The Evolution of Late-Life Income and Assets: Measurement in IRS Tax Data and Three Household Surveys

James J. Choi
Yale University and NBER

Lucas Goodman
US Department of the Treasury

Justin Katz
Harvard University

David Laibson
Harvard University and NBER

Shanthi Ramnath
Federal Reserve Bank of Chicago

November 23, 2020

Abstract: Using a 5% random sample of administrative IRS tax records covering households born from 1933 to 1952, we evaluate how three widely-used household surveys—the Health and Retirement Study, the Survey of Income and Program Participation, and the Current Population Survey—capture the level of and trends in late-life income and assets. First, relative to the tax data, survey data underestimate total income levels and overestimate declines in income at the median during the initial transition from working life to retirement. Survey estimates of median income at age 73 are lower than tax data estimates by an average of 4.5% in the HRS, 14.2% in the SIPP, and 25.1% in the CPS. Median total income declined from 58 to 68 by an average of only 11.7% in the tax data, compared with 24.4% in HRS, 16.8% in SIPP, and 29.0% in CPS. Second, survey sources overestimate income growth across birth cohorts at older ages but do a better job of capturing these trends at younger ages. Third, lower-income households have not experienced income growth across birth cohorts outside of the Social Security system. Averaging across ages 68 to 74, the 25th percentile income excluding Social Security fell by 16.5% from the 1933 birth cohort to the 1943 birth cohort in the tax data. These declines are larger in the HRS (26.9%) and SIPP (45.5%) and smaller in the CPS (11.1%). The fraction of households in the tax data with no non-Social Security income and no assets at age 72 rose from 18.9% to 20.5% from cohorts born in 1933 to 1945. The fraction of such households is captured well by the HRS and SIPP, but overstated by the CPS.

How well do widely used US household surveys—the Health and Retirement Study (HRS), the Survey of Income and Program Participation (SIPP), and the Current Population Survey (CPS)—capture the level of and trends in late-life income and assets? In this paper, we compare estimates from these surveys to those from a 5% random sample of 2001-2017 IRS tax records covering the 1933-1952 birth cohorts. IRS data contain administrative records on income sources such as pensions and IRAs, which may be severely misreported in household surveys (Bee and Mitchell 2017). IRS data hence offer a unique benchmark to assess bias in survey estimates.

We have three main findings. First, for middle-income households' pre-tax income (excluding capital gains), IRS data show higher levels in early retirement and smaller declines during the initial transition to retirement compared to estimates from survey data. Averaging across the 1933-1939 birth cohorts, survey estimates of median pre-tax income at age 73 are lower than tax data estimates by 4.5% in the HRS, 14.2% in the SIPP, and 25.1% in CPS. Averaging across the 1943-1949 birth cohorts, median pre-tax income declined from ages 58 to 68 by only 11.7% in the tax data, compared with 24.4% in HRS, 16.8% in SIPP, and 29.0% in CPS. Therefore, relying exclusively on survey data to evaluate income levels throughout retirement or trends as households age during early retirement will tend to produce overly pessimistic results about the median.

These patterns are driven by surveys overestimating working-life income and underestimating retirement income. Specifically, wage income at age 58 is overestimated in all three surveys. In contrast, all three surveys tend to significantly underestimate pension income and slightly underestimate Social Security income. HRS estimates of the overall pre-tax income distribution tend to align most closely with the tax data, especially at older ages, but the SIPP often more closely matches the tax data when measuring individual income components. This discrepancy is due to offsetting patterns over- and under-estimation of individual income components in the HRS.

We also compare the distribution of IRA balances in the tax data with levels reported in the HRS and the SIPP. We find that reported levels in the HRS are quite accurate, especially at older ages, while the SIPP significantly underestimates balances.

Second, surveys tend to overestimate income growth across birth cohorts at older ages, but do a better job of capturing these trends at younger ages. Averaging across ages 68-74, median pre-tax income (excluding capital gains) grew 16.3% in the tax data from the 1933 birth cohort to

the 1943 birth cohort, compared to an increase of 21.7% in the HRS, 23.5% in the SIPP, and 23.6% in the CPS. On the other hand, averaging across ages 58-67, median pre-tax income (excluding capital gains) grew by -0.2% from the 1944 birth cohort to the 1950 birth cohort, compared to -7.8% in the HRS, -0.7% in the SIPP, and -0.9% in the CPS. Overall, relying exclusively on survey data will produce an overly *optimistic* picture of across-cohort trends in retirement income.

Finally, both IRS and survey data suggest that later-life income outside of the Social Security system has been falling for low-income households with each successive birth cohort. Averaging across ages 68 to 74, the 25th percentile income excluding Social Security fell by 16.5% from the 1933 birth cohort to the 1943 birth cohort in the tax data. Survey data also show declines that are larger in the HRS (26.9%) and SIPP (45.5%) and smaller in the CPS (11.1%). IRS data show that the fraction of tax units at age 72 with no financial assets and no non-Social Security income rose from 18.9% in the 1933 birth cohort to 20.5% in the 1945 birth cohort, the most recent year for which measurements are available. Survey data trends are qualitatively similar. The tax data also reveal substantial geographic heterogeneity in the fraction with no assets and no non-Social Security income—for the 1933 birth cohort at age 72, this fraction ranges from 11.1% at the 10th percentile state (Kansas) to 22.1% at the 90th percentile state (Georgia).

This paper builds on previous work evaluating the accuracy of survey data using administrative records. Compared with papers that explicitly link survey respondents to administrative data (e.g. Bee and Mitchell 2017, Dushi, Iams, and Trenkamp, 2017), we are able to: (i) evaluate trends over a longer time period and in a wider variety of surveys, as we are not restricted by the ability to link to survey data; and (ii) compare population-level trends, as we are not restricted to the subset of linked respondents. Compared with papers that compare aggregate income in household surveys to public administrative aggregates (e.g. Chen, Munnell, and Sanzenbacher, 2018), our access to administrative microdata allows for more precise distributional comparisons of a larger set of income and asset variables.

I. Constructing harmonized datasets in tax and survey sources
A. Data and variable construction
IRS tax data. We construct our IRS sample using tax data from 2001-2017 and the Death Master File from the Social Security Administration (SSA). The SSA data include information on date of birth, sex, and date of death. We draw a 5% random sample of individuals born from 1933-1952

who appear in the tax data at least once during our sample period using the last four digits of a person's Tax Identification Number (TIN),[1] dropping individuals after they die. For state-level analysis, we supplement this sample with a larger draw such that we observe at least 7,000 living people within each birth cohort, state, and year.[2]

We construct income variables using (self-reported) tax return data and (third-party reported) information returns data. All dollar amounts are deflated to 2010 dollars using the Consumer Price Index. To characterize sources of income as households age, we group income into five categories (specifying the form from which the income is reported in parentheses):

(i)   Wage income (Form W-2)

(ii)   Pension income: the sum of distributions from defined benefit and defined contribution plans (Form 1099-R)[3]

(iii)   Social Security income, including OASI and DI (Form SSA-1099)

(iv)   Capital income: the sum of taxable and tax-exempt interest income (Form 1040), taxable dividend income (Form 1040), and capital gains (Schedule D)

(v)   Other income: the sum of income from rent, royalties, partnerships and S-corporations, estates, and trusts (Schedule E), income from sole-proprietorship income (Schedule C), income from unemployment insurance (Form 1040), and net alimony income (Form 1040)

For the income items derived from Form 1040, we override the Form 1040 amount in certain cases. If an individual is a non-filer or reports zero for a particular component, we typically use the information return amount in its place. Specifically, we use Form 1099-DIV for dividend income, Form 1099-INT for taxable interest income, Form 1099-G for unemployment compensation, and Form 1099-MISC (non-employee compensation) for sole proprietorship income.[4]

---

[1] Technically, we used a person's masked TIN, which is a deterministic scrambling of their TIN.
[2] State-level results concern the fraction of tax units with no IRA and negligible non-Social Security income. Our target of 7,000 observations per birth cohort and year ensures the standard error on the estimate of this fraction does not exceed 0.6%. In some small states, a 100% pull results in fewer than 7,000 observations.
[3] Pension income includes non-taxable distributions but excludes direct rollovers, Section 1035 exchanges, withdrawals of impermissible contributions, and recharacterized contributions.
[4] The one exception is tax-exempt interest, where we lack relevant information throughout part of our study period. For consistency, we only use self-reported values for tax-exempt interest income. Additionally, we do not attempt to use information returns to calculate Schedule D or Schedule E income.

Based on these income components, we construct several aggregate variables. Pre-tax income sums income in categories (i)-(v). After tax-income is pre-tax income, minus federal tax liability reported on Form 1040 and estimated FICA taxes on wages, plus any earned income tax credit or refundable child tax credits reported on Form 1040. For comparisons with survey data, we construct versions of pre-tax income and after-tax income that exclude capital gains income reported on Schedule D. All total income measures exclude Supplemental Security Income (SSI), private insurance payouts, and SNAP, WIC, TANF, and VA benefits because these are not available in the tax data.

We measure IRA balances held in both traditional and Roth IRAs from an information return (Form 5498).

We classify households as receiving income in a certain category if income in that category exceeds $500. We classify households as having capital income either if capital income exceeds $500 or if interest income exceeds $10,000 times the average effective federal funds rate in the year of measurement. This smooths out variation in interest income on fixed asset holdings due to exogenous changes in the interest rate environment.

All variables are constructed at the tax unit level, summing individuals' income and assets with the income and assets of their spouse. We adjust for household size by dividing income and assets by the square root of the number of people in the tax unit (either 1 or 2, since we do not consider dependents). We are not able to link spouses in years in which a person does not file a return. In such years, we impute spousal links using tax returns from other years when they appear as either a primary or secondary filer on a joint return. Specifically, for a non-filer in year t, we check whether the person appeared on a joint return from $t-4$ through $t+1$, and assign a spouse based on the joint filing closest in time to t, prioritizing status in $t-1$ over status in $t+1$.[5] If the person does not appear on a joint filing during that window, we assume they are single in all years.

Survey data. We prepare versions of the HRS, SIPP, and CPS datasets to align as closely as possible with the tax data. For HRS, we use the RAND HRS 2016 Longitudinal and Detailed Imputation files, RAND HRS Fat Files from 1992-2016, and the HRS public use dataset produced by the University of Michigan. For CPS, we pull the March ASEC supplement from 1992-2018

---

[5] Data on tax filers runs from 1996-2018, meaning this imputation does not introduce truncation issues at the start or end of the panel.

produced by IPUMS at the University of Minnesota. For SIPP, we merge the 1996, 2001, 2004, 2008, and 2014 panels using files maintained at the NBER. While we focus on people born from 1933-1952 observed from 2001-2017—consistent with the tax data sample—we also compute results for all available cohorts and years to analyze longer-horizon trends.

First, we construct the unit of analysis in the survey data to approximate tax units in the IRS data. We cannot directly identify tax units in the survey data, as we typically do not observe filing status. To maximize comparability, we group two survey respondents into a single economic unit if they are identified in the survey as being married, which typically requires that a married couple is cohabiting at the time of the survey.[6] This grouping will fail to link married couples living separately but filing jointly, who *will* be linked in a single tax unit in the tax data. This grouping will incorrectly link cohabiting married non-filers who did not file jointly in the previous four years or subsequent year, who will *not* be linked in a single tax unit in the tax data. Throughout the paper, we refer to units of analysis in the tax data and survey data as "households." We often identify fewer two-person households in the tax data than in the survey data. For example, for the 1933 birth cohort at age 68, 59.7% of IRS tax units have two persons, versus 65.1% in the HRS, 67.3% in the SIPP, and 70.1% in the CPS. On the other hand, the tax data are more aligned at older ages; for the 1933 birth cohort at age 76, the two-person household share is 54.3% in the IRS data, 58.5% in the HRS, 52.6% in the SIPP, and 60.8% in the CPS.

In the CPS, we lack information on date of birth, so cannot always associate individuals with a unique birth cohort. We do know the exact age of respondents up to age 79.[7] Households are interviewed in March and asked about their income over the previous calendar year; we assign birth cohorts by assuming that respondent age has not changed since December 31 of the income reference year. Therefore, around a quarter of individuals assigned to, for example, the 1935 birth cohort were in fact born in 1936. We drop observations of individuals aged 80 and up.

Next, we construct income and asset variables to match definitions in the tax data. In all sources, we exclude income from SSI, VA pensions, and other government transfers. All survey

---

[6] There are some exceptions to this in SIPP and HRS due to their longitudinal structure. If members of a married couple are cohabiting when they are initially selected to participate in the survey, but are no longer cohabiting in later waves, SIPP and HRS will follow each member of the original couple. It is possible to link the survey records for such married couples if they stay married even once they are no longer cohabiting. In such circumstances, we do link non-cohabiting married couples in SIPP and HRS.

