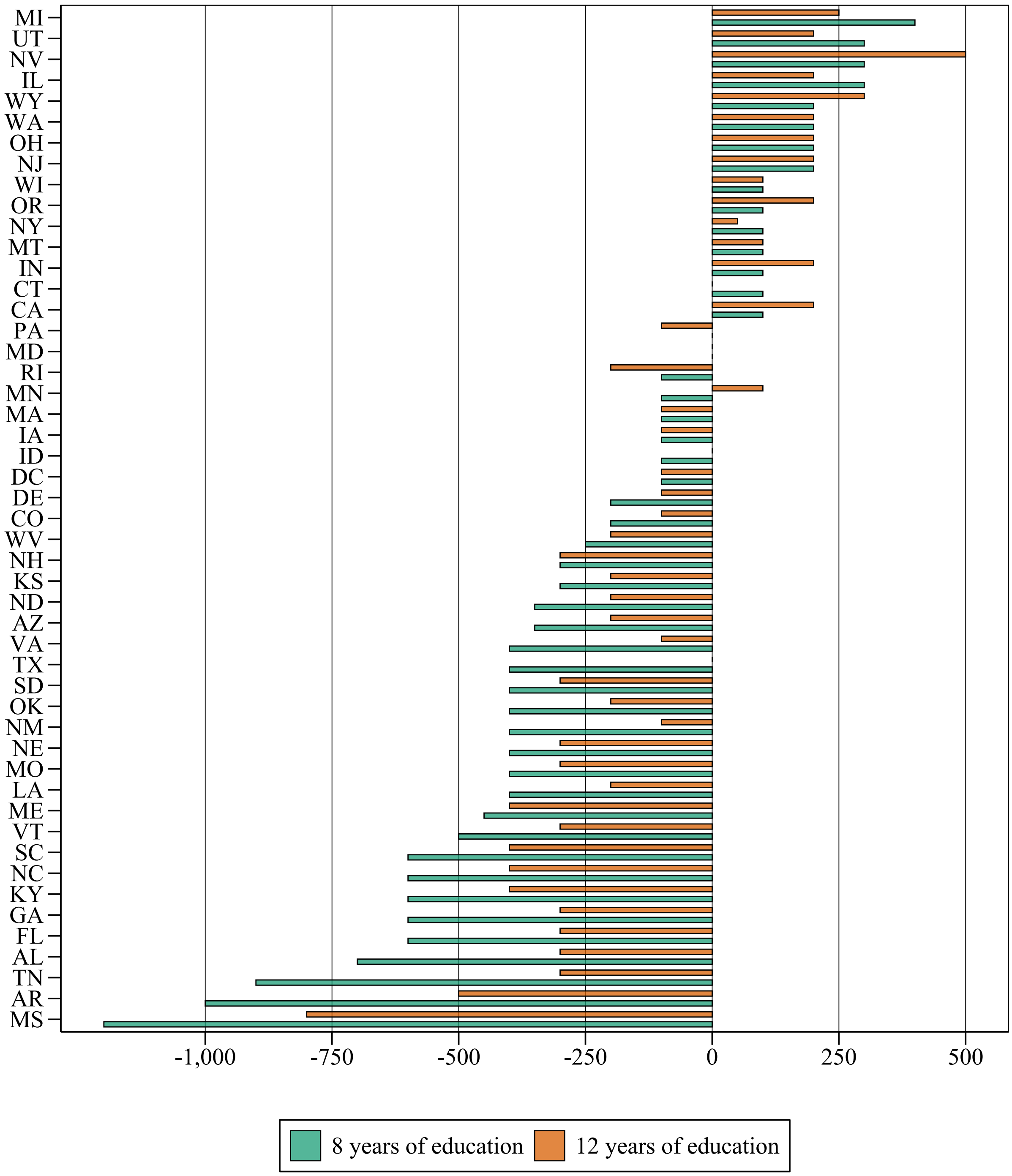
Notes: The top row in this figure shows the baseline coefficients from Table 7 (column 4). The next two rows show estimates using alternative measures of Social Security exposure (variant of the baseline measure that assumes eligibility starts in 1954 instead of 1950 for 1950-covered workers, or expected average monthly benefits between ages 65 and 84 based on the earnings history of a median earner, in tens of 2000 dollars). The bottom four rows show estimates under variants of the baseline specification that include various additional controls (3-digit industry fixed effects, 3-digit occupation-by-Social Security coverage category fixed effects, fixed effects for all outcomes in Table 3 except home values which is included as a linear control, or fixed effects for parents' 1930 county of residence). Colors/shapes indicate separate estimates for sons and daughters. All regressions include fixed effects for children's sex-by-birth cohort and fathers' 3-digit occupation-by-birth cohort, Social Security coverage category, race, and birthplace. Observations weighted using inverse propensity score weights that re-weight the sample to match the population of interest in 1930. 95% confidence intervals based on robust standard errors clustered by fathers' 3-digit industry.

Figure A.12: The impact of father's exposure to Social Security on sons' joint migration status in 1950 and late-life ZIP quintile in terms of 2001 mean AGI



Notes: This figure shows the coefficients on father's Social Security exposure (i.e., negative of earliest age of Social Security eligibility) from equation (3), where the sample consists of sons in our Numident death-census linked sample who are additionally linked to the 1950 Census and the dependent variable is an indicator for the joint outcome of living in or outside parents' 1930 census division of residence by 1950 and living in a specified ZIP quintile at the time of death (where ZIP quintiles are based on mean AGI in 2001). All regressions include fixed effects for sons' birth cohort and fathers' 3-digit occupation-by-birth cohort, Social Security coverage category, race, and birthplace. Observations weighted using inverse propensity score weights that re-weight the sample to match the population of interest in 1930. 95% confidence intervals based on robust standard errors clustered by fathers' 3-digit industry.

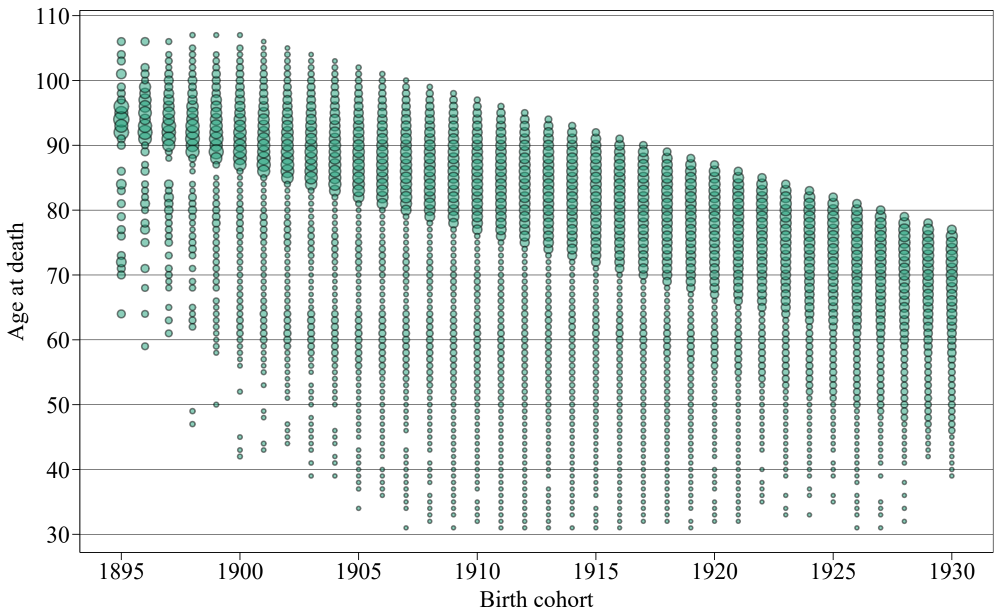
Figure A.13: Median earnings by state in 1950 relative to national median, men aged 25–54 with 8 years of education vs. 12 years of education