[7] Ages are reported in bins for 80-85, 85-90, etc.

measures of total income and capital income exclude capital gains and losses,[8] so when making direct comparisons with survey data, we use pre-tax income in the tax data that excludes capital gains and losses. Due to the structure of each survey, some definitions vary slightly as follows:

1. In the CPS and SIPP, pension income includes income from paid-up insurance policies, which are not observed in the tax data or the HRS.[9]

2. For some years, CPS pension income is constructed differently than in other sources. The HRS, SIPP, and tax data separately report pension income from each component—income from defined benefit and defined contribution employer-sponsored pensions; income from IRA withdrawals; and income from annuities. Until 2014, the CPS asked broadly about up to two sources of regular retirement income. Households were not asked about each income source separately, irregular withdrawals from retirement accounts, or income beyond the first two sources mentioned. From 2014 onwards, the CPS asked separate questions about retirement account withdrawals and distributions, consistent with other sources.[10]

3. In the HRS, pension income and other income are constructed differently than in other sources. Regarding pension income, the HRS only asks about total IRA withdrawals since the previous interview—a period of approximately two years. We adjust this amount to reflect withdrawals over a 12-month period, assuming a constant withdrawal rate. Regarding other income, the HRS stopped separately asking about alimony income starting in 2004 (HRS wave 7), and instead allowed respondents to report

---

[8] The CPS does not ask about income from capital gains and losses. The HRS asks questions that would allow us to calculate changes in the value of real estate and net asset holdings across waves. However, we lack data on initial purchase prices for financial assets, making any measure of capital gains not comparable with the tax data. In some panels, the SIPP asked about capital gains reported on tax forms. We do not use SIPP capital gains variables for two reasons: (i) respondents who report capital losses are not asked the amount of capital losses reported, inconsistent with the tax data; and (ii) the SIPP stopped asking about capital gains income in the 2014 panel, meaning capital gains income is missing from 2013-2017.

[9] Both surveys ask about total income from annuities and/or paid-up life insurance policies in a single question, meaning we cannot separate annuity income from paid-up life insurance policy income.

[10] In 2014, the CPS divided randomly divided the March ASEC Supplement sample into two groups. Three-eighths of the sample received an experimental questionnaire with redesigned pension income questions that asked separately about withdrawals from retirement accounts, and the remaining 5/8 received the old versions of questions. The CPS released separate survey weights to make either sample—when used individually—representative of the full US civilian noninstitutionalized population. We use the appropriately weighted three-eighths sample for 2014. Starting in 2015, the redesigned questionnaire was sent to all households.

alimony income in response to a question asking generally about "other income."[11] Alimony income is potentially significant for single divorced households. To include alimony income while preserving consistency in HRS across years, we include all income reported in response to the generic "other income" questions in our HRS measure of other income in all years.

4. Only the HRS and the SIPP include information on assets. In addition to IRA balances, we construct a more comprehensive variable describing household financial assets by summing balances held in IRAs, defined contribution employer-sponsored pensions, stocks, bonds, and generically specified "other savings."

B. Adjusting estimates to represent the US civilian noninstitutionalized population

The tax data do not represent the same population as each survey. To produce comparable estimates, we target the US civilian noninstitutionalized population in each dataset. This population includes all residents of the 50 states and Washington DC but excludes members of the military[12] and people in institutions such as prisons and nursing homes.

The CPS and the SIPP are already designed to represent the US civilian noninstitutionalized population, and hence do not require any additional restrictions. The HRS initially samples from the US civilian noninstitutionalized population, but continues to interview individuals as they move into nursing homes or move abroad. We therefore exclude observations for years in which HRS participants live outside the US or in nursing homes.

The tax data contain records for any individual who files US taxes or receives a US information return. This *excludes* members of the US civilian noninstitutionalized population who do not have contact with the tax system, but *includes* individuals who are outside of the US civilian noninstitutionalized population and hence not represented in all survey datasets—most importantly, people living abroad and residents of institutions such as nursing homes. The tax data also contain members of the military, who are not included in the survey data if they live in military

---

[11] Specifically, the HRS asks: "(Other than income you have already told me about,) did you [or your husband/wife/partner] receive any other income in [Last Calendar Year], for example, from private disability insurance payments, consulting fees, rent from your home or second home, odd jobs, and so forth? Do not include financial support from relatives or friends."

[12] Both the SIPP and the CPS include military personnel living in civilian housing units (not in military barracks).

barracks. Such individuals are exceedingly rare among the older population that we study, so we ignore the small discrepancy across datasets potentially induced by such individuals.[13]

We adjust the tax data by dropping observations for years in which members of our tax sample live outside the 50 US states and Washington DC, as identified by having a foreign address on either Form 1040 or Form 1042-S (an information return). We then add individuals who have no contact with the tax system to make the tax data representative of the US resident population. Finally, we adjust our estimates to exclude the institutionalized population.

Adding individuals with no contact with the tax system. Any person who does not appear in the tax data must have zero income and assets. We add enough synthetic zero-income individuals to the dataset such that the number of individuals at each age measured in each year matches estimates of the US resident population published by the Census Bureau.[14]

Adjusting estimates to exclude institutionalized individuals. The tax data include a non-negligible fraction of individuals living in institutions—most commonly, nursing homes—at older ages. However, we are not able to directly identify and remove institutionalized individuals from the tax data. We take a different approach to correct for potential bias introduced by the institutionalized population when estimating distributions of income and assets. First, we construct bounds on the estimates, imposing the extreme assumptions that the institutionalized population lies at either the top or bottom of income or asset distributions. Second, we compute baseline estimates by using estimates from survey data of the institutionalized population's relative position within the income and asset distributions.

To be concrete, we aim to estimate the distribution of income or asset variable Y for the non-institutionalized population (denoted by $I = 0$) in birth cohort τ observed at age a, given by the conditional CDF, $F(y|a,\tau,I = 0) = \Pr(Y \leq y|a,\tau,I = 0)$. We do not directly observe $F(y|a,\tau,I = 0)$. Instead, we observe the conditional CDF for the resident population, $F(y|a,\tau)$, which can be decomposed as follows:

---

[13] In 2018, there were only 3,349 active enlisted members across the entire Department of Defense aged 50-55, the oldest age for which counts are reported (US Department of Defense 2018, Table B-15). This is about 0.03% of the civilian labor force within that age range.

[14] Intercensal estimates for single year of age are anchored on the decennial Census, with annual administrative adjustments for births, deaths, and net migration abroad. There are some age-year cells where the count in the tax data exceeds the Census resident population estimate. We do not adjust the tax data in these instances.

$$F(y|a,\tau) = \Pr(Y \leq y|a,\tau, I = 0) \times \Pr(I = 0|a,\tau) + \Pr(Y \leq y|a,\tau, I = 1) \times \Pr(I = 1|a,\tau).$$

We estimate $\Pr(I = 0|a,\tau)$ and $\Pr(I = 1|a,\tau)$ using data published by the Census Bureau,[15] so with values for $\Pr(Y \leq y|a,\tau, I = 1)$, we can recover the conditional CDF for the non-institutionalized population. To obtain lower bounds for percentiles and levels of receipt, we assume $\Pr(Y \leq y|a,\tau, I = 1) = 1$, and to obtain upper bounds, we assume $\Pr(Y \leq y|a,\tau, I = 1) = 0$.

For our baseline estimates, we use values for $\Pr(Y \leq y|a,\tau, I = 1)$ estimated using data on the nursing home population in the HRS. Measurement error may bias estimates of income and asset levels in the HRS. We therefore proceed under the assumption that, conditional on birth cohort and age, the HRS correctly measures the *relative* position of the institutionalized population within the full resident distribution.[16] For a fine mesh of $n \in (0, 100)$, we compute the probability that an institutionalized HRS individual of age a in birth cohort τ has income lower than the nth percentile of HRS-measured resident income for age a and cohort τ. We do this by regressing a dummy for an institutionalized HRS individual having income less than the nth HRS resident percentile on a cubic polynomial in age and a cubic polynomial in birth cohort.[17] We then assume that an institutionalized individual has this same probability of having less than the nth percentile of income in the tax data.[18]

Appendix Figure A1 shows how in the tax data, for the 1941 birth cohort, the bounds and baseline estimates of the pre-tax income distribution of the resident noninstitutionalized population compare with the pre-tax income distribution of the resident population. Each panel plots the difference in log points between the estimated lower bound, baseline estimate, or upper bound of the resident noninstitutionalized population and the estimates from the resident population. For the 25th, 50th, and 75th percentiles, bounds are quite tight until around age 75—typically much less than ±3%. At later ages, as the institutionalized population grows, bounds become considerably

---

[15] The Census Bureau estimates the annual institutionalized population for each single year of age by anchoring on decennial Census values and adjusting annual estimates using administrative records for net migration into institutions.
[16] Our baseline estimates are robust to any rank-preserving measurement error, as well as age and birth cohort trends in measurement error that may not preserve rank.
[17] We lack sufficient observations in the HRS nursing home population to use a more flexible specification.
[18] The resulting estimates should be weakly monotone: $u < u'$ implies that $\Pr(Y \leq F^{-1}(u)|a,\tau, I = 1) \leq \Pr(Y \leq F^{-1}(u')|a,\tau, I = 1)$. Noise in coefficient estimates means this property may not be satisfied. We enforce monotonicity by re-ranking estimates using the method described in Chernozhukov, Fernandez-Val, and Galichon (2010).

wider. The relative position of baseline estimates indicates that the institutionalized population typically has pre-tax income between the 10th and 50th percentiles of the resident pre-tax income distribution. At the 10th percentile, baseline estimates are close to the lower bound, indicating that the lower bound assumption—that the institutionalized population has income above the 10th resident income percentile—is an accurate approximation. At and above the 25th percentile, baseline estimates tend to hug the upper bound, indicating that the upper bound assumption is a more accurate. Unless otherwise noted, we report baseline estimates in results.

II. Comparing levels and trends as households age

A. Levels and trends as households age in tax data

We first examine how the pre-tax income distribution evolves as households age in the tax data. Figure 1A plots baseline estimates for the 10th, 25th, 50th, 75th, and 90th pre-tax income percentiles[19] at each age for a selection of birth cohorts. Figure 1A also displays baseline estimates for the fraction of households receiving pre-tax income above $500 (in 2010 dollars). Total income steadily declines with age throughout most of the distribution. At the median, total income drops from around $45,000 at age 55 to around $25,000 at age 80; at the 25th percentile, total income drops from around $22,000 at age 55 to around $14,000 at age 80. The 10th percentile does not decline monotonically, jumping upwards at the Social Security claiming age around 62-64 before plateauing and slowly declining at later ages. The fraction of households receiving at least $500 in pre-tax income exhibits this pattern as well.

Our pre-tax total income measure does not measure the levels of after-tax liquid income available for consumption. Tax liability may fall as households age, attenuating the decline in total income. Figure 1B investigates this possibility by plotting the distribution of after-tax income. We continue to see a pattern of declining income with age.

To investigate changes in the distribution of income as households age more systematically, we divide our sample into two groups: (i) households in younger birth cohorts (1943-1949) whom we observe at ages 58-68, and (ii) households in older birth cohorts (1933-1939) whom we observe at ages 68-78. We then estimate the proportional change in each percentile of the pre-tax and after-tax income distributions from age 58 to age 68 for the younger cohort, and

---

[19] To maintain confidentiality, all "percentiles" in the tax data are constructed by averaging the 10 observations nearest to the exact percentile cuts.

from age 68 to age 78 for the older cohort. Table 1A reports pre-tax income results, along with the across-birth-cohort average of the proportional change at each percentile. From age 58 to age 68, the 25th-90th total income percentiles decline by about 12-13%. The 10th percentile *increases* by about 30%. From ages 68-78, the proportional decline in total income is larger for the lower percentiles than for the upper percentiles. The 90th percentile decreases by about 9%, compared with a 14.5% decrease at the median, a 16.6% decrease at the 25th percentile, and a 35.9% decrease at the 10th percentile. Table 1B reports analogous results for the after-tax income distribution. From 58-68, proportional declines are about half as large compared with the pre-tax income distribution for the 25th-90th percentiles. From ages 68-78, proportional declines are slightly smaller, but similar at and below the median.

B. Comparing levels in the tax data and survey data

Next, we compare these results with estimates from the survey data. Since the survey data typically lack information on capital gains and tax liability, we compare total pre-tax income in each survey with pre-tax income excluding capital gains in the tax data. (For brevity, we will henceforth refer to pre-tax income excluding capital gains as simply "pre-tax income.") To get a preliminary sense for estimates in each source, Figure 2 plots the pre-tax income distribution by age, as estimated by the tax data and survey data, for the 1941 birth cohort. Several patterns emerge. First, a higher fraction of households in the survey data have pre-tax income exceeding $500 than in the tax data. Levels of receipt in the SIPP and the CPS are closer to the tax data than HRS estimates. This partly accounts for the higher estimates of 10th percentile income in the survey data than in the tax data. Second, from the 25th- 90th percentiles, the SIPP and the CPS tend to underestimate levels of total income compared with the tax data, especially at older ages. Third, at the 25th-90th percentiles, the HRS matches the tax data relatively closely at older ages. For younger ages at the 75th and 90th percentile, the HRS estimates are substantially higher than in the tax data.

Table 2 compares the survey data to the tax data more comprehensively. For each birth cohort and age, we compute the percent difference between percentiles in the tax data and the survey data, and average these percentages across birth cohorts. We also compute the relative difference in the probability of receiving at least $500 of pre-tax income across datasets. Table 2A focuses on measurements at ages 58, 63, and 68 among younger birth cohorts (1943-1949). For context, we also report the across-cohort average values from each survey. Bracketed values are

the minimum and maximum level or percent difference for individual birth cohorts within the 1943-1949 range.

All surveys tend to overestimate income at the 10th percentile relative to the tax data; by age 68, reported values are on average 25.0%, 15.6%, and 20.1% higher in the survey data than in the tax data in the HRS, the SIPP, and the CPS, respectively. At higher percentiles, the HRS tends to overestimate income, while the SIPP and CPS tend to underestimate income.

Comparing the accuracy of each survey at ages 63 and 68, the HRS tends to align most closely with the tax data, except at the 10th percentile, where the SIPP performs better. However, at age 58, the HRS often significantly overestimates pre-tax income relative to the tax data—by an average of 16.3% at the 25th percentile, 15.9% at the 75th percentile, and 17.5% at the 90th percentile. At such points in the distribution, the HRS estimates tend to be less accurate than those in the SIPP and the CPS. For all ages, the SIPP outperforms the CPS at the 25th percentile, but the CPS does better at and above the median.

Table 2B produces similar comparisons of survey and tax data estimates of the pre-tax income distribution at ages 68, 73, and 78, averaging across the group of older birth cohorts (1933-1939). As at younger ages, the survey data report much higher estimates for the 10th percentile than the tax data. Above the 10th percentile, HRS estimates are closest to the tax data, and are often quite accurate in absolute terms. At the 25th percentile, the average percent difference between the tax data and the HRS estimates is 1.5%, -2.5%, and 0.4% at age 68, 73, and 78, respectively; at the median, the average percent difference is -3.8, -4.5%, and -10.7% at age 68, 73, and 78, respectively. Both the SIPP and the CPS tend to underestimate levels of total income. However, the SIPP is uniformly better aligned with the tax data than the CPS. For example, at age 73, the average percent difference between the SIPP and the tax data is -8.9% compared with -13.1% in the CPS; at the median, this difference is -14.2% in the SIPP and -25.1% in the CPS.

Overall, surveys overestimate income at the far left tail. The SIPP and CPS underestimate income for the remainder of the distribution. HRS has a tendency to underestimate income in the remainder of the distribution for older birth cohorts, but a tendency to overestimate income in the remainder of the distribution for younger birth cohorts. Across all birth cohorts, HRS has a slight tendency to underestimate median income.

C. Comparing trends as households age in the tax data and survey data

How do these patterns of under- and over-estimation in survey sources affect survey estimates of declines in income as households age? To investigate, Table 3 compares changes in each percentile of the pre-tax income distribution, as well as changes in the probability of receiving at least $500 of pre-tax income, across different ages for fixed birth cohorts. Panel I reports changes from age 58 to age 68, averaging across the group of younger birth cohorts (1943-1949). Panel II reports changes from age 68 to age 78, averaging across the group of older birth cohorts (1933-1939).

In Panel I, the average proportional declines in the 25th-90th percentiles from age 58 to 68 tend to be larger in the survey data than in the tax data. For instance, the tax data median declines by an average of 11.7%, compared with 24.4% in the HRS, 16.8% in the SIPP, and 29.0% in the CPS. The proportional declines in the SIPP align most closely with the tax data, followed by estimates from the HRS, then the CPS. Panel II shows that survey data estimates of proportional declines from age 68 to age 78 tend to be less uniformly biased upwards—sometimes overestimating declines, and sometimes underestimating declines. Survey estimates at these ages more closely approximate the 12.8% average decline in the tax data median: the HRS estimates a 18.1% decline, the SIPP estimates a 10.2% decline, and the CPS estimates a 16.1% decline.

D. Comparing income in individual categories in the tax data and survey data

The results from Panel I of Table 3 suggest that the survey data often overestimate declines in income during the initial transition from working life to retirement. To produce this pattern, surveys must have either higher levels of working life income than the tax data, lower levels of retirement income than the tax data, or a combination of the two.

We explore these possibilities by measuring distributions separately for each income category in the tax data and survey data. Table 4 reports results. Each row shows estimates for the 10th-90th percentiles of the income distribution in a given category at age 58, 68, or 78, along with the probability of this category of income exceeding $500, averaged over either the younger birth cohorts in Panel I (1933-1939) or the older birth cohorts in Panel II (1943-1949). For each survey source, we additionally report the average percent difference between tax data estimates and survey estimates.[20]

---

[20] In calculating this average, we omit age-cohort measurements with tax data estimates of zero, for which percent differences are not defined.

We first examine sources of income that are more important while households are still working, namely wage income and "other" income. Other income includes income from self-employment and sole-proprietorships, and hence represents a source of labor income. We find that all surveys have a tendency to overestimate wage income, and the HRS overestimates other income, while the SIPP and the CPS tend to underestimate other income. Overall, the SIPP is more closely aligned with the tax data than the HRS or the CPS within these income categories.

Table 4A reports wage income results. Panel I shows that in the tax data, the fraction of households with wage income exceeding $500 falls substantially, from an average of 77% at age 58 to an average of 39% at age 68. All survey sources accurately capture these extensive margin levels and trends. The HRS and CPS overestimate wage income percentiles across the distribution at both age 58 and age 68. At age 58, HRS 50th-90th percentile estimates are on average between 10.6% and 15.0% higher than tax data estimates, and corresponding CPS estimates are on average between 14.2% and 21.8% higher than tax data estimates. The SIPP is comparatively more aligned with the tax data; while 25th and 50th percentile estimates are too high at age 58, estimates for the 75th and 90th percentiles are quite accurate both at age 58 and age 68. In Panel II, we see that the fraction of households receiving wage income in the tax data continues to fall to an average of 13% by age 78, a trend accurately replicated in the survey data. The HRS and CPS continue to overestimate the 75th and 90th wage income percentiles at age 68 and the 90th income percentile at age 78, although the HRS is comparatively more accurate. At both age 68 and age 78, the SIPP overestimates wage income percentiles as well, with comparable accuracy to the HRS. The fact that extensive margin estimates are largely accurate, but percentile estimates are often too high, suggests that survey respondents accurately report whether they receive any wage income, but tend to overestimate the amount they receive.

Table 4B reports other-income results. The tax data show that receipt of other income is relatively uncommon—29% on average at age 58, which declines to 15% by age 78. The survey data also exhibit declining levels of receipt, although the HRS estimates are too high and estimates from the SIPP and CPS are too low. Turning to percentiles, HRS estimates at the 75th and 90th percentiles are much higher on average than estimates in the tax data. By contrast, estimates from the SIPP and CPS are much lower than estimates in the tax data. These differences are economically significant. For instance, in Panel I, the HRS estimate of the 90th percentile at age

58 is on average $32,326, compared with an average estimate of $7,981 in the SIPP, $10,352 in the CPS, and $13,260 in the tax data.

HRS estimates of self-employment income may be more accurate than estimates in the tax data and other surveys. First, there is evidence that households substantially underreport self-employment income to the IRS (Feldman and Slemrod 2007). Second, the HRS questionnaire design makes it more likely to capture self-employment income than other surveys. After asking generally about whether a household earned income from work in the past year, the HRS immediately asks specifically about earnings from self-employment. By contrast, the CPS asks generally about income from "own businesses," and the SIPP asks respondents to classify already-reported earnings as coming from work for an employer or from self-employment.

Next, we study sources of income that are relatively more important in retirement, such as pension income, Social Security income, and capital income. All three surveys typically underestimate income from these sources compared with estimates from the tax data.

Table 4C compares the distribution of pension income in the tax data and survey sources. Unsurprisingly, pension income becomes much more important as households age. Panel I shows that the fraction of households with pension income over $500 in the tax data grows from an average of 33% at 58 to an average of 62% by 68. Pension income at the 75th percentile grows from an average of $5,837 at 58 to an average of $21,672 at 68; at the 90th percentile, the average increases from $26,745 at 58 to $42,708 at 68. All survey data sources underestimate both pension income percentiles and probabilities of receipt. The HRS and the SIPP underestimate pension income by a similar amount, while the CPS performs worse. For all surveys, estimates at age 68 are closer to the tax data than estimates at age 58.

Panel II exhibits the continued growth in pension income from age 68 to age 78, albeit less dramatic than the increase from age 58 to age 68. By age 78, an average of 70% of households have pension income over $500, and the 50th, 75th, and 90th percentiles of the distribution are on average $7,462, $22,124, and $42,567, respectively. Patterns of underreporting are similar to those in Panel I. The HRS and the SIPP suffer from similar levels of underreporting, and the CPS performs worse. Within each survey source, estimates at age 68 tend to be closer to the tax data than estimates at age 78. This suggests that survey reports of pension income become less accurate at older ages, in contrast to results from Panel I, where it appeared that estimates improved with age.

16

Table 4D reports results for Social Security income. In Panel I, the tax data show that receipt of Social Security income goes from uncommon at age 58, with an average of 19% of households receiving income over $500, to near-universal by age 68, with 91% of households receiving income. Focusing on results at age 68, once Social Security becomes a meaningful source of income for most households, the SIPP produces estimates most closely aligned with the tax data. Results on the extensive margin are on average nearly identical to the tax data, as are estimates of the 25th-90th percentiles. The HRS follows close behind the SIPP in terms of accuracy. Average levels of receipt are about 4 percentage points higher in the HRS than in the tax data; estimates of the 10th and 25th percentiles are higher on average in the HRS than in the tax data; and HRS estimates for upper percentiles are about 5-6% lower on average than tax data estimates. The CPS performs the worst, underestimating the extensive margin by an average of 6.5%, and underestimating the 25th-90th percentiles by between 5.3% and 16.1% on average.

In Panel II, the relative accuracy of the surveys is more ambiguous. As in Panel I, the HRS overestimates the extensive margin relative to the tax data and on average overestimates the 10th and 25th percentiles at both ages. Estimates at the 50th-90th percentiles are lower than the tax data but closer than equivalent estimates in Panel I. In the SIPP, age 68 estimates are between 6.5% and 10.0% lower on average than in the tax data at the 25th-90th percentiles, considerably less accurate than in Panel I. However, at age 78, estimates are uniformly *higher* than in the tax data by between 1.5% and 10.0% on average. CPS estimates are typically more accurate at both age 68 and age 78 than in Panel I. In fact, CPS estimates at age 78 are, on average, quite aligned with the tax data, especially above the 10th percentile.

Lastly, Table 4E compares the capital income distribution across sources. The tax data show that while around 40-50% of households earn capital income, the amounts received are rarely substantial. Even the average 75th capital income percentile is often below $1,000. In Panel I, the HRS and the SIPP underestimate and the CPS overestimates percentiles and levels of receipt. Average percent differences between the tax and survey estimates are typically large, but since income levels are low, dollar differences are rarely economically significant. In Panel II, all survey sources tend to underestimate capital income. The CPS and HRS tend to perform best, followed by the SIPP. At the 90th percentile, income levels tend to be larger in the tax data—$11,543 at age 68 and $8,775 at age 78 on average—such that large average percent differences, especially in the SIPP, are often economically significant.

E. Comparing IRA balances in the tax data and survey data

Consumption need not track income in retirement, as households can spend down accumulated assets. Therefore, accurately measuring asset holdings during late life is important to understanding overall financial wellbeing. The tax data lack comprehensive asset measures, but do have information on IRA balances. We therefore compare the age- and cohort-specific distribution of IRA balances in the tax data with estimates in the survey data to assess the accuracy of survey-derived asset measures.

Figure 3 shows the age-specific distribution of IRA balances in the tax data, in addition to the fraction of households with positive IRA balances, for select birth cohorts. At most ages, fewer than 50% of households have positive IRA balances; percentiles at and below the median are mostly zero and are omitted from the figure. Several patterns emerge. First, the 75th and 90th percentiles of the IRA balance distribution, and to a lesser extent the fraction of households with positive balances, rises from the mid-50s to age 70. Focusing on the 1947 birth cohort from age 54 to age 70, the 75th percentile grows from around $22,000 to $125,000, the 90th percentile grows from around $93,000 to $392,000, and the fraction with positive balances rises from 44% to 51%. This increase reflects continued IRA contributions at later ages, asset growth within accounts, and the rolling over of employer-sponsored accounts into IRAs. Second, starting when required distributions begin around age 70, percentiles tend to stay flat with age, while the fraction of households with positive balances begins to decline with age. This suggests that while some households completely exhaust their IRA assets, those with positive balances tend not to spend them down. Finally, amounts held in IRAs at fixed ages has grown in more recent birth cohorts. For instance, at age 70, 75th percentile holdings have increased from about $46,000 in the 1933 birth cohort to around $125,000 in the 1947 birth cohort. The growth in the fraction of households with positive IRA balances is more modest, increasing at age 70 from 44% in the 1933 birth cohort to 51% in the 1947 birth cohort.