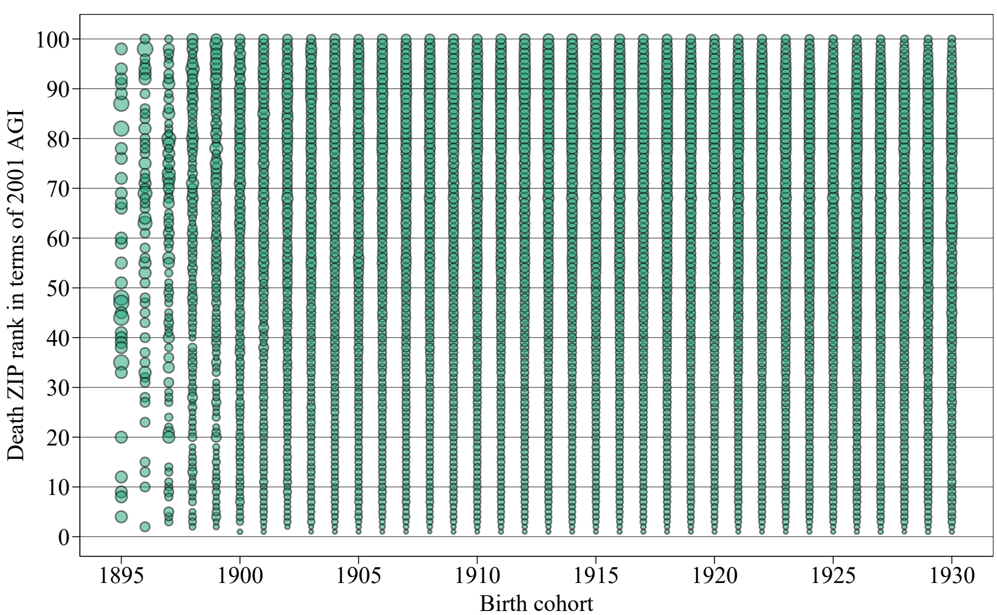
Notes: This figure plots deviations in median earned income by state in 1950 relative to the national median, separately for men aged 25–54 with 8 years of education vs. 12 years of education. Earned income is defined as the sum of wage income and business or farm income. All amounts expressed in 1949 dollars.

Figure A.14: Age at death and late-life ZIP AGI rank distribution by birth cohort in Numident death-census linked sample

(a) Age at death distribution



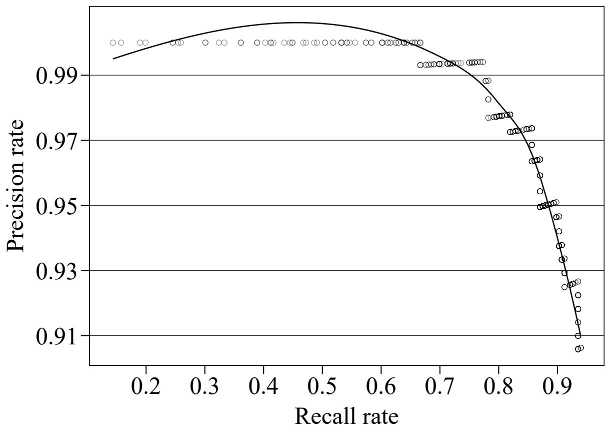
(b) Late-life ZIP AGI rank distribution



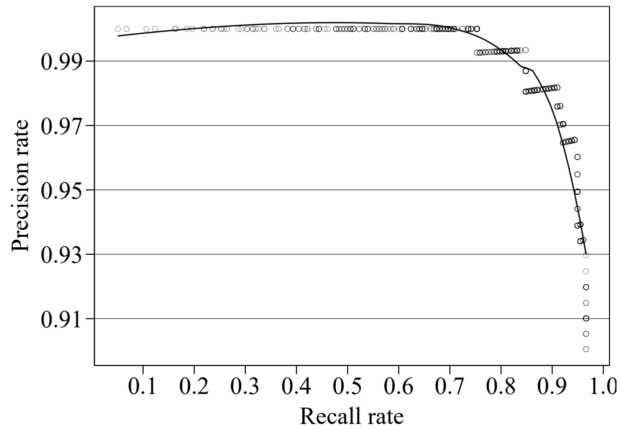
Notes: Panel (a) plots the age at death distribution by birth cohort among children in our Numident death-census linked sample (where the size of each dot is proportional to the cohort-specific share). Panel (b) similarly plots the distribution of late-life ZIP AGI percentile ranks by birth cohort among children our Numident death-census linked sample.

Figure A.15: Model performance: Precision-recall curves

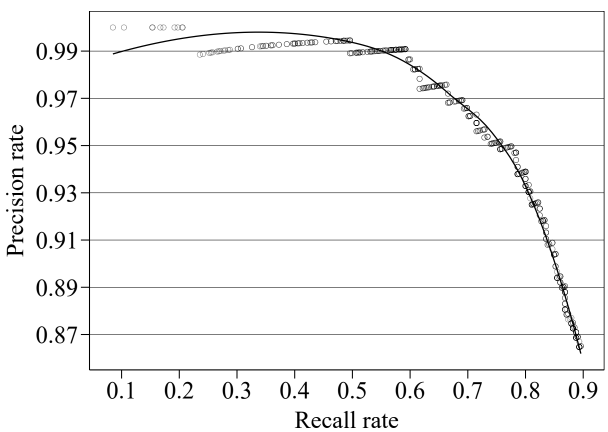
(a) Numident SS-5-1930 Census linking



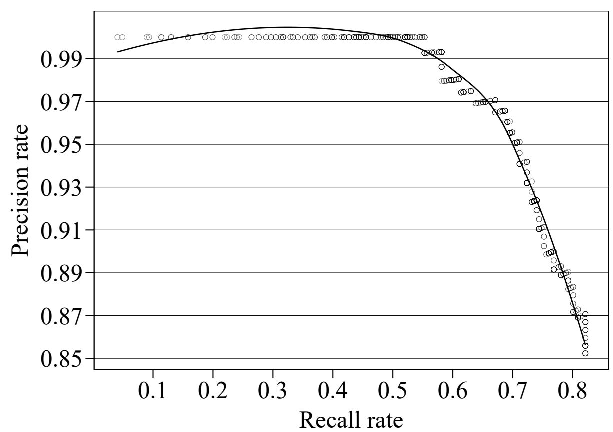
(b) 1930-1920 Census linking



(c) Numident claims-1930 Census linking (records with non-missing birthplace)



(d) Numident claims-1930 Census linking (records with missing birthplace)



Notes: These figures plot the performance of each model as measured by the cross-validated precision-recall curve in the training set. Each dot corresponds to the precision and recall rate associated with a specific probability threshold.