We next examine how well the survey data measures IRA holdings. Since the CPS does not ask about IRA balances, we restrict our analysis to the HRS and the SIPP. Table 5 reports results using a similar structure to Table 2. For each birth cohort and each age, we estimate the percent difference between estimates in the tax and survey data. Panel I reports averages across younger birth cohorts (1943-1949) at ages 58, 63, and 68, where brackets below each estimate

indicate minimum and maximum estimates within individual cohorts. Panel II reports analogous estimates constructed by averaging across older birth cohorts (1933-1939) at ages 68, 73, and 78.

The results show that the HRS is quite well-aligned with the tax data, especially at older ages. In Panel I, the HRS underestimates the fraction of households with positive holdings by an average of at most 5.1%. Estimates of the 75th and 90th percentiles are typically too high, sometimes by large amounts at age 58. In Panel II, HRS estimates are more accurate. On the extensive margin, the average difference between tax and survey data estimates ranges from -2.9% to 0.7%; at the 75th percentile, the average difference ranges from -2.6% to 2.0%; and at the 90th percentile, the average difference ranges from -13.7% to 0.3%. The SIPP performs much worse by comparison. In both panels, the fraction of households with positive IRA balances is underestimated by an average of around 20-25%; both the 75th and 90th percentiles are underestimated by an average of between around 40-70%.

III. Comparing trends in income across birth cohorts

A. Trends in total and after-tax income across birth cohorts using tax data

We start by using the tax data to estimate trends in pre-tax income (including capital gains) and after-tax income across birth cohorts. Specifically, we measure changes at fixed ages from the 1933 cohort to the 1943 cohort, measured at ages 68-74, and from the 1944 cohort to the 1950 cohort, measured at ages 58-67.

Table 6A reports levels and proportional changes at various ages and percentiles across the 1944 to 1950 birth cohorts and the 1933 to 1943 birth cohorts for the pre-tax income distribution. In Panel I, we see that pre-tax income at ages 58-67 has stagnated at the median while falling substantially at the 10th and 25th percentiles, even as it rose at the 75th and 90th percentiles. At the older ages shown in Panel II, we see consistent growth across the entire distribution. From the 1933-1943 birth cohorts, the median increases by an average of 14.1%. Increases in percentiles below the median tend to be slightly lower, while increases in percentiles above the median tend to be slightly higher. Table 6B reports equivalent results for the distribution of after-tax income. Average proportional changes in the after-tax income distribution across cohorts are qualitatively similar but tend to be slightly lower in magnitude. The notable exception is the change in the 10th percentile from the 1944 to 1950 birth cohorts, where the after-tax decline of 20.8% is larger than the before-tax decline of 19.6%.

B. Comparing trends across birth cohorts in the tax data and survey data

We now study whether the survey data accurately capture tax data trends in the pre-tax income (excluding capital gains) distribution across birth cohorts. In each survey source, we compute the proportional change in each percentile of the pre-tax income distribution across birth cohorts at fixed ages. Table 7 shows the results. Panel I reports changes from the 1944 cohort to the 1950 birth cohort, averaging across younger ages (58-67). Panel II reports changes from the 1933 cohort to the 1943 cohort, averaging across older ages (68-74). Average changes in the tax data distribution of pre-tax income excluding capital gains are reported for comparison.

In Panel I, we see that the CPS and SIPP do a good job capturing the trends from the 1944 to 1950 birth cohorts. On the other hand, the HRS overestimates growth at the 10th and 25th percentiles, and underestimates growth at the 50th percentile and above. Panel II shows that for the older cohorts, at the 50th and 75th percentiles, all the survey sources tend to overestimate growth in the distribution of total income. The average proportional increase in the median from the 1933 cohort to the 1943 cohort is 16.3% in the tax data, compared to 21.7% in the HRS, 23.5% in the SIPP, and 23.6% in the CPS. At the 75th percentile, the average proportional increase is 19.9% in the tax data, 23.3% in the HRS, 28.9% in the SIPP, and 27.1% in the CPS. Below the median, results are different for each survey. At the 25th percentile, the HRS underestimates the average increase of 14.6% reported in the tax data, while the increase is overestimated in the SIPP and appropriately estimated in the CPS. All three survey sources underestimate the average increase of 17.5% in the tax data at the 10th percentile.

IV. Stagnating non-Social Security income at and below the median

The results on older ages from the previous section suggest improving retirement preparedness across the income distribution in more recent birth cohorts. In this section, we examine the extent to which this is due to income growth *outside* of the Social Security system.

A. Changes in the distribution of non-Social Security income across birth cohorts

In Table 8, we estimate in the tax data levels and proportional changes in non-Social Security income across birth cohorts for fixed ages. Estimates for the 10th percentile are omitted,

as non-Social Security income levels there are almost universally zero. We focus here on results in Panel II, which shows changes from the 1933 birth cohort to the 1943 birth cohort.

At and below the median, the proportional changes in non-Social Security income from the 1933 cohort to the 1943 cohort are much lower than the proportional changes in total pre-tax income. On average, the proportional change at the median is 9.4%, compared with the 14.1% for total pre-tax income shown in Table 6A. At the 25th percentile, non-Social Security income *falls* by 16.5%, compared with an increase of 11.0% for total pre-tax income. The proportional changes in non-Social Security income at the 75th and 90th percentiles are similar to the proportional changes in total pre-tax income. This suggests that the lower income percentiles are becoming more reliant on Social Security income in a way that the upper income percentiles are not.

Table 9 compares these results to survey data trends in the distribution of non-Social Security income, adopting a similar structure to Table 7. In Panel II, the survey sources tend to overestimate growth in non-Social Security income at the median and above. The surveys qualitatively capture the decline in the 25th percentile across cohorts, although the HRS and the SIPP overestimate and the CPS underestimates the magnitude of this decline on average.

B. Changes in the fraction of households with no income and no IRA balances across birth cohorts

The previous results suggest that reliance on Social Security has grown in more recent birth cohorts for low-income households. However, it is possible that households with limited non-Social Security income have substantial other savings available for consumption. While we cannot quantify total available savings in the tax data, we can identify households with no financial assets with relative precision. We directly observe IRA balances, and can detect substantial holdings in remaining accounts through income from interest, dividend, capital gains, and distributions from tax-exempt retirement accounts.[21] We classify households in the tax data as entirely reliant on Social Security if non-Social Security income does not exceed $500 in 2010 dollars and IRA balances are zero. For these households, income is presumably nearly identical to consumption.

Figure 4A plots the fraction of households in the tax data with no non-Social Security income and no IRA balances for each cohort at selected ages. For each age, the fraction is flat to increasing over time. For example, at age 72, the fraction increases from 18.9% in the 1933 birth

---

[21] By age 72, all households with positive balances in tax-advantaged employer-sponsored savings accounts such as 401(k)s must take an annual distribution.

cohort to 20.5% in the 1945 birth cohort. Reliance on Social Security also substantially increases with age, from around 13% at age 64 to around 25% at age 80.

Figure 4B focuses on age 72 households, and compares trends in Social Security reliance in the tax data and survey data. Since the HRS and the SIPP have information on assets, we are able to construct estimates of the fraction of households with no non-Social Security income and no IRA balances that are comparable to the tax data. While estimates from the HRS and the SIPP are noisier than the tax data, they show similar levels and trends. We also compute an additional measure that instead excludes households in the HRS or SIPP with total financial assets (including IRA balances) exceeding $3,500.[22] This lowers the fraction of households entirely reliant on Social Security by about 1-3 percentage points in the HRS and about 5 percentage points in the SIPP; however, overall trends across cohorts are similar even in the modified series. The CPS lacks data on IRA balances. For comparability, we compute a version of our Social Security reliance measure in the tax data that does not screen out households with positive IRA balances, and compare this series to the fraction of households in the CPS with non-Social Security income below $500. The CPS significantly overestimates the fraction of households that meet even this less restrictive criterion by about 3-5 percentage points.

Lastly, we examine differences across geography in Social Security reliance. Figure 5 displays the fraction of households with no non-Social Security income and no IRA balances in each US state, estimated from the tax data. Social Security reliance is highest in the Deep South states like Mississippi and Louisiana, and lowest in Upper Midwest states like Wisconsin and Minnesota. Table 10 ranks states by the fraction of households with no non-Social Security income and no IRA balances at age 72 in the 1933 birth cohort, and computes the changes in this fraction from the 1933 to 1943 birth cohorts. There is considerable geographic heterogeneity in Social Security reliance, ranging from to 11.1% at the 10th percentile state (Kansas) to 22.1% at the 90th percentile state (Georgia). Different states have also experienced considerably different trends in Social Security reliance from the 1933 to 1943 birth cohorts, from a 10th percentile change of -1.4 percentage points on a base of 19.0% (North Carolina) to a 90th percentile change of +1.7 percentage points on a base of 11.1% (Kansas). Regressing the change in Social Security reliance

_____

[22] We choose the $3,500 threshold because the average monthly Social Security benefit is about $2,500 for married couples, which households who otherwise have no assets might have in their checking account when interviewed, to which we arbitrarily add $1,000.

at age 72 from the 1933 to 1943 birth cohort on its level in the 1933 birth cohort yields a coefficient of -0.15 ($p < 0.001$), indicating geographic convergence over time.

V. Conclusion

The tax data show considerable declines in income across the distribution as households age. However, relying on survey data alone will overstate this decline at the median during the initial transition to retirement, leading to overly dire assessments of middle-income household retirement preparedness. By contrast, while the tax data show income growth across birth cohorts, growth at older ages across cohorts is overstated in survey sources. Leaning on survey data will therefore present an excessively optimistic picture of cohort trends in retirement preparedness. All sources indicate declining income outside of the Social Security system across birth cohorts for low-income households, as well as an increasing share of households who are entirely reliant on Social Security to finance their consumption.

References

Bee, Adam and Joshua Mitchell, 2017. "Do Older Americans Have More Income Than We Think?" SESHD Working Paper #2017-39.

Chen, Anqi, Alicia H. Munnell, and Geoffrey T. Sanzenbacher, 2018. "How Much Income Do Retirees Actually Have? Evaluating the Evidence from Five National Datasets." CRR Working Paper #2018-14.

Chernozhukov, Victor, Iván Fernández-Val, and Alfred Galichon, 2010. "Quantile and Probability Curves without Crossing." *Econometrica* 78(3): 1093-125. Available at: www.jstor.org/stable/40664520.

Dushi, Irena, Howard M. Iams, and Brad Trenkamp, 2017. "The Importance of Social Security Benefits to the Income of the Aged Population." *Social Security Bulletin* 77(2).

Feldman, Naomi E. and Joel Slemrod, 2007. "Estimating tax noncompliance with evidence from unaudited tax returns." *The Economic Journal*, 117(518), 327-352. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0297.2007.02020.x

Guvenen, Fatih, Greg Kaplan, Jae Song, and Justin Weidner, 2017. "Lifetime Incomes in the United States over Six Decades." National Bureau of Economic Research Working Paper #23371. Available at: https://www.nber.org/papers/w23371.pdf

US Department of Defense, 2018. "Population Representation in the Military Services: Appendix B tables." Available at: https://www.cna.org/pop-rep/2018/appendixb/appendixb.pdf.