Table A.1: Representativeness of Numident death-census linked sample

	Mean			Difference (two-sided t-test)	
	Population	Linked sample (unweighted)	Linked sample (weighted)	Linked sample (unweighted)	Linked sample (weighted)
	(1)	(2)	(3)	(4)	(5)
Panel (a): Father characteristics in 1930					
Black	0.074	0.028	0.072	-0.046	-0.002
				[0]	[0.001]
Born in 1875–1879	0.296	0.264	0.297	-0.032	0.001
				[0]	[0.161]
Born in 1880–1884	0.375	0.383	0.375	0.007	-0.001
				[0]	[0.461]
Born in 1885–1888	0.329	0.354	0.329	0.025	0
				[0]	[0.499]
Born abroad	0.309	0.302	0.309	-0.007	0
				[0]	[0.59]
Born in Northeast	0.203	0.221	0.203	0.017	0
				[0]	[0.564]
Born in Midwest	0.25	0.274	0.25	0.024	0
				[0]	[0.81]
Born in South	0.215	0.182	0.215	-0.034	-0.001
				[0]	[0.198]
Born in West	0.023	0.022	0.023	-0.001	0
				[0.002]	[0.623]
Literate	0.941	0.955	0.942	0.015	0.001
				[0]	[0.001]
Covered by Social Security in 1935	0.893	0.899	0.891	0.005	-0.003
				[0]	[0]
Covered by Social Security in 1950	0.107	0.101	0.109	-0.005	0.003
				[0]	[0]
Farmer	0.002	0.003	0.003	0	0
				[0]	[0.041]
Unskilled worker	0.265	0.234	0.26	-0.031	-0.005
				[0]	[0]
Skilled/semi-skilled worker	0.474	0.506	0.492	0.032	0.018
				[0]	[0]
White-collar worker	0.259	0.257	0.245	-0.001	-0.013
				[0.014]	[0]
(table continues on next page)					
N	3,920,008	858,612			

Notes: Columns (3) and (5) weighted using inverse propensity score weights. p-values for two-sided t-test of equality of means in brackets. Sample restricted to ever-married men who are household heads (or spouses of heads) in 1930.

Table A.1 (cont.): Representativeness of Numident death-census linked sample

	Mean			Difference (two-sided *t*-test)	
	Population	Linked sample (unweighted)	Linked sample (weighted)	Linked sample (unweighted)	Linked sample (weighted)
	(1)	(2)	(3)	(4)	(5)
Panel (b): Household characteristics in 1930					
Live in Northeast	0.369	0.387	0.378	0.018	0.009
				[0]	[0]
Live in Midwest	0.321	0.346	0.33	0.025	0.009
				[0]	[0]
Live in South	0.202	0.176	0.2	-0.027	-0.002
				[0]	[0]
Live in West	0.108	0.092	0.092	-0.016	-0.016
				[0]	[0]
Live on farm	0.051	0.06	0.051	0.009	0
				[0]	[0.535]
Live in urban area	0.739	0.712	0.737	-0.027	-0.002
				[0]	[0.018]
Own home	0.488	0.552	0.492	0.064	0.004
				[0]	[0]
Own a radio	0.475	0.509	0.477	0.034	0.002
				[0]	[0.005]
N	3,920,008	858,612			

Notes: Columns (3) and (5) weighted using inverse propensity score weights. *p*-values for two-sided *t*-test of equality of means in brackets. Sample restricted to ever-married men who are household heads (or spouses of heads) in 1930.

Table A.2: Representativeness of Numident claims-census linked sample

	Mean			Difference (two-sided *t*-test)	
	Population	Linked sample (unweighted)	Linked sample (weighted)	Linked sample (unweighted)	Linked sample (weighted)
	(1)	(2)	(3)	(4)	(5)
Panel (a): Own characteristics in 1930					
Black	0.074	0.026	0.076	-0.048	0.002
				[0]	[0.01]
Born in 1875–1879	0.296	0.357	0.287	0.062	-0.009
				[0]	[0]
Born in 1880–1884	0.375	0.392	0.374	0.017	-0.001
				[0]	[0.148]
Born in 1885–1888	0.329	0.213	0.339	-0.116	0.01
				[0]	[0]
Born abroad	0.309	0.268	0.311	-0.041	0.002
				[0]	[0.047]
Born in Northeast	0.203	0.24	0.202	0.037	-0.001
				[0]	[0.187]
Born in Midwest	0.25	0.311	0.248	0.062	-0.002
				[0]	[0.02]
Born in South	0.215	0.155	0.216	-0.06	0.001
				[0]	[0.35]
Born in West	0.023	0.025	0.023	0.002	0
				[0]	[0.731]
Literate	0.941	0.969	0.941	0.028	0
				[0]	[0.853]
Covered by Social Security in 1935	0.893	0.938	0.924	0.044	0.03
				[0]	[0]
Covered by Social Security in 1950	0.107	0.062	0.076	-0.044	-0.03
				[0]	[0]
Farmer	0.002	0.002	0.002	-0.001	0
				[0]	[0.003]
Unskilled worker	0.265	0.195	0.241	-0.07	-0.024
				[0]	[0]
Skilled/semi-skilled worker	0.474	0.527	0.513	0.053	0.039
				[0]	[0]
White-collar worker	0.259	0.277	0.244	0.018	-0.014
				[0]	[0]
(table continues on next page)					
N	3,920,008	367,732			

Notes: Columns (3) and (5) weighted using inverse propensity score weights. *p*-values for two-sided *t*-test of equality of means in brackets. Sample restricted to ever-married men who are household heads (or spouses of heads) in 1930.

Table A.2 (cont.): Representativeness of Numident claims-census linked sample

	Mean			Difference (two-sided *t*-test)	
	Population	Linked sample (unweighted)	Linked sample (weighted)	Linked sample (unweighted)	Linked sample (weighted)
	(1)	(2)	(3)	(4)	(5)
Panel (b): Household characteristics in 1930					
Live in Northeast	0.369	0.374	0.364	0.005	-0.005
				[0]	[0]
Live in Midwest	0.321	0.37	0.338	0.049	0.017
				[0]	[0]
Live in South	0.202	0.142	0.189	-0.06	-0.013
				[0]	[0]
Live in West	0.108	0.114	0.109	0.006	0.001
				[0]	[0.065]
Live on farm	0.051	0.038	0.051	-0.013	0
				[0]	[0.647]
Live in urban area	0.739	0.767	0.74	0.028	0
				[0]	[0.636]
Own home	0.488	0.556	0.488	0.067	0
				[0]	[0.763]
Own a radio	0.475	0.562	0.473	0.087	-0.001
				[0]	[0.184]
N	3,920,008	367,732			

Notes: Columns (3) and (5) weighted using inverse propensity score weights. *p*-values for two-sided *t*-test of equality of means in brackets. Sample restricted to ever-married men who are household heads (or spouses of heads) in 1930.

Table A.3: Responses to 1940 Social Security Number question by Social Security coverage year

	Share of male workers in 1940 (%)	
	Without an SSN	With an SSN
	(1)	(2)
Panel (a): All emp. categories (excl. public emergency), male workers aged 25–64		
Covered in 1935 based on 1940 employment	9.3	90.7
Covered in 1950 or later based on 1940 employment	74.3	25.7
Panel (b): Emp. categories in analysis (excl. public emergency), male workers aged 51–64		
Covered in 1935 based on 1940 employment	12.8	87.2
Covered in 1950 based on 1940 employment	76.5	23.5

Notes: Panel (a) shows the share of male workers aged 25–64 reporting having an SSN vs. not in the 1940 Census (sample-line respondents only), separately for workers whose employment was covered by Social Security in 1935 vs. 1950 or later. Panel (b) further restricts the sample to the population of interest in the analysis: male workers aged 51–64 whose employment was covered in 1935 or 1950, excluding the self-employed. We exclude public emergency workers from this analysis as their coverage status is hard to determine.

Table A.4: Top 20 occupations with some Social Security coverage year variation in 1930

	Men born in 1875–1888 in 1930 Census						
	Population			Death-census linked sample		Claims-census linked sample	
	Number	Share covered in 1935 (%)	Share (%)	Share covered in 1935 (%)	Share (%)	Share covered in 1935 (%)	Share (%)
Occupation	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Laborers (n.e.c.)	586,498	90.6	14.96	90.8	14.99	93.4	14.28
Managers, officials, and proprietors (n.e.c.)	289,471	98.3	7.38	98	7.18	98.7	6.97
Carpenters	204,143	98.5	5.21	98.5	5.52	98.7	5.58
Foremen (n.e.c.)	119,435	98.4	3.05	98.2	3.28	99.2	3.38
Clerical and kindred workers (n.e.c.)	73,422	92.3	1.87	92	1.76	95.2	1.87
Painters, construction and maintenance	69,527	97.2	1.77	97.4	1.65	97.4	1.69
Truck and tractor drivers	68,466	95.4	1.75	95.4	1.86	96.5	1.82
Stationary engineers	63,644	92.7	1.62	92.6	1.64	95.4	1.65
Janitors and sextons	57,344	56.9	1.46	51.9	1.5	67.3	1.43
Plumbers and pipe fitters	38,794	96.8	0.99	96.6	1	97.5	1
Electricians	34,909	95.3	0.89	94.8	0.88	97.2	0.93
Mechanics and repairmen (n.e.c.)	30,130	97.1	0.77	97.1	0.77	98.1	0.84
Teachers (n.e.c.)	28,506	4.3	0.73	3.4	0.83	5.8	0.44
Stationary firemen	28,068	93.6	0.72	93.6	0.8	95.6	0.77
Guards, watchmen, and doorkeepers	27,065	88.1	0.69	88.2	0.73	91.8	0.63
Accountants and auditors	23,524	94.8	0.6	94.8	0.55	95.6	0.61
Bookkeepers	23,382	96.2	0.6	96.1	0.58	97.1	0.65
Taxicab drivers and chauffers	23,062	43.7	0.59	43.8	0.55	41.3	0.56
Cooks, except private household	22,531	95.9	0.57	94.7	0.36	95.8	0.52
Gardeners, except farm and groundskeepers	22,236	58	0.57	58	0.55	60.1	0.46

Notes: This table shows summary statistics for the top 20 occupations with some variation in Social Security coverage year in the 1930 Census (defined as occupations with a share of workers covered in 1935 between 1 percent and 99 percent). Population consists of men born in 1875–1888 whose current employment was covered in 1935 or 1950 (excluding the self-employed). Columns 4 and 5 restrict the sample to men who appear in our Numident death-census linked sample. Columns 6 and 7 restrict the sample to men who appear in our Numident claims-census linked sample. Observations in columns 4–7 weighted using inverse propensity score weights that re-weight the sample to match the population of interest in 1930.

Table A.5: Selection of 1930 households and children into our Numident death-census linked sample, by child cohorts

	Dependent variable: Any/number of children linked					
	Children born in 1895–1910		Children born in 1911–1920		Children born in 1921–1930	
	Any	Number	Any	Number	Any	Number
	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): All children						
Social Security exposure	-0.0003	-0.0004	-0.0001	0.0001	0.0004	0.0006
	(0.0003)	(0.0003)	(0.0004)	(0.0005)	(0.0003)	(0.0005)
Mean of dep. variable	0.043	0.049	0.178	0.262	0.071	0.093
Panel (b): Sons						
Social Security exposure	-0.0000	-0.0000	0.0000	0.0002	0.0002	0.0002
	(0.0002)	(0.0002)	(0.0003)	(0.0004)	(0.0002)	(0.0003)
Mean of dep. variable	0.023	0.025	0.112	0.138	0.045	0.052
Panel (c): Daughters						
Social Security exposure	-0.0003**	-0.0004***	-0.0001	-0.0001	0.0003	0.0004**
	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Mean of dep. variable	0.022	0.023	0.102	0.124	0.036	0.041
N	3,920,008	3,920,008	3,920,008	3,920,008	3,920,008	3,920,008

Notes: This table shows the coefficients on Social Security exposure (i.e., negative of earliest age of Social Security eligibility) from equation (2), where the sample consists of all men in our population of interest in the 1930 Census and the outcome is either an indicator for any children/sons/daughters born in a specified cohort range appearing in our Numident death-census linked sample or a discrete variable for the number of such children/sons/daughters appearing in our linked sample. All regressions include fixed effects for men's 3-digit occupation-by-birth cohort, Social Security coverage category, race, and birthplace. Robust standard errors in parentheses, clustered by men's 3-digit industry. *** 1%, ** 5%, * 10% significance.

Table A.6: The impact of father's exposure to Social Security on parents' educational attainment and fathers' wage income and employment status in 1940

	Dependent variable: Parents' outcomes in 1940			
	Years of education	Wage income (log)	In labor force	Employed
	(1)	(2)	(3)	(4)
Panel (a): Men in 1930–1940 CLP linked sample (ABE conservative algorithm)				
Social Security exposure	0.0052	-0.0001	-0.0012	-0.0015
	(0.0085)	(0.0035)	(0.0009)	(0.0010)
Mean of dep. variable	7.6	1,037	0.89	0.81
N	725,039	547,692	740,257	740,257
Panel (b): Fathers in Numident death-census sample who are additionally linked to 1940				
Social Security exposure	0.0117***	-0.0020	0.0000	-0.0014
	(0.0045)	(0.0039)	(0.0008)	(0.0010)
Mean of dep. variable	7.4	1,040	0.90	0.81
N	509,596	395,646	521,023	521,023
Panel (c): Mothers in Numident death-census linked sample who are additionally linked to 1940				
Social Security exposure	0.0000	—	—	—
	(0.0073)			
Mean of dep. variable	7.5			
N	516,147			

Notes: This table shows the coefficients on father's Social Security exposure (i.e., negative of earliest age of Social Security eligibility) from equation (2), where the outcome is indicated in the column title. In panel (a), the sample consists of men in our population of interest in the 1930 Census who are linked to the 1940 Census via linkages from the Census Linking Project using the ABE conservative algorithm (Abramitzky et al., 2025). In panels (b) and (c), the sample consists of fathers and mothers in our Numident death-census linked sample who are additionally linked to the 1940 Census at 97 percent precision via linkages from Bailey et al. (2026). All regressions include fixed effects for fathers' 3-digit occupation-by-birth cohort, Social Security coverage category, race, and birthplace. In column 2, means of wage income in 1939 dollars are shown. Observations weighted using inverse propensity score weights that re-weight the sample to match the population of interest in 1930. Robust standard errors in parentheses, clustered by fathers' 3-digit industry. *** 1%, ** 5%, * 10% significance.

Table A.7: The relationship between late-life ZIP quality and late-career earnings

	Dependent variable: Late-life ZIP quality	
	Mean AGI in 2001 (log)	Median home values in 2000 (log)
	(1)	(2)
Late-career earnings (log)	0.085***	0.092***
	(0.003)	(0.005)
R^2	0.071	0.045
N	8,100	7,900

Notes: Each entry in this table corresponds to the estimated slope coefficient from a linear regression of a measure of late-life ZIP quality (log of ZIP-level mean AGI in 2001 or ZIP-level median home values in 2000) on a constant and the log of late-career earnings between ages 50 and 65, where the sample consists of male HRS respondents aged 65 or older when they were last interviewed and for which we have ZIP of residence and SSA earnings information (see Appendix A.5 for details). Observations weighted using HRS respondent weights. Robust standard errors in parentheses. Sample sizes have been rounded to the nearest hundredth following HRS disclosure guidelines. *** 1%, ** 5%, * 10% significance.