Table 1A. Changes in the tax data distribution of pre-tax income across ages

Panel I. Changes across ages for younger birth cohorts (1943-1949)

	10th percentile			25th percentile			50th percentile			75th percentile			90th percentile		
Cohort	Age 58	Age 68	Change	Age 58	Age 68	Change	Age 58	Age 68	Change	Age 58	Age 68	Change	Age 58	Age 68	Change
1943	6,869	7,451	8.5%	21,301	18,554	-12.9%	41,874	35,724	-14.7%	69,406	58,881	-15.2%	110,483	93,612	-15.3%
1944	5,497	6,080	10.6%	20,280	17,219	-15.1%	41,092	35,067	-14.7%	68,388	59,100	-13.6%	107,213	95,491	-10.9%
1945	2,698	5,828	116.0%	18,655	17,153	-8.1%	40,212	35,755	-11.1%	68,096	60,988	-10.4%	107,876	98,335	-8.8%
1946	6,138	7,517	22.5%	21,221	18,738	-11.7%	42,917	37,724	-12.1%	71,704	63,738	-11.1%	113,648	102,735	-9.6%
1947	6,830	8,183	19.8%	21,948	19,700	-10.2%	44,122	38,911	-11.8%	73,686	65,411	-11.2%	117,966	105,338	-10.7%
1948	6,871	7,836	14.0%	21,828	18,780	-14.0%	44,159	37,756	-14.5%	74,171	63,648	-14.2%	119,385	102,462	-14.2%
1949	5,906	7,447	26.1%	20,881	18,300	-12.4%	43,814	37,777	-13.8%	74,348	64,569	-13.2%	120,329	104,149	-13.4%
Avg			**31.1%**			**-12.0%**			**-13.2%**			**-12.7%**			**-11.9%**

Panel II. Changes across ages for older birth cohorts (1933-1939)

Cohort	Age 68	Age 78	Change	Age 68	Age 78	Change	Age 68	Age 78	Change	Age 68	Age 78	Change	Age 68	Age 78	Change
1933	6,818	4,591	-32.7%	16,861	14,466	-14.2%	31,950	26,811	-16.1%	51,134	44,061	-13.8%	82,210	71,561	-13.0%
1934	6,577	3,775	-42.6%	16,695	13,766	-17.5%	31,741	27,020	-14.9%	50,305	45,049	-10.4%	79,298	74,021	-6.7%
1935	5,891	3,673	-37.6%	15,885	13,829	-12.9%	31,241	27,707	-11.3%	50,122	46,399	-7.4%	80,259	76,700	-4.4%
1936	6,767	4,480	-33.8%	16,898	14,261	-15.6%	32,584	28,400	-12.8%	52,802	48,229	-8.7%	85,458	80,660	-5.6%
1937	6,909	4,475	-35.2%	17,316	14,645	-15.4%	33,643	29,368	-12.7%	55,049	49,477	-10.1%	90,232	82,173	-8.9%
1938	7,679	5,859	-23.7%	18,429	15,232	-17.3%	35,180	29,575	-15.9%	58,125	50,107	-13.8%	96,608	82,943	-14.1%
1939	7,589	4,126	-45.6%	18,678	14,360	-23.1%	35,928	29,566	-17.7%	59,592	50,672	-15.0%	99,767	86,435	-13.4%
Avg			**-35.9%**			**-16.6%**			**-14.5%**			**-11.3%**			**-9.4%**

Source: Authors' calculations. Pre-tax income is the sum of (i) wage income; (ii) pension income; (iii) Social Security income; (iv) capital income; and (v) other income. All values are baseline estimates deflated to 2010 dollars obtained by positioning the institutionalized population in the total income distribution using HRS data.

Table 1B. Changes in the tax data distribution of after-tax income across ages

Panel I. Changes across ages for younger birth cohorts (1943-1949)

	10th percentile			25th percentile			50th percentile			75th percentile			90th percentile		
Cohort	Age 58	Age 68	Change	Age 58	Age 68	Change	Age 58	Age 68	Change	Age 58	Age 68	Change	Age 58	Age 68	Change
1943	6,168	7,205	16.8%	19,154	18,215	-4.9%	35,541	34,414	-3.2%	56,037	53,025	-5.4%	86,557	80,338	-7.2%
1944	4,758	5,762	21.1%	18,610	16,815	-9.6%	35,379	33,813	-4.4%	55,965	53,124	-5.1%	84,354	81,561	-3.3%
1945	1,999	5,531	176.7%	17,274	16,828	-2.6%	34,849	34,281	-1.6%	56,516	54,434	-3.7%	86,766	83,562	-3.7%
1946	5,541	7,302	31.8%	19,521	18,363	-5.9%	37,023	35,949	-2.9%	59,347	56,703	-4.5%	91,231	87,214	-4.4%
1947	6,253	8,042	28.6%	20,143	19,308	-4.1%	37,968	36,956	-2.7%	60,907	57,999	-4.8%	94,808	89,064	-6.1%
1948	6,444	7,675	19.1%	20,112	18,438	-8.3%	38,042	35,923	-5.6%	61,409	56,575	-7.9%	96,324	86,674	-10.0%
1949	5,442	7,299	34.1%	19,362	17,940	-7.3%	37,772	35,843	-5.1%	61,589	57,147	-7.2%	96,769	87,979	-9.1%
Avg			**46.9%**			**-6.1%**			**-3.6%**			**-5.5%**			**-6.3%**

Panel II. Changes across ages for older birth cohorts (1933-1939)

Cohort	Age 68	Age 78	Change	Age 68	Age 78	Change	Age 68	Age 78	Change	Age 68	Age 78	Change	Age 68	Age 78	Change
1933	6,694	4,185	-37.5%	16,568	14,294	-13.7%	30,700	26,487	-13.7%	46,185	41,678	-9.8%	69,561	64,411	-7.4%
1934	6,368	3,562	-44.1%	16,342	13,645	-16.5%	30,620	26,660	-12.9%	45,994	42,571	-7.4%	67,747	66,456	-1.9%
1935	5,745	3,493	-39.2%	15,607	13,717	-12.1%	30,336	27,349	-9.8%	46,099	43,580	-5.5%	69,737	68,747	-1.4%
1936	6,638	4,323	-34.9%	16,614	14,126	-15.0%	31,565	27,974	-11.4%	48,086	45,093	-6.2%	73,884	71,817	-2.8%
1937	6,780	4,360	-35.7%	17,057	14,551	-14.7%	32,560	28,978	-11.0%	49,895	46,149	-7.5%	77,758	73,196	-5.9%
1938	7,568	5,717	-24.5%	18,159	15,118	-16.7%	33,905	29,162	-14.0%	52,421	46,669	-11.0%	83,124	73,665	-11.4%
1939	7,449	4,015	-46.1%	18,361	14,251	-22.4%	34,569	29,122	-15.8%	53,597	47,116	-12.1%	85,868	76,602	-10.8%
Avg			**-37.4%**			**-15.9%**			**-12.7%**			**-8.5%**			**-5.9%**

Source: Authors' calculations. After-tax income is pre-tax income, minus reported federal tax liability and estimated FICA taxes, plus any earned income tax credits. All values are baseline estimates deflated to 2010 dollars obtained by positioning the institutionalized population in the total income distribution using HRS data.

Table 2A. Differences in pre-tax income excluding capital gains between tax data and survey data, 1943-1949 birth cohorts

	Age	HRS		SIPP		CPS	
		Level	Avg diff vs. tax	Level	Avg diff vs. tax	Level	Avg diff vs. tax
p10	58	8,988	37.9%	8,267	21.3%	8,872	31.1%
		[7,881 to 10,040]	[0.8% to 95.0%]	[6,986 to 10,386]	[-6.5% to 59.7%]	[7,551 to 10,623]	[0.2% to 86.1%]
	63	9,509	9.7%	8,818	2.0%	8,401	-2.2%
		[8,192 to 10,582]	[-1.8% to 17.2%]	[8,188 to 9,629]	[-6.1% to 14.2%]	[7,388 to 9,311]	[-12.4% to 7.3%]
	68	10,592	25.0%	9,796	15.6%	10,087	20.1%
		[9,872 to 11,258]	[7.8% to 51.8%]	[7,762 to 11,433]	[-1.2% to 32.7%]	[8,893 to 11,398]	[4.6% to 49.2%]
p25	58	24,872	16.3%	19,900	-7.7%	23,882	10.6%
		[20,235 to 29,135]	[-8.2% to 45.5%]	[17,771 to 22,069]	[-20.5% to 10.2%]	[22,086 to 26,030]	[2.7% to 22.2%]
	63	20,385	-4.2%	19,231	-9.5%	18,674	-11.6%
		[18,616 to 22,705]	[-10.2% to 7.5%]	[18,146 to 21,145]	[-14.1% to -5.4%]	[17,311 to 20,000]	[-19.5% to -1.8%]
	68	19,008	-4.7%	18,862	-5.0%	17,344	-12.0%
		[17,540 to 19,771]	[-13.9% to 5.4%]	[16,099 to 24,052]	[-14.6% to 15.0%]	[16,074 to 18,585]	[-18.7% to -3.9%]
p50	58	47,198	9.9%	39,071	-9.1%	45,610	6.1%
		[43,550 to 49,209]	[2.9% to 15.4%]	[37,387 to 42,575]	[-15.5% to -2.1%]	[42,866 to 47,870]	[0.6% to 15.3%]
	63	41,520	1.8%	36,302	-10.9%	36,897	-9.2%
		[40,405 to 42,724]	[0.1% to 5.2%]	[34,107 to 39,329]	[-15.5% to -5.0%]	[35,650 to 38,404]	[-11.6% to -4.6%]
	68	35,145	-7.4%	32,607	-14.1%	32,378	-14.7%
		[33,570 to 36,938]	[-11.9% to -0.1%]	[29,886 to 36,049]	[-19.1% to -8.0%]	[31,042 to 35,374]	[-19.5% to -7.4%]
p75	58	82,746	15.9%	64,403	-9.8%	76,426	7.0%
		[80,442 to 87,237]	[11.8% to 18.5%]	[61,716 to 68,957]	[-13.4% to -4.2%]	[73,558 to 78,574]	[2.3% to 11.8%]
	63	71,228	5.5%	59,832	-11.2%	63,605	-5.4%
		[67,231 to 76,180]	[-0.2% to 12.0%]	[52,371 to 64,328]	[-20.9% to -4.1%]	[61,281 to 66,480]	[-7.8% to -0.1%]
	68	64,595	3.0%	53,247	-15.2%	57,283	-8.9%
		[63,131 to 66,437]	[1.4% to 3.9%]	[47,925 to 59,611]	[-20.6% to -8.2%]	[54,521 to 61,054]	[-14.4% to -4.7%]
p90	58	132,162	17.5%	96,875	-13.9%	117,815	4.8%
		[131,115 to 133,089]	[12.5% to 21.7%]	[90,843 to 101,797]	[-17.6% to -7.5%]	[111,528 to 122,633]	[-0.4% to 11.5%]
	63	111,969	4.9%	90,612	-15.1%	102,463	-4.0%
		[105,027 to 120,385]	[-1.0% to 12.9%]	[77,999 to 97,887]	[-25.8% to -7.8%]	[96,648 to 108,400]	[-8.4% to 1.0%]
	68	110,751	11.3%	85,548	-14.1%	92,560	-7.2%
		[102,789 to 125,224]	[4.5% to 21.1%]	[74,894 to 100,838]	[-22.4% to -2.5%]	[82,374 to 99,312]	[-14.5% to -1.6%]
has	58	0.96	1.8%	0.96	1.2%	0.95	0.9%
		[0.95 to 0.98]	[0.1% to 2.9%]	[0.94 to 0.97]	[-0.9% to 2.8%]	[0.94 to 0.96]	[-1.0% to 2.4%]
	63	0.98	1.6%	0.97	1.2%	0.95	-1.2%
		[0.97 to 0.99]	[0.2% to 2.9%]	[0.96 to 0.98]	[-0.3% to 1.9%]	[0.94 to 0.96]	[-1.9% to 0.6%]
	68	0.99	2.5%	0.97	1.2%	0.97	0.5%
		[0.98 to 0.99]	[2.0% to 3.3%]	[0.96 to 0.99]	[-0.3% to 3.0%]	[0.95 to 0.99]	[-1.2% to 2.7%]

Source: Authors' calculations. Each row displays the average percentile level in the survey and the average percent difference between the survey percentile estimate and the baseline tax data percentile estimate for the 1943-1949 birth cohorts. The "has" rows indicate the probability of having pre-tax income excluding capital gains that exceeds $500 in 2010 dollars. Brackets report the minimum and maximum estimates for individual cohorts.