Table A.8: The impact of father's exposure to Social Security on children's post-1950 migration patterns

	Dependent variable: Children's location at death vs. location in 1950							
	Different location					Distance between counties		
	County	County (same state)	State	Census division	Census region	> 0 and $\leqslant 100$ miles	> 100 and $\leqslant 500$ miles	> 500 miles
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel (a): All children								
Social Security exposure	-0.0011	0.0005	-0.0016	-0.0018	-0.0020*	0.0006	0.0000	-0.0016*
	(0.0013)	(0.0007)	(0.0012)	(0.0013)	(0.0011)	(0.0006)	(0.0010)	(0.0009)
Mean of dep. variable	0.51	0.23	0.28	0.21	0.18	0.21	0.13	0.17
N	507,232	507,232	507,936	507,936	507,936	507,232	507,232	507,232
Panel (b): Sons								
Social Security exposure	-0.0022	0.0000	-0.0022*	-0.0026*	-0.0028***	-0.0006	-0.0007	-0.0010
	(0.0016)	(0.0014)	(0.0013)	(0.0014)	(0.0009)	(0.0013)	(0.0009)	(0.0008)
Mean of dep. variable	0.51	0.22	0.29	0.22	0.18	0.20	0.13	0.18
N	253,174	253,174	253,564	253,564	253,564	253,174	253,174	253,174
Panel (c): Daughters								
Social Security exposure	-0.0002	0.0013	-0.0015	-0.0012	-0.0017	0.0017	0.0006	-0.0026**
	(0.0015)	(0.0011)	(0.0013)	(0.0013)	(0.0014)	(0.0012)	(0.0013)	(0.0011)
Mean of dep. variable	0.50	0.23	0.27	0.20	0.17	0.21	0.13	0.16
N	254,058	254,058	254,372	254,372	254,372	254,058	254,058	254,058

Notes: This table shows the coefficients on father's Social Security exposure (i.e., negative of earliest age of Social Security eligibility) from equation (3), where the sample consists of children in our Numident death-census linked sample who are additionally linked to the 1950 Census and the outcome is an indicator for migrating to a different county/state/division/region or migrating a specified number of miles away between 1950 and the time of death. Panels show separate estimates for all children, sons and daughters. All regressions include fixed effects for children's sex-by-birth cohort and fathers' 3-digit occupation-by-birth cohort, Social Security coverage category, race, and birthplace. Observations weighted using inverse propensity score weights that re-weight the sample to match the population of interest in 1930. Robust standard errors in parentheses, clustered by fathers' 3-digit industry. *** 1%, ** 5%, * 10% significance.

Table A.9: The impact of father's exposure to Social Security on children's long-run mortality

	Dependent variable:	
	Died between ages 65 and 77	Age at death (conditional on dying between ages 65 and 77)
	(1)	(2)
Panel (a): All children (born in 1923–1930)		
Social Security exposure	-0.001	0.035***
	(0.001)	(0.010)
Mean of dep. variable	0.58	71.6
N	225,731	131,277
Panel (b): Sons (born in 1923–1930)		
Social Security exposure	-0.001	0.047**
	(0.002)	(0.018)
Mean of dep. variable	0.58	71.5
N	128,313	74,225
Panel (c): Daughters (born in 1923–1930)		
Social Security exposure	0.001	0.020
	(0.002)	(0.016)
Mean of dep. variable	0.58	71.8
N	97,418	57,052

Notes: This table shows the coefficients on father's Social Security exposure (i.e., negative of earliest age of Social Security eligibility) from equation (3), where the sample consists of children born in 1923–1930 in our Numident death-census linked sample. In column 1, the outcome is an indicator for dying between ages 65 and 77. In column 2, the outcome is age at death, for the subset of children who died between ages 65 and 77. Panels show separate estimates for all children, sons and daughters. All regressions include fixed effects for children's sex-by-birth cohort and fathers' 3-digit occupation-by-birth cohort, Social Security coverage category, race, and birthplace. Observations weighted using inverse propensity score weights that re-weight the sample to match the population of interest in 1930. Robust standard errors in parentheses, clustered by fathers' 3-digit industry. *** 1%, ** 5%, * 10% significance.

Table A.10: The impact of father's exposure to Social Security on children's living arrangements, labor market outcomes, marital status, and fertility in 1950

	Dependent variable: Outcome in 1950						
	Living with parent	Living with father	Living with mother	In labor force	Occscore (log)	Married	Any children
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel (a): All children							
Social Security exposure	-0.0012**	-0.0010*	-0.0010*	0.0010	-0.0005	-0.0001	0.0010
	(0.0005)	(0.0005)	(0.0006)	(0.0008)	(0.0010)	(0.0007)	(0.0009)
Mean of dep. variable	0.19	0.12	0.16	0.63	2,706	0.80	0.65
N	601,321	601,321	601,321	601,321	357,190	601,321	601,321
Panel (b): Sons							
Social Security exposure	-0.0014	-0.0012	-0.0010	0.0006	0.0001	0.0009	0.0011
	(0.0009)	(0.0014)	(0.0009)	(0.0008)	(0.0010)	(0.0014)	(0.0017)
Mean of dep. variable	0.18	0.12	0.16	0.94	2,855	0.80	0.63
N	309,682	309,682	309,682	309,682	270,215	309,682	309,682
Panel (c): Daughters							
Social Security exposure	-0.0009	-0.0009	-0.0008	0.0008	0.0001	-0.0009	0.0003
	(0.0013)	(0.0010)	(0.0012)	(0.0015)	(0.0020)	(0.0013)	(0.0012)
Mean of dep. variable	0.20	0.12	0.17	0.32	2,254	0.80	0.67
N	291,639	291,639	291,639	291,639	86,975	291,639	291,639

Notes: This table shows the coefficients on father's Social Security exposure (i.e., negative of earliest age of Social Security eligibility) from equation (3), where the sample consists of children in our Numident death-census linked sample who are additionally linked to the 1950 Census and the outcome is indicated in the column title. Panels show separate estimates for all children, sons and daughters. All regressions include fixed effects for children's sex-by-birth cohort and fathers' 3-digit occupation-by-birth cohort, Social Security coverage category, race, and birthplace. In column 5, means of 1950-based IPUMS occupational income scores in 1949 dollars are shown. Observations weighted using inverse propensity score weights that re-weight the sample to match the population of interest in 1930. Robust standard errors in parentheses, clustered by fathers' 3-digit industry. *** 1%, ** 5%, * 10% significance.

Table A.11: Median earnings by state in 1950 relative to national median, men aged 25–54

	Median earnings by state in 1950 relative to US median, selected education levels							
	All years	4 years	6 years	8 years	10 years	12 years	14 years	16 years
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
United States	2,450	1,550	1,950	2,450	2,650	2,850	3,050	3,050
Alabama	-800	-400	-600	-700	-600	-300	-400	-300
Arizona	-300	0	-100	-350	-200	-200	-475	-100
Arkansas	-900	-500	-700	-1,000	-800	-500	-500	-300
California	300	400	100	100	200	200	0	0
Colorado	0	0	-100	-200	-100	-100	-300	-200
Connecticut	200	700	400	100	0	0	0	0
Delaware	-200	100	-300	-200	-100	-100	-50	500
District of Columbia	0	400	100	-100	-100	-100	-175	-50
Florida	-600	-200	-400	-600	-500	-300	-500	-450
Georgia	-900	-300	-500	-600	-600	-300	-400	-100
Idaho	100	400	50	-100	-100	0	-50	0
Illinois	400	700	500	300	400	200	0	0
Indiana	200	500	200	100	200	200	0	0
Iowa	100	400	100	-100	0	-100	0	0
Kansas	0	300	100	-300	-200	-200	-200	0
Kentucky	-700	-400	-600	-600	-350	-400	-500	-500
Louisiana	-700	-200	-400	-400	-300	-200	-300	0
Maine	-400	0	-100	-450	-400	-400	-500	-200
Maryland	0	100	100	0	0	0	0	100
Massachusetts	100	500	200	-100	-100	-100	-200	0
Michigan	500	1,000	700	400	400	250	50	100
Minnesota	100	400	0	-100	175	100	0	100
Mississippi	-1,200	-600	-900	-1,200	-1,000	-800	-500	-100
Missouri	-400	-300	-200	-400	-200	-300	-300	-200
Montana	150	300	100	100	200	100	-200	200
Nebraska	-100	150	-100	-400	-200	-300	-100	0
Nevada	600	800	100	300	400	500	500	850
New Hampshire	-100	200	100	-300	-300	-300	-300	0
New Jersey	300	500	500	200	200	200	0	400

(table continues on next page)

Notes: This table shows deviations in median earned income by state in 1950 relative to the national median among men aged 25–54, separately for selected education levels. Earned income is defined as the sum of wage income and business or farm income. All amounts expressed in 1949 dollars.

Table A.11 (cont.): Median earnings by state in 1950 relative to national median, men aged 25–54

	Median earnings by state in 1950 relative to US median, selected education levels							
	All years (1)	4 years (2)	6 years (3)	8 years (4)	10 years (5)	12 years (6)	14 years (7)	16 years (8)
New Mexico	-400	-100	-400	-400	-300	-100	0	200
New York	150	500	200	100	100	50	0	0
North Carolina	-700	-200	-400	-600	-600	-400	-500	-250
North Dakota	-200	100	-100	-350	-150	-200	-150	200
Ohio	300	600	450	200	300	200	50	50
Oklahoma	-400	-300	-400	-400	-300	-200	-250	0
Oregon	350	500	400	100	200	200	0	0
Pennsylvania	100	500	300	0	-100	-100	-100	0
Rhode Island	-100	500	100	-100	-200	-200	-200	300
South Carolina	-900	-400	-400	-600	-600	-400	-500	-200
South Dakota	-300	-50	-400	-400	-300	-300	-200	100
Tennessee	-800	-400	-600	-900	-600	-300	-350	-300
Texas	-400	-50	-200	-400	-200	0	0	0
Utah	500	500	500	300	300	200	0	400
Vermont	-300	-100	-200	-500	-400	-300	-500	0
Virginia	-400	-100	-200	-400	-200	-100	-50	300
Washington	450	400	300	200	300	200	100	100
West Virginia	-200	400	100	-250	-200	-200	-400	-50
Wisconsin	200	500	400	100	200	100	0	0
Wyoming	500	600	300	200	400	300	200	600

Notes: This table shows deviations in median earned income by state in 1950 relative to the national median among men aged 25–54, separately for selected education levels. Earned income is defined as the sum of wage income and business or farm income. All amounts expressed in 1949 dollars.

Table A.12: The impact of father's exposure to Social Security on sons' educational attainment, veteran status and union status

	Dependent variable:					
	Years of education			Veteran status		Union status
	1940	1950	1973	1950	1973	1973
	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Sons born in 1907–1930 in 1973 OCG sample						
Social Security exposure	—	—	0.058	—	0.004	-0.027
			(0.168)		(0.027)	(0.022)
Mean of dep. variable			11.2		0.61	0.34
N			2,929		3,037	3,007
Panel (b): Sons born in 1907–1919 in 1973 OCG sample						
Social Security exposure	—	—	0.029	—	-0.010	-0.030
			(0.174)		(0.032)	(0.034)
Mean of dep. variable			10.7		0.48	0.33
N			1,693		1,754	1,732
Panel (c): Sons in Numident death-census linked sample who are additionally linked to 1940 or 1950						
Social Security exposure	-0.018	-0.026	—	0.000	—	—
	(0.017)	(0.019)		(0.004)		
Mean of dep. variable	11.0	11.1		0.61		
N	83,507	51,101		51,149		

Notes: This table shows the coefficients on father's Social Security exposure (i.e., negative of earliest age of Social Security eligibility) from equation (3), where the outcome is indicated in the column title. In panels (a) and (b), the samples are sons born in 1907–1930 or 1907–1919 in our 1973 OCG sample. In panel (c), the sample consists of sons in our Numident death-census linked sample who are additionally linked to the 1950 Census (columns 2 and 4) or to the 1940 Census via linkages from Bailey et al. (2026) (column 1). In column 1, the sample is restricted to sons aged 25 or older and not attending school in 1940. In column 2, the sample is restricted to sample-line sons aged 25 or older and not attending school in 1950. All regressions include fixed effects for sons' birth cohort and fathers' 3-digit occupation-by-birth cohort, Social Security coverage category, race, and birthplace. Observations weighted using inverse propensity score weights that re-weight the sample to match the population of interest in 1930. Robust standard errors in parentheses, clustered by fathers' 3-digit industry. *** 1%, ** 5%, * 10% significance.

Table A.13: Linking families in SS-5 records to households in the 1930 Census: Model features

Stage 1 model features (at the set level)
Description
Gap between highest and second-highest JW score
Max JW score
Mean JW score
Absolute difference in number of matching children (method 1) between highest and second-highest JW score potential
Gap between highest and second-highest father last JW score
Max father last JW score
Mean father last JW score
Absolute difference in number of matching children (method 1) between highest and second-highest father last JW score potential
Unique potential with 1+ matching children (method 1)
Unique potential with 2+ matching children (method 1)
Unique potential with max JW score or max father last JW score & 1+ matching children (method 1)
Unique potential with max JW score or max father last JW score & 2+ matching children (method 1)
Unique potential with max JW score or max father last JW score & father or mother middle initial match & 1+ matching children (method 1)
Unique potential with max JW score or max father last JW score & father or mother middle initial match & 2+ matching children (method 1)
Unique potential with 1+ matching children (method 1) & 1 matching child (method 2)
Unique potential with 1+ matching children (method 1) & 2+ matching children (method 2)
Unique potential with max JW score or max father last JW score & 1+ matching children (method 1) & 1 matching child (method 2)
Unique potential with max JW score or max father last JW score & 1+ matching children (method 1) & 2+ matching children (method 2)
Unique potential with max JW score or max father last JW score & father or mother middle initial match & 1+ matching children (method 1) & 1 matching child (method 2)
Unique potential with max JW score or max father last JW score & father or mother middle initial match & 1+ matching children (method 1) & 2+ matching children (method 2)
Unique potential with max JW score or max father last JW score & 1+ matching children (methods 1 or 2) & 1+ matching children (method 3)
Unique potential with max JW score or max father last JW score & 1+ matching children (methods 1 or 2) & father or mother middle initial match & 1+ matching children (method 3)
Unique potential with max JW score or max father last JW score & 1+ matching children (methods 1 or 2) & 1+ matching children (method 4)
Unique potential with max JW score or max father last JW score & 1+ matching children (methods 1 or 2) & father or mother middle initial match & 1+ matching children (method 4)

Stage 2 model features (at the pair level)	
Description	Description
JW score	Modal race match
Father last JW score	Father middle initial match
Number of matching children (method 1)	Mother middle initial match
Number of matching children (method 2)	Max JW score or max father last JW score & 1+ matching children (method 1)
Number of matching children (method 3)	Max JW score or max father last JW score & 1+ matching children (method 2)
Number of matching children (method 4)	Max JW score or max father last JW score & 1+ matching children (method 3)
Rank	Max JW score or max father last JW score & 1+ matching children (method 4)
Rank = 1	

Notes: "JW score" is the average of three Jaro-Winkler scores comparing the following strings: (1) father first names, (2) father last names, and (3) mother first names. Two children match according to method 1 if three conditions hold: (1) the Jaro-Winkler score between their first names is above 0.8 or the Levenshtein distance between their first names is within 2, (2) their absolute age difference is within 2 years, and (3) their birthplaces match. Method 2 allows for inverted first and middle names, method 3 allows for matches on initials (both first and middle name initials must match), and method 4 is a combination of methods 2 and 3.