Table 2B. Differences in pre-tax income excluding capital gains between tax data and survey data, 1933-1939 birth cohorts

	Age	HRS Level	HRS Avg diff vs. tax	SIPP Level	SIPP Avg diff vs. tax	CPS Level	CPS Avg diff vs. tax
p10	68	10,502	38.7%	9,555	25.7%	9,257	21.9%
		[9,407 to 11,211]	[27.9% to 52.5%]	[8,281 to 10,753]	[-1.9% to 40.3%]	[9,026 to 9,544]	[7.5% to 37.6%]
	73	10,968	42.7%	9,098	20.0%	9,627	27.6%
		[10,695 to 11,233]	[31.7% to 62.5%]	[7,901 to 10,517]	[2.6% to 38.4%]	[8,592 to 11,330]	[17.3% to 46.2%]
	78	10,092	62.0%	8,351	33.3%	8,522	37.2%
		[9,484 to 10,904]	[45.4% to 77.2%]	[7,133 to 10,395]	[1.8% to 72.9%]	[6,910 to 9,931]	[13.7% to 71.0%]
p25	68	18,385	1.5%	16,959	-6.3%	15,696	-13.4%
		[17,472 to 19,146]	[-1.3% to 4.6%]	[15,817 to 17,805]	[-14.0% to 1.1%]	[15,181 to 16,394]	[-17.0% to -9.4%]
	73	17,166	-2.5%	15,874	-8.9%	15,072	-13.1%
		[16,748 to 17,731]	[-5.9% to -0.9%]	[14,607 to 17,661]	[-19.2% to -1.3%]	[14,263 to 16,073]	[-16.8% to -3.0%]
	78	15,517	0.4%	15,185	-1.8%	14,618	-5.4%
		[15,222 to 15,714]	[-0.7% to 1.5%]	[13,827 to 16,538]	[-12.6% to 8.4%]	[13,536 to 15,966]	[-12.6% to 6.5%]
p50	68	32,415	-3.8%	28,351	-15.7%	27,461	-18.4%
		[29,283 to 34,561]	[-9.4% to 0.3%]	[25,593 to 29,520]	[-20.8% to -10.4%]	[26,426 to 30,410]	[-23.8% to -15.2%]
	73	30,568	-4.5%	27,464	-14.2%	23,879	-25.1%
		[29,930 to 31,463]	[-8.0% to -1.6%]	[24,743 to 28,892]	[-20.1% to -11.4%]	[21,642 to 25,140]	[-31.4% to -22.6%]
	78	25,951	-10.7%	25,397	-13.0%	23,013	-21.6%
		[24,037 to 27,605]	[-14.6% to -8.6%]	[23,313 to 28,781]	[-19.1% to -1.5%]	[19,807 to 26,424]	[-29.7% to -8.3%]
p75	68	55,972	4.2%	47,183	-12.3%	48,020	-10.5%
		[52,030 to 58,297]	[-0.5% to 11.6%]	[41,781 to 52,437]	[-18.2% to -7.1%]	[45,888 to 50,284]	[-17.2% to -1.3%]
	73	50,142	-2.8%	42,669	-17.5%	41,533	-19.6%
		[48,876 to 51,778]	[-6.1% to 3.4%]	[38,755 to 45,699]	[-23.0% to -13.0%]	[36,311 to 45,943]	[-27.9% to -12.2%]
	78	42,173	-11.7%	41,257	-14.5%	40,181	-17.2%
		[38,263 to 45,495]	[-16.9% to -8.9%]	[36,977 to 47,583]	[-21.7% to -6.3%]	[35,180 to 47,210]	[-24.1% to -0.1%]
p90	68	95,833	12.2%	71,300	-16.3%	84,706	-0.4%
		[86,776 to 106,011]	[7.5% to 14.8%]	[60,751 to 86,878]	[-26.4% to -1.5%]	[80,353 to 87,262]	[-8.2% to 8.6%]
	73	79,457	-3.1%	61,754	-25.2%	70,613	-14.6%
		[71,270 to 85,538]	[-15.1% to 10.2%]	[55,167 to 64,678]	[-32.0% to -23.0%]	[64,309 to 77,224]	[-23.4% to -8.5%]
	78	71,741	-7.5%	67,435	-13.8%	65,406	-17.2%
		[66,610 to 80,858]	[-11.5% to -0.4%]	[57,873 to 84,034]	[-23.5% to 7.0%]	[53,477 to 75,118]	[-29.3% to -7.1%]
has	68	0.99	3.0%	0.98	1.9%	0.97	1.0%
		[0.98 to 0.99]	[2.5% to 3.6%]	[0.97 to 0.99]	[0.3% to 4.0%]	[0.96 to 0.98]	[-0.3% to 2.6%]
	73	0.99	3.9%	0.98	3.2%	0.97	2.0%
		[0.99 to 0.99]	[3.4% to 4.7%]	[0.98 to 0.99]	[2.6% to 3.9%]	[0.96 to 0.98]	[1.0% to 3.2%]
	78	0.99	5.2%	0.97	2.4%	0.96	1.9%
		[0.99 to 1.00]	[4.1% to 6.0%]	[0.94 to 1.00]	[-1.3% to 6.0%]	[0.94 to 0.98]	[-0.9% to 4.7%]

Source: Authors' calculations. Each row displays the average percentile level in the survey and the average percent difference between the survey percentile estimate and the baseline tax data percentile estimate for the 1933-1939 birth cohorts. The "has" rows indicate the probability of having pre-tax income excluding capital gains that exceeds $500 in 2010 dollars. Brackets report the minimum and maximum estimates for individual cohorts.

Table 3. Comparisons between changes in the pre-tax income excluding capital gains distribution across various ages in the tax data and survey data

Panel I. Average changes from 58-68 for younger birth cohorts (1943-1949)

	Tax data			HRS			SIPP			CPS		
	58	68	Avg. % change	58	68	Avg. % change	58	68	Avg. % change	58	68	Avg. % change
p10	6,948	8,446	23.2%	8,988	10,592	16.7%	8,267	9,796	21.5%	8,872	10,087	15.4%
			[15.3% to 52.5%]			[6.0% to 25.3%]			[-12.8% to 63.6%]			[-1.8% to 50.9%]
p25	21,625	19,726	-8.8%	24,872	19,008	-24.1%	19,900	18,862	-5.1%	23,882	17,344	-27.1%
			[-10.8% to -6.3%]			[-32.1% to -13.3%]			[-27.1% to 19.7%]			[-33.0% to -17.9%]
p50	43,032	37,972	-11.7%	47,198	35,145	-24.4%	39,071	32,607	-16.8%	45,610	32,378	-29.0%
			[-12.8% to -10.4%]			[-27.9% to -22.4%]			[-25.0% to -5.1%]			[-32.4% to -26.1%]
p75	71,456	62,892	-12.0%	82,746	64,595	-20.5%	64,403	53,247	-17.4%	76,426	57,283	-25.0%
			[-13.0% to -10.5%]			[-21.5% to -19.0%]			[-23.9% to -8.3%]			[-29.3% to -19.1%]
p90	112,544	99,635	-11.5%	132,162	110,751	-16.4%	96,875	85,548	-11.8%	117,815	92,560	-21.3%
			[-13.4% to -9.3%]			[-22.1% to -5.5%]			[-24.7% to 4.2%]			[-32.8% to -15.2%]
has	0.94	0.96	2.0%	0.96	0.99	2.6%	0.96	0.97	2.2%	0.95	0.97	1.7%
			[1.5% to 2.8%]			[1.5% to 3.7%]			[0.7% to 3.6%]			[-0.7% to 4.8%]

Panel II. Average changes from 68-78 for older birth cohorts (1933-1949)

	68	78	Avg. % change	68	78	Avg. % change	68	78	Avg. % change	68	78	Avg. % change
p10	7,632	6,251	-17.8%	10,502	10,092	-3.0%	9,555	8,351	-10.3%	9,257	8,522	-7.8%
			[-29.8% to -7.5%]			[-11.4% to 5.1%]			[-26.2% to 10.7%]			[-25.2% to 9.3%]
p25	18,139	15,468	-14.6%	18,385	15,517	-14.3%	16,959	15,185	-10.0%	15,696	14,618	-6.7%
			[-20.8% to -10.9%]			[-18.4% to -10.1%]			[-20.5% to -4.8%]			[-12.7% to 4.6%]
p50	33,667	29,336	-12.8%	32,415	25,951	-18.1%	28,351	25,397	-10.2%	27,461	23,013	-16.1%
			[-16.3% to -10.5%]			[-18.8% to -17.6%]			[-19.2% to -2.5%]			[-25.9% to 0.0%]
p75	53,767	48,534	-9.6%	55,972	42,173	-23.7%	47,183	41,257	-11.6%	48,020	40,181	-16.5%
			[-13.3% to -7.3%]			[-26.5% to -19.8%]			[-20.2% to -8.5%]			[-24.9% to -6.1%]
p90	85,220	78,862	-7.3%	95,833	71,741	-22.5%	71,300	67,435	-5.6%	84,706	65,406	-22.9%
			[-11.3% to -4.4%]			[-26.2% to -18.1%]			[-14.7% to 6.9%]			[-33.4% to -13.1%]
has	0.96	0.94	-1.9%	0.99	0.99	0.2%	0.98	0.97	-1.5%	0.97	0.96	-1.0%
			[-2.4% to -1.1%]			[-0.2% to 0.6%]			[-4.3% to 1.1%]			[-3.4% to 1.0%]

Source: Authors' calculations. Each row displays the average percentile level of pre-tax income excluding capital gains at specified ages, along with the average percent change in these percentile levels between those ages, averaged over either the 1943-1949 birth cohorts or the 1933-1949 birth cohorts. The "has" rows indicate the probability of having pre-tax income excluding capital gains that exceeds $500 in 2010 dollars. Brackets report the minimum and maximum estimates for individual cohorts.

Table 4A. Comparisons between the distribution of wage income in tax data and survey data at select ages

Panel I. Average levels and percent differences for younger birth cohorts (1943-1949)

	Tax data		HRS		SIPP		CPS	
	58	68	58	68	58	68	58	68
p10	0	0	0	0	0	0	0	0
(% diff vs. tax)			-	-	-	-	-	-
p25	1,870	0	945	0	8,035	0	7,246	0
(% diff vs. tax)			5.3%	-	627.9%	-	558.7%	-
p50	27,504	0	30,264	0	29,573	0	33,461	0
(% diff vs. tax)			10.6%	-	7.8%	-	21.8%	-
p75	54,528	12,493	61,929	14,593	54,233	13,026	62,965	19,865
(% diff vs. tax)			13.9%	19.1%	-0.5%	6.1%	15.6%	59.6%
p90	86,985	39,995	99,877	54,599	85,007	38,553	99,265	52,487
(% diff vs. tax)			15.0%	37.8%	-2.3%	-2.5%	14.2%	30.9%
has	0.77	0.39	0.75	0.38	0.81	0.39	0.79	0.41
(% diff vs. tax)			-1.7%	-1.9%	5.5%	1.7%	3.1%	6.7%

Panel II. Average levels and percent differences for older birth cohorts (1933-1939)

	Tax data		HRS		SIPP		CPS	
	68	78	68	78	68	78	68	78
p10	0	0	0	0	0	0	0	0
(% diff vs. tax)			-	-	-	0	-	-
p25	0	0	0	0	0	0	0	0
(% diff vs. tax)			-	-	-	-	-	-
p50	0	0	0	0	0	0	0	0
(% diff vs. tax)			-	-	-	-	-	-
p75	8,150	0	9,672	0	11,025	0	11,490	0
(% diff vs. tax)			19.1%	-	35.6%	-	44.6%	-
p90	28,472	3,737	35,679	4,742	32,538	4,403	39,746	12,950
(% diff vs. tax)			25.8%	28.9%	14.6%	16.5%	40.6%	245.5%
has	0.36	0.13	0.36	0.14	0.40	0.13	0.36	0.16
(% diff vs. tax)			-0.3%	5.0%	11.5%	-1.0%	0.8%	21.7%

Source: Authors' calculations. Households are defined as having wage income if wage income exceeds $500 in 2010 dollars. Each row displays average wage income estimates in each source at specified ages. Survey data results additionally display average percent differences between age-birth cohort estimates and the tax data, over either 1943-1949 birth cohorts or 1933-1939 birth cohorts. Zeroes are included in estimates of average levels. When tax data estimates equal zero, percent differences are recorded as missing and excluded from estimates of average percent changes. For percentiles with no non-missing percent changes, average percent changes are omitted from the table.

Table 4B. Comparisons between the distribution of other income in tax data and survey data at select ages

Panel I. Average levels and percent differences for younger birth cohorts (1943-1949)

	Tax data		HRS		SIPP		CPS	
	58	68	58	68	58	68	58	68
p10	0	0	0	0	0	0	0	0
(% diff vs. tax)			-	-	-	-	-	-
p25	0	0	0	0	0	0	0	0
(% diff vs. tax)			-	-	-	-	-	-
p50	0	0	0	0	0	0	0	0
(% diff vs. tax)			-	-	-	-	-	-
p75	1,322	12	3,644	293	334	0	0	0
(% diff vs. tax)			213.8%	558.2%	-76.8%	-100.0%	-100.0%	-100.0%
p90	13,260	8,048	32,386	17,276	7,981	3,994	10,352	5,044
(% diff vs. tax)			146.8%	107.4%	-39.8%	-50.5%	-21.7%	-37.2%
has	0.29	0.22	0.33	0.24	0.23	0.15	0.20	0.16
(% diff vs. tax)			17.4%	11.2%	-17.7%	-29.5%	-29.2%	-27.0%

Panel II. Average levels and percent differences for older birth cohorts (1933-1939)

	Tax data		HRS		SIPP		CPS	
	68	78	68	78	68	78	68	78
p10	0	0	0	0	0	0	0	0
(% diff vs. tax)			-	-	-	-	-	-
p25	0	0	0	0	0	0	0	0
(% diff vs. tax)			-	-	-	-	-	-
p50	0	0	0	0	0	0	0	0
(% diff vs. tax)			-	-	-	-	-	-
p75	0	0	744	0	65	0	0	0
(% diff vs. tax)			-	-	-	-	-	-
p90	6,076	2,990	19,940	11,533	4,715	1,989	2,077	928
(% diff vs. tax)			232.5%	293.0%	-22.9%	-34.1%	-66.0%	-68.1%
has	0.20	0.15	0.25	0.20	0.19	0.12	0.12	0.10
(% diff vs. tax)			24.5%	35.3%	-6.6%	-16.5%	-40.1%	-29.9%

Source: Authors' calculations. Households are defined as having other income if other income exceeds $500 in 2010 dollars. Each row displays average other income estimates in each source at specified ages. Survey data results additionally display average percent differences between age-birth cohort estimates and the tax data, over either 1943-1949 birth cohorts or 1933-1939 birth cohorts. Zeroes are included in estimates of average levels. When tax data estimates equal zero, percent differences are recorded as missing and excluded from estimates of average percent changes. For percentiles with no non-missing percent changes, average percent changes are omitted from the table.

Table 4C. Comparisons between the distribution of pension income in tax data and survey data at select ages

Panel I. Average levels and percent differences for younger birth cohorts (1943-1949)

	Tax data		HRS		SIPP		CPS	
	58	68	58	68	58	68	58	68
p10	0	0	0	0	0	0	0	0
(% diff vs. tax)			-	-	-	-	-	-
p25	0	0	0	0	0	0	0	0
(% diff vs. tax)			-	-	-	-	-	-
p50	0	5,082	0	1,680	0	1,487	0	0
(% diff vs. tax)			-	-68.3%	-	-71.1%	-	-100.0%
p75	5,837	21,672	1,123	14,823	2,850	14,222	-	11,589
(% diff vs. tax)			-79.4%	-31.9%	-50.7%	-34.1%	-100.0%	-46.4%
p90	26,745	42,708	20,865	35,395	19,510	32,478	14,269	27,865
(% diff vs. tax)			-21.7%	-16.9%	-27.1%	-23.7%	-46.8%	-34.8%
has	0.33	0.62	0.26	0.55	0.27	0.53	0.18	0.43
(% diff vs. tax)			-21.8%	-12.6%	-18.5%	-14.1%	-44.7%	-31.2%

Panel II. Average levels and percent differences for older birth cohorts (1933-1939)

	Tax data		HRS		SIPP		CPS	
	68	78	68	78	68	78	68	78
p10	0	0	0	0	0	0	0	0
(% diff vs. tax)			-	-	-	-	-	-
p25	0	0	0	0	0	0	0	0
(% diff vs. tax)			-	-	-	-	-	-
p50	5,372	7,462	2,783	3,430	3,393	3,799	0	275
(% diff vs. tax)			-47.7%	-53.2%	-36.6%	-48.3%	-100.0%	-96.4%
p75	19,397	22,124	14,512	14,510	14,677	15,871	9,638	10,463
(% diff vs. tax)			-25.0%	-33.4%	-24.3%	-27.4%	-50.3%	-53.0%
p90	37,274	42,567	30,108	31,686	29,417	33,412	23,736	26,391
(% diff vs. tax)			-18.6%	-23.9%	-21.3%	-20.8%	-36.3%	-38.3%
has	0.63	0.70	0.56	0.60	0.61	0.61	0.43	0.46
(% diff vs. tax)			-11.6%	-13.9%	-2.8%	-13.2%	-32.6%	-33.6%

Source: Authors' calculations. Households are defined as having pension income if pension income exceeds $500 in 2010 dollars. Each row displays average pension income estimates in each source at specified ages. Survey data results additionally display average percent differences between age-birth cohort estimates and the tax data, over either 1943-1949 birth cohorts or 1933-1939 birth cohorts. Zeroes are included in estimates of average levels. When tax data estimates equal zero, percent differences are recorded as missing and excluded from estimates of average percent changes. For percentiles with no non-missing percent changes, average percent changes are omitted from the table.

Table 4D. Comparisons between the distribution of Social Security income in tax data and survey data at select ages

Panel I. Average levels and percent differences for younger birth cohorts (1943-1949)

	Tax data		HRS		SIPP		CPS	
	58	68	58	68	58	68	58	68
p10	0	2,436	0	5,668	0	1,510	0	0
(% diff vs. tax)			-	127.3%	-	-50.1%	-	-100.0%
p25	0	9,537	0	10,297	0	9,489	0	7,994
(% diff vs. tax)			-	5.6%	-	-1.7%	-	-16.1%
p50	0	16,056	0	15,298	0	15,710	0	14,358
(% diff vs. tax)			-	-5.9%	-	-3.1%	-	-10.5%
p75	0	21,379	0	20,154	0	21,485	0	19,745
(% diff vs. tax)			-	-5.7%	-	0.1%	-	-7.6%
p90	10,478	25,763	10,151	24,141	8,220	26,142	9,158	24,399
(% diff vs. tax)			-3.0%	-6.0%	-21.6%	1.2%	-12.6%	-5.3%
has	0.19	0.91	0.22	0.95	0.18	0.90	0.17	0.85
(% diff vs. tax)			15.5%	3.3%	-6.3%	-1.4%	-9.8%	-6.5%

Panel II. Average levels and percent differences for older birth cohorts (1933-1939)

	Tax data		HRS		SIPP		CPS	
	68	78	68	78	68	78	68	78
p10	3,827	3,193	6,320	6,582	4,174	4,984	1,187	3,477
(% diff vs. tax)			66.2%	115.8%	10.3%	55.8%	-66.6%	11.0%
p25	9,412	10,023	10,143	10,661	8,726	11,006	8,748	10,508
(% diff vs. tax)			8.0%	7.0%	-7.2%	10.0%	-6.8%	4.9%
p50	14,604	15,796	14,215	14,958	13,139	15,959	13,557	15,574
(% diff vs. tax)			-2.5%	-4.0%	-10.0%	1.5%	-7.0%	-1.4%
p75	18,437	19,557	17,903	18,041	16,824	20,609	17,448	19,477
(% diff vs. tax)			-2.8%	-5.9%	-8.7%	6.1%	-5.3%	-0.4%
p90	21,697	23,229	21,094	22,478	20,296	24,466	20,900	23,252
(% diff vs. tax)			-2.8%	-1.1%	-6.5%	6.2%	-3.5%	0.1%
has	0.93	0.92	0.97	0.97	0.95	0.95	0.90	0.92
(% diff vs. tax)			4.2%	5.3%	1.7%	2.7%	-3.7%	-0.6%

Source: Authors' calculations. Households are defined as having Social Security income if Social Security income exceeds $500 in 2010 dollars. Each row displays average Social Security income estimates in each source at specified ages. Survey data results additionally display average percent differences between age-birth cohort estimates and the tax data, over either 1943-1949 birth cohorts or 1933-1939 birth cohorts. Zeroes are included in estimates of average levels. When tax data estimates equal zero, percent differences are recorded as missing and excluded from estimates of average percent changes. For percentiles with no non-missing percent changes, average percent changes are omitted from the table.

Table 4E. Comparisons between the distribution of capital income in tax data and survey data at select ages

Panel I. Average levels and percent differences for younger birth cohorts (1943-1949)

	Tax data		HRS		SIPP		CPS	
	58	68	58	68	58	68	58	68
p10	0	0	0	0	0	0	0	0
(% diff vs. tax)			-	-	-	-	-	-
p25	0	0	0	0	1	0	0	0
(% diff vs. tax)			-	-	-	-	-	-
p50	58	9	36	3	64	14	113	34
(% diff vs. tax)			-38.6%	-69.4%	26.9%	71.1%	120.7%	333.2%
p75	955	425	764	326	617	251	1,556	1,050
(% diff vs. tax)			-19.4%	-27.8%	-21.4%	-38.3%	81.4%	148.7%
p90	6,568	6,113	4,045	3,276	2,636	2,925	7,796	7,769
(% diff vs. tax)			-37.5%	-42.3%	-54.8%	-51.3%	25.8%	31.6%
has	0.39	0.46	0.33	0.41	0.33	0.48	0.42	0.52
(% diff vs. tax)			-13.2%	-19.5%	-15.5%	-1.6%	8.4%	13.6%

Panel II. Average levels and percent differences for older birth cohorts (1933-1939)

	Tax data		HRS		SIPP		CPS	
	68	78	68	78	68	78	68	78
p10	0	0	0	0	0	0	0	0
(% diff vs. tax)			-	-	-	-	-	-
p25	0	0	0	0	1	0	0	0
(% diff vs. tax)			-	-	-	-	-	-
p50	146	20	79	5	126	11	116	17
(% diff vs. tax)			-47.8%	-77.5%	-6.8%	-51.7%	-11.0%	7.3%
p75	2,212	948	1,531	908	1,169	239	2,533	878
(% diff vs. tax)			-29.3%	-6.7%	-40.7%	-75.2%	19.6%	-5.2%
p90	11,543	8,775	7,019	7,798	3,934	2,818	10,770	6,071
(% diff vs. tax)			-38.2%	-5.7%	-63.6%	-66.6%	-4.5%	-30.1%
has	0.47	0.51	0.40	0.44	0.40	0.46	0.43	0.47
(% diff vs. tax)			-15.4%	-20.9%	-15.0%	-14.5%	-7.9%	-6.3%

Source: Authors' calculations. Households are defined as having capital income if nominal capital income exceeds $10,000 times the annual effective federal funds rate in the year of measurement. Each row displays average capital income estimates in each source at specified ages. Survey data results additionally display average percent differences between age-birth cohort estimates and the tax data, over either 1943-1949 birth cohorts or 1933-1939 birth cohorts. Zeroes are included in estimates of average levels. When tax data estimates equal zero, percent differences are recorded as missing and excluded from estimates of average percent changes. For percentiles with no non-missing percent changes, average percent changes are omitted from the table.

Table 5. Differences in IRA balances between tax data and survey data
Panel I. Average levels and percent differences for younger birth cohorts (1943-1949)

	Age	HRS		SIPP	
		Level	Avg diff vs. tax	Level	Avg diff vs. tax
p75	58	40,648	18.5%	14,266	-60.1%
		[39,953 to 41,209]	[0.0% to 44.4%]	[6,173 to 18,932]	[-78.2% to -51.6%]
	63	67,451	8.6%	22,112	-65.8%
		[50,249 to 80,483]	[-7.8% to 26.9%]	[13,006 to 30,491]	[-79.6% to -51.9%]
	68	95,473	2.4%	32,117	-67.0%
		[80,932 to 113,927]	[-21.4% to 14.5%]	[11,888 to 55,255]	[-85.1% to -46.3%]
p90	58	183,167	31.4%	77,136	-44.3%
		[136,213 to 229,811]	[24.4% to 42.8%]	[53,154 to 89,791]	[-51.4% to -30.7%]
	63	242,826	12.1%	102,316	-54.6%
		[209,024 to 297,898]	[1.6% to 29.5%]	[76,505 to 130,261]	[-65.8% to -40.9%]
	68	307,290	-2.3%	151,578	-51.8%
		[252,969 to 348,404]	[-6.4% to 5.3%]	[73,882 to 229,685]	[-72.7% to -30.9%]
has	58	0.47	0.5%	0.37	-19.8%
		[0.44 to 0.50]	[-6.5% to 8.3%]	[0.34 to 0.40]	[-27.5% to -15.3%]
	63	0.46	-5.1%	0.38	-21.2%
		[0.44 to 0.49]	[-12.1% to -1.0%]	[0.35 to 0.42]	[-29.3% to -16.1%]
	68	0.48	-1.7%	0.37	-25.4%
		[0.48 to 0.49]	[-2.5% to -0.2%]	[0.32 to 0.43]	[-33.5% to -14.6%]

Panel II. Average levels and percent differences for older birth cohorts (1933-1939)

	Age	HRS		SIPP	
p75	68	56,253	-2.6%	20,612	-63.2%
		[48,750 to 69,267]	[-30.8% to 15.9%]	[16,326 to 28,494]	[-76.7% to -41.7%]
	73	58,247	2.0%	22,025	-64.8%
		[49,532 to 64,872]	[-21.6% to 27.1%]	[10,820 to 33,583]	[-79.9% to -49.7%]
	78	58,175	0.3%	15,411	-70.3%
		[38,140 to 78,423]	[-20.2% to 34.8%]	[8,469 to 27,423]	[-85.4% to -38.1%]
p90	68	162,842	-13.7%	105,084	-45.3%
		[154,374 to 174,554]	[-31.8% to 1.5%]	[84,993 to 163,615]	[-63.0% to -33.8%]
	73	196,304	0.3%	107,966	-48.7%
		[157,523 to 219,678]	[-15.8% to 28.8%]	[79,336 to 123,483]	[-57.6% to -38.3%]
	78	187,158	-10.9%	109,095	-44.7%
		[156,549 to 212,230]	[-19.5% to -1.6%]	[84,002 to 123,483]	[-58.0% to -30.3%]
has	68	0.47	-1.0%	0.37	-21.3%
		[0.45 to 0.50]	[-4.3% to 4.1%]	[0.33 to 0.41]	[-29.5% to -14.1%]
	73	0.44	-2.9%	0.35	-23.8%
		[0.41 to 0.47]	[-10.3% to 4.9%]	[0.33 to 0.37]	[-29.7% to -18.1%]
	78	0.44	0.7%	0.33	-24.5%
		[0.39 to 0.49]	[-10.0% to 11.9%]	[0.29 to 0.39]	[-32.1% to -10.0%]

Source: Authors' calculations. Households are defined as having IRA balances if IRA are positive. Each row displays the average level and average percent difference between survey estimates and baseline tax data across specified birth cohorts. Brackets report the minimum and maximum estimates for individual cohorts.

Table 6A. Changes in the tax data distribution of pre-tax income across birth cohorts

Panel I. Changes across birth cohorts at younger ages

	10th percentile			25th percentile			50th percentile			75th percentile			90th percentile		
Age	1944	1950	Change	1944	1950	Change	1944	1950	Change	1944	1950	Change	1944	1950	Change
58	5,497	3,880	-29.4%	20,280	19,149	-5.6%	41,092	41,564	1.1%	68,388	70,809	3.5%	107,213	112,039	4.5%
59	3,746	2,927	-21.9%	18,962	18,127	-4.4%	39,585	39,897	0.8%	66,557	68,981	3.6%	105,645	108,369	2.6%
60	4,450	3,821	-14.1%	19,257	18,496	-4.0%	39,983	40,098	0.3%	67,956	69,441	2.2%	109,341	110,728	1.3%
61	5,329	1,787	-66.5%	19,570	17,345	-11.4%	39,960	38,454	-3.8%	67,721	67,010	-1.1%	110,617	107,507	-2.8%
62	6,711	3,494	-47.9%	20,141	17,479	-13.2%	40,186	38,176	-5.0%	68,163	66,377	-2.6%	111,786	107,912	-3.5%
63	7,587	6,090	-19.7%	20,535	18,189	-11.4%	40,562	38,494	-5.1%	68,498	66,614	-2.7%	112,939	107,631	-4.7%
64	7,027	6,555	-6.7%	18,963	17,962	-5.3%	37,489	38,107	1.6%	62,563	66,023	5.5%	100,614	108,075	7.4%
65	6,944	6,938	-0.1%	18,635	18,277	-1.9%	36,738	38,338	4.4%	60,591	66,176	9.2%	95,570	107,517	12.5%
66	7,155	6,862	-4.1%	18,696	17,988	-3.8%	37,319	38,158	2.2%	61,750	65,456	6.0%	97,937	104,984	7.2%
67	6,357	7,245	14.0%	17,839	18,241	2.3%	36,047	38,474	6.7%	60,211	66,119	9.8%	95,522	106,768	11.8%
Avg			-19.6%			-5.9%			0.3%			3.4%			3.6%

Panel II. Changes across birth cohorts at older ages

	10th percentile			25th percentile			50th percentile			75th percentile			90th percentile		
Age	1933	1943	Change	1933	1943	Change	1933	1943	Change	1933	1943	Change	1933	1943	Change
68	6,818	7,451	9.3%	16,861	18,554	10.0%	31,950	35,724	11.8%	51,134	58,881	15.1%	82,210	93,612	13.9%
69	6,972	7,054	1.2%	16,734	17,854	6.7%	31,163	34,977	12.2%	48,941	58,157	18.8%	76,477	94,056	23.0%
68	6,818	7,451	9.3%	16,861	18,554	10.0%	31,950	35,724	11.8%	51,134	58,881	15.1%	82,210	93,612	13.9%
69	6,972	7,054	1.2%	16,734	17,854	6.7%	31,163	34,977	12.2%	48,941	58,157	18.8%	76,477	94,056	23.0%
70	6,205	7,208	16.2%	15,529	18,007	16.0%	30,037	35,424	17.9%	47,860	59,239	23.8%	76,735	96,266	25.5%
71	6,207	7,385	19.0%	15,491	18,073	16.7%	30,236	35,752	18.2%	49,101	60,608	23.4%	80,428	100,564	25.0%
72	6,189	7,402	19.6%	15,353	18,112	18.0%	30,215	35,667	18.0%	49,724	60,193	21.1%	82,595	99,581	20.6%
73	6,681	7,149	7.0%	16,117	17,572	9.0%	30,485	34,582	13.4%	50,772	58,269	14.8%	85,138	96,017	12.8%
74	6,626	6,986	5.4%	16,292	17,232	5.8%	30,779	34,202	11.1%	51,906	58,289	12.3%	88,447	97,485	10.2%
Avg			9.8%			11.0%			14.1%			18.1%			18.6%

Source: Authors' calculations. Pre-tax income is the sum of (i) wage income; (ii) pension income; (iii) Social Security income; (iv) capital income; and (v) other income. All values are baseline estimates deflated to 2010 dollars obtained by positioning the institutionalized population in the total income distribution using HRS data.

Table 6B. Changes in the tax data distribution of after-tax income across birth cohorts

Panel I. Changes across birth cohorts at younger ages

Age	10th percentile			25th percentile			50th percentile			75th percentile			90th percentile		
	1944	1950	Change	1944	1950	Change	1944	1950	Change	1944	1950	Change	1944	1950	Change
58	4,758	3,319	-30.2%	18,610	17,852	-4.1%	35,379	35,863	1.4%	55,965	58,601	4.7%	84,354	89,744	6.4%
59	3,040	2,293	-24.6%	17,583	17,097	-2.8%	34,482	34,895	1.2%	55,539	57,653	3.8%	85,321	87,437	2.5%
60	3,732	3,197	-14.3%	17,939	17,475	-2.6%	35,042	35,186	0.4%	56,982	58,104	2.0%	88,615	89,454	0.9%
61	4,748	1,324	-72.1%	18,340	16,646	-9.2%	35,255	34,463	-2.2%	57,154	57,307	0.3%	90,349	88,525	-2.0%
62	6,315	3,048	-51.7%	19,053	16,747	-12.1%	35,843	34,449	-3.9%	57,985	56,989	-1.7%	92,082	88,904	-3.5%
63	7,325	5,656	-22.8%	19,628	17,485	-10.9%	36,730	34,799	-5.3%	58,947	56,828	-3.6%	93,787	88,093	-6.1%
64	6,734	6,215	-7.7%	18,229	17,309	-5.1%	34,381	34,627	0.7%	54,380	56,760	4.4%	83,931	88,925	5.9%
65	6,619	6,701	1.2%	18,095	17,686	-2.3%	34,355	35,212	2.5%	53,429	57,213	7.1%	80,544	89,029	10.5%
66	6,860	6,635	-3.3%	18,226	17,473	-4.1%	35,259	35,512	0.7%	54,647	57,115	4.5%	82,505	87,546	6.1%
67	6,036	7,071	17.1%	17,509	17,846	1.9%	34,547	36,194	4.8%	53,927	58,044	7.6%	81,665	89,490	9.6%
Avg			**-20.8%**			**-5.1%**			**0.0%**			**2.9%**			**3.0%**

Panel II. Changes across birth cohorts at older ages

Age	1933	1943	Change	1933	1943	Change	1933	1943	Change	1933	1943	Change	1933	1943	Change
68	6,694	7,205	7.6%	16,568	18,215	9.9%	30,700	34,414	12.1%	46,185	53,025	14.8%	69,561	80,338	15.5%
69	6,786	6,805	0.3%	16,363	17,532	7.1%	30,159	33,827	12.2%	44,960	52,576	16.9%	65,799	81,184	23.4%
68	6,694	7,205	7.6%	16,568	18,215	9.9%	30,700	34,414	12.1%	46,185	53,025	14.8%	69,561	80,338	15.5%
69	6,786	6,805	0.3%	16,363	17,532	7.1%	30,159	33,827	12.2%	44,960	52,576	16.9%	65,799	81,184	23.4%
70	6,090	7,045	15.7%	15,320	17,713	15.6%	29,319	34,190	16.6%	44,481	53,460	20.2%	67,365	82,960	23.1%
71	6,078	7,222	18.8%	15,304	17,813	16.4%	29,548	34,549	16.9%	45,598	54,583	19.7%	70,500	86,439	22.6%
72	6,079	7,276	19.7%	15,210	17,870	17.5%	29,582	34,604	17.0%	46,070	54,441	18.2%	72,799	86,083	18.2%
73	6,621	7,028	6.1%	15,977	17,354	8.6%	29,898	33,647	12.5%	46,897	52,999	13.0%	75,003	83,008	10.7%
74	6,529	6,883	5.4%	16,163	17,053	5.5%	30,202	33,324	10.3%	47,910	53,014	10.7%	78,035	84,724	8.6%
Avg			**9.1%**			**10.9%**			**13.5%**			**16.1%**			**17.9%**

Source: Authors' calculations. After-tax income is pre-tax income, minus reported federal tax liability and estimated FICA taxes, plus any earned income tax credits. All values are baseline estimates deflated to 2010 dollars obtained by positioning the institutionalized population in the total income distribution using HRS data.

Table 7. Comparisons between changes in the pre-tax income excluding capital gains distribution across birth cohorts in the tax data and survey data

Panel I. Average changes from 1944-1950 birth cohorts at younger fixed ages (58-67)

	Tax data			HRS			SIPP			CPS		
	1944	1950	Avg. % change	1944	1950	Avg. % change	1944	1950	Avg. % change	1944	1950	Avg. % change
p10	6,589	7,481	15.5% [-1.3% to 50.3%]	8,353	10,076	22.7% [-2.0% to 58.5%]	7,294	8,408	15.9% [-5.5% to 51.1%]	7,559	8,544	13.4% [-11.3% to 38.5%]
p25	19,302	20,481	6.2% [1.5% to 11.3%]	19,954	21,400	10.8% [-2.3% to 25.1%]	17,791	19,216	8.1% [3.3% to 12.7%]	18,500	19,276	4.1% [-6.6% to 22.4%]
p50	39,946	39,837	-0.2% [-3.2% to 3.1%]	46,002	40,916	-7.8% [-19.5% to 1.4%]	35,715	35,422	-0.7% [-8.4% to 4.3%]	38,185	37,825	-0.9% [-11.2% to 8.1%]
p75	68,051	65,955	-3.1% [-5.3% to 0.4%]	76,495	68,936	-5.1% [-9.6% to 1.0%]	60,241	59,933	1.2% [-12.0% to 10.9%]	67,084	65,312	-2.7% [-14.2% to 8.5%]
p90	108,492	104,554	-3.6% [-5.9% to 0.0%]	131,028	113,511	-10.8% [-21.1% to 1.1%]	97,437	87,757	-8.5% [-19.5% to 2.2%]	107,386	103,450	-3.5% [-19.0% to 14.6%]
has	0.94	0.95	0.7% [-1.1% to 1.9%]	0.97	0.98	0.6% [0.0% to 1.3%]	0.95	0.97	1.6% [0.2% to 3.3%]	0.94	0.96	1.2% [0.2% to 2.5%]

Panel II. Average changes from 1933-1943 birth cohorts at older fixed ages (68-74)

	1933	1943	Avg. % change	1933	1943	Avg. % change	1933	1943	Avg. % change	1933	1943	Avg. % change
p10	7,227	8,486	17.5% [7.8% to 24.6%]	9,678	10,725	13.6% [2.8% to 32.9%]	9,190	10,135	12.2% [-7.8% to 39.5%]	9,114	9,825	7.9% [-4.6% to 18.3%]
p25	16,862	19,325	14.6% [8.0% to 19.3%]	16,985	18,055	5.0% [0.4% to 14.1%]	15,435	18,416	20.8% [9.7% to 37.8%]	14,614	16,928	15.9% [6.6% to 24.8%]
p50	31,227	36,303	16.3% [12.4% to 19.2%]	28,608	34,649	21.7% [14.6% to 26.5%]	25,073	31,069	23.5% [11.6% to 34.9%]	23,981	29,570	23.6% [16.5% to 32.2%]
p75	49,938	59,867	19.9% [13.8% to 23.4%]	48,445	60,634	23.3% [21.3% to 25.6%]	39,536	51,651	28.9% [11.4% to 42.8%]	40,695	51,532	27.1% [18.0% to 36.6%]
p90	79,757	96,291	20.9% [12.1% to 26.0%]	80,183	107,353	37.7% [20.1% to 62.4%]	56,384	82,515	44.2% [17.2% to 77.1%]	72,894	82,945	14.4% [2.3% to 26.7%]
has	0.96	0.97	0.6% [0.3% to 1.0%]	0.99	0.99	0.3% [0.0% to 0.9%]	0.98	0.98	0.2% [-1.9% to 1.4%]	0.97	0.97	0.3% [-0.8% to 1.4%]

Source: Authors' calculations. Households are defined as having pre-tax income if it exceeds $500 in 2010 dollars. Each row displays average pre-tax income levels at specified birth cohorts, along with average percent changes between those cohorts, over either ages 58-67 or ages 68-74. Brackets report the minimum and maximum estimates for individual ages.

Table 8. Changes in the tax data distribution