Table A.14: Key assumptions and sources for the derived effects on lifetime earnings, family gains, and fiscal implications (Section 6)

Assumption	Baseline value	Source / rationale
Sons' earnings gains		
Baseline late-career gain g	28.2% (SE 9.5%)	ZIP-income effect 0.024 / HRS gradient 0.085 (Appendix D)
Working-life span	Ages 20–65, 1910 cohort	Standard
Interpolation knots	0 at first job & 1950; g in 1962 & 1972; held at g after	Mirrors occscore profile (Section 5.2)
Baseline earnings path	Median male four-quarter-worker annual earnings	SSB Annual Statistical Supplements (Social Security Administration, 1965, 1975)
Real discount rate	3%, discount to 1950	Standard
Son survival	1910 male cohort, conditional on alive in 1930	SSA cohort life table; matches sampling frame
Sons per recipient family	2.3	Sample summary statistics
Social Security benefits		
Earnings / coverage	Median covered wage, 1877 cohort	Simulation assumption
PIA	Statutory formulas by year, 1937 vs. 1951 eligibility	SSA
Reform	10-year-earlier earliest eligibility	1935- vs. 1950-covered employment for 1877 cohort
Recipient labor supply	1930 LFP profile, −25% after 1937 earliest eligibility	Baseline assumption
Couple / survivor benefit	150% / 75% of PIA	Early-program rule
Parent survival	Conditional on couple alive in 1930 (ages 53 / 50)	Sampling frame; Human Mortality Database cohort tables
Spouse age gap	Wife 3 years younger	Sample summary statistics
Fiscal / welfare		
First stage ϕ	1 (upper bound)	Conservative; overstates Social Security benefits and fiscal costs relative to children's gains
Recipient tax rate $\overline{\tau}_r$	18.8%	Historical tax policies (see Appendix E.1)
Son tax rate $\overline{\tau}_k$	25.3%	Historical tax policies (see Appendix E.1)

Notes: Baseline specification. See Appendix E for details.

Table A.15: Robustness of the derived effects on lifetime earnings, family gains, and fiscal implications

		Labor supply responses				Policy parameters			Benefit and gain magnitudes		
	Benchmark	LFP 40%	LFP 50%	Son resp.	Recip. resp.	No c taxes	OASI offset UB	Survivor	SS $\times$1.25	SS $\times$1.5	Gain 33/21
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Implied individual earnings gain											
Late career (1962 onward)	28.2%	28.2%	28.2%	21.2%	28.2%	28.2%	28.2%	28.2%	28.2%	28.2%	27.0%
Full-career average	11.9%	11.9%	11.9%	8.9%	11.9%	11.9%	11.9%	11.9%	11.9%	11.9%	11.7%
EPV per son	$130K	$130K	$130K	$97K	$130K	$130K	$130K	$130K	$130K	$130K	$132K
Implied family gains (per recipient family) and fiscal implications											
EPV of children's earnings gain	$299K	$299K	$299K	$224K	$299K	$299K	$299K	$299K	$299K	$299K	$303K
$\div$ ΔEPV(SS)	6.4	5.6	5.1	4.8	6.4	6.4	6.4	5.6	5.1	4.3	6.5
Net family earnings $\div$ ΔEPV(SS)	5.7	4.5	3.9	4.1	5.7	5.7	5.7	5.0	4.5	3.8	5.7
Net govt. cost $\div$ mech. effect	−0.55	−0.25	−0.08	−0.08	−0.39	−0.11	−0.34	−0.34	−0.21	0.02	−0.57

Notes: Amounts in 2020 dollars, present values as of 1950. Gain estimate based on ZIP income gain throughout; benchmark scenario except as noted (see Table 9 and Appendix Section E), including the conservative assumption of a first stage of one (i.e., differences in predicted eligibility translate one-for-one into differences in actual eligibility). Each column changes one assumption relative to the benchmark. LFP 40%: fathers' post-65 labor-force-participation reduction increased to 40% (benchmark 25%). LFP 50%: fathers' post-65 labor-force-participation reduction increased to 50%. Son resp.: sons' estimated earnings response scaled by 0.75. Recip. resp.: recipients' estimated earnings response doubled (e.g., from doubling their counterfactual earnings path while holding fixed their LFP response). No c taxes: consumption/property taxes excluded; rates are 12.3% for recipients and 18.5% for sons (vs. baseline rates of 18.8% and 25.3%, respectively). OASI offset upper bound (UB): sons' future OASI benefit offset set to the family (spouse and survivor) upper bound rather than the worker-only amount; rates are 18.8% for recipients and 22.5% for sons (vs. baseline rates of 18.8% and 25.3%, respectively). Survivor: survivor benefit is 100% of PIA (benchmark 75%). SS $\times$1.25: recipients' Social Security benefits scaled to 125%. SS $\times$1.5: recipients' Social Security benefits scaled to 150%. Gain 33/21: unequal late-career gains (linear decline from 33% in 1962 to 21% in 1972 and continued linear decline thereafter; based on OCG estimates in Table 5).

Table A.16: Historical average net tax rates on earnings responses

Fiscal component	Rate
Panel A. Recipients	
Federal income and OASDI/HI payroll taxes	10.3%
State and local income taxes	0.3%
Unemployment insurance taxes	1.7%
Earnings-side subtotal	12.3%
Broad consumption and property taxes	6.6%
Preferred net tax rate	18.8%
Panel B. Sons	
Federal income and OASDI/HI payroll taxes	18.7%
State and local income taxes	1.9%
Unemployment insurance taxes	0.0%
Earnings-side subtotal before OASI benefit accrual	20.6%
OASI benefit accrual	−2.1%
Earnings-side subtotal, net	18.5%
Broad consumption and property taxes	6.8%
Preferred net tax rate	25.3%

Notes: Each annual rate is the change in tax liability per dollar of the modeled earnings change and is weighted by that year's present-value earnings response. OASDI/HI payroll taxes include employee and employer OASDI/HI shares and apply annual taxable maxima. The recipient state-income component is a BEA average effective rate; the son component is the Barro–Redlick marginal rate. UI taxes include state UI contributions and net FUTA contributions where applicable. The preferred son UI component is zero because the sons' incremental earnings are above the federal wage base and, in nearly all relevant state-year cells, above the state wage base. The baseline son household claims two dependents. Broad consumption and property taxes assume full eventual consumption of disposable earnings. Components may not sum exactly because of rounding. See Appendix Section E.1 for details.

Online Appendix References

Abramitzky, Ran, Leah Boustan, Katherine Eriksson, Myera Rashid, and Santiago Pérez, *Census Linking Project: Version 3.0 [dataset]*, Harvard Dataverse, 2025.

Altmeyer, Arthur J., "The Desirability of Extending Social Security to Employees of Nonprofit Institutions," *Social Security Bulletin*, 1944, *7* (8).

_ , "Social Security for State and Local Government Employees," *Social Security Bulletin*, 1945, *8* (4), 4–7.

American Hospital Association, "For Flexibility: Changes in Pension Plan By-Laws," *Hospitals: the Journal of the American Hospital Association*, 1949, *23* (2), 61.

Bailey, Martha J., A. R. Shariq Mohammed, and Paul Mohnen, "The Evolution of U.S. Educational Mobility over the 20th Century and the Role of Public Education," Atlanta Fed Working Paper 2026-1 2026.

__ **, Peter Z. Lin, A. R. Shariq Mohammed, Paul Mohnen, Jared Murray, Mengying Zhang, and Alexa Prettyman**, "The creation of LIFE-M: The Longitudinal, Intergenerational, Family Electronic Micro-Database project," *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 2023, *56* (3), 138–159.

Barro, Robert J. and Charles J. Redlick, "Macroeconomic Effects from Government Purchases and Taxes," *Quarterly Journal of Economics*, 2011, *126* (1), 51–102.

Bugbee, George, "Pensions Plan Is Now Open for Enrollment," *Hospitals: the Journal of the American Hospital Association*, 1946, *20* (9), 59–60.

Bureau of Economic Analysis, "National Income and Product Accounts, Tables 1.1.5, 2.1, 3.4, and 3.5," Downloadable NIPA tables. https://apps.bea.gov/iTable/?ReqID=19&step=2 2026. Accessed June 19, 2026.

Clark, Robert L., Lee A. Craig, and Jack W. Wilson, *A History of Public Sector Pensions in the United States*, University of Pennsylvania Press, 2003.

__ **,** __ **, and John Sabelhaus**, *State and Local Retirement Plans in the United States*, Edward Elgar, 2011.

Collins, William J. and Marianne H. Wanamaker, "African American Intergenerational Economic Mobility since 1880," *American Economic Journal: Applied Economics*, 2022, *14* (3), 84–117.

Fang, Chichun, "Health and Retirement Study Imputation of Lifetime Earnings," Working paper, Institute for Social Research, University of Michigan 2021.

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

Hayhow, Edgar C., "Pensions and Life Insurance Plans for Hospital Employees," *Hospitals: the Journal of the American Hospital Association*, 1940, *14* (11), 73–77.

Health and Retirement Study, *(Imputation of Lifetime Earnings Records) restricted dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740)*, Ann Arbor, MI, 2021.

__ , *(Cross-Wave Geographic Information (Detail) [1992–2020]) restricted dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740)*, Ann Arbor, MI, 2023.

__ , *(Cross-Wave Tracker File [1992–2020]) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740)*, Ann Arbor, MI, 2024.

Internal Revenue Service, "Prior Year Forms and Instructions," https://www.irs.gov/prior-year-forms-and-instructions 2026. Accessed June 19, 2026.

Kestenbaum, Bert, "State and Local Government Employees Covered Under Social Security, 1977–81," *Social Security Bulletin*, 1982, *45* (12), 11–13.

Leibowitz, George J., "Old-Age and Survivors Insurance: Coverage Under the 1950 Amendments," *Social Security Bulletin*, 1950, *13* (12), 3–10, 21.

Marquis, James E., "Old-Age and Survivors Insurance: Coverage Under the 1954 Amendments," *Social Security Bulletin*, 1955, *18* (1), 3–10.

McCamman, Dorothy, "State and Local Employees Covered by Government Retirement Systems," *Social Security Bulletin*, 1943, *6* (7), 31–41.

__ , "Higher Educational Institutions and Social Security," *Social Security Bulletin*, 1948, *11* (1), 5–11.

—, "Retirement Protection for State and Local Employees: Ten Years of Growth," *Social Security Bulletin*, 1953, *16* (5), 3–10,24.

—, "State and Local Government Employment Under OASI, October 1955," *Social Security Bulletin*, 1956, *19* (2), 15–17.

McDaniel, Cara, "Average Tax Rates on Consumption, Investment, Labor and Capital in the OECD, 1950–2003," Technical Report, Arizona State University 2007.

Murray, Merrill G. and Ilse M. Smith, "Higher Educational Institutions and the Social Security Act," *Social Security Bulletin*, 1940, *3* (12), 21–28.

Nakamura, Emi, Jósef Sigurdsson, and Jón Steinsson, "The Gift of Moving: Intergenerational Consequences of a Mobility Shock," *The Review of Economic Studies*, 2022, *89* (3), 1557–1592.

National Bureau of Economic Research, "Internet TAXSIM Version 35," https://taxsim.nber.org/taxsim35/ 2026. Accessed June 19, 2026.

National Education Association, *Statutory Provisions for Statewide Retirement Systems*, Washington, D.C.: National Education Association of the United States, 1946.

Nelson, William, "Employment Covered Under the Social Security Program, 1935-84," *Social Security Bulletin*, 1985, *48* (4), 33–39.

Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Grace Cooper, Julia A. Rivera Drew, Stephanie Richards, Renae Rogers, Jonathan Schroeder, and Kari C.W. Williams, *IPUMS USA: Version 16.0 [dataset]*, Minneapolis, MN: IPUMS, 2025.

Schottland, Charles I., "Social Security Amendments of 1956: A Summary and Legislative History," *Social Security Bulletin*, 1956, *19* (9), 3–31.

Social Security Administration, "Social Security Bulletin, Annual Statistical Supplement, 1965," 1965. U.S. Government Printing Office.

—, "Social Security Bulletin, Annual Statistical Supplement, 1975," 1975. U.S. Government Printing Office.

—, "Social Security Bulletin, Annual Statistical Supplement 1977–79," 1980. U.S. Government Printing Office.

—, "History of OASDI Benefit Computation and Automatic Adjustments," Annual Statistical Supplement, Tables 2.A8–2.A19. https://www.ssa.gov/policy/docs/statcomps/supplement/2023/2a8-2a19.html 2023.

—, "Contribution and Benefit Base," Office of the Chief Actuary. https://www.ssa.gov/oact/cola/cbb.html 2026. Accessed June 19, 2026.

—, "FICA and SECA Tax Rates," Office of the Chief Actuary. https://www.ssa.gov/oact/progdata/taxRates.html 2026. Accessed June 19, 2026.

Subcommittee on Social Security, "Background Material on Social Security Coverage of Governmental Employees and Employees of Nonprofit Organizations," in "Coverage and Termination of Coverage of Government and Nonprofit Organization Employees under the Social Security System: Public Hearings before the Subcommittee on Social Security of the Committee on Ways and Means, U.S. House of Representatives, Ninety-Fourth Congress, Second Session," U.S. Government Printing Office, 1976.

Tax Foundation, "Historical U.S. Federal Individual Income Tax Rates and Brackets, 1862–2025," https://taxfoundation.org/data/all/federal/historical-income-tax-rates-brackets/ 2025. Accessed June 19, 2026.

United States Bureau of the Census, *Retirement Systems for State and Local Government Employees: 1941*, U.S. Department of Commerce, Bureau of the Census, 1943.

__ , *Retirement Coverage of State and Local Government Employees*, Washington, D.C.: U.S. Department of Commerce, Bureau of the Census, 1953.

__ , "National Intercensal Tables, 1900–1990: Population by age, sex, and race," https://www.census.gov/data/tables/time-series/demo/popest/pre-1980-national.html 2004. Accessed: 2020-08-10.

Urban-Brookings Tax Policy Center, "Historical Individual Income Tax Parameters," https://taxpolicycenter.org/statistics/historical-individual-income-tax-parameters 2023. Updated May 11, 2023; accessed June 19, 2026.

U.S. Bureau of the Census, *Alphabetical Index of Occupations and Industries: 1930*, Washington, D.C.: U.S. Government Printing Office, 1950.

U.S. Department of Labor, "A Comparative Analysis of Unemployment Insurance Financing Methods," Technical Report ETAOP 2018-08, Employment and Training Administration 2018.

__ , "ET Handbook No. 394: Unemployment Insurance Financial Data," Employment and Training Administration. https://oui.doleta.gov/dmstree/handbooks/394/lnkus.html 2026. Historical annual tables; accessed June 19, 2026.

Whittaker, Julie M., "Unemployment Compensation: The Fundamentals of the Federal Unemployment Tax (FUTA)," Technical Report R44527, Congressional Research Service 2016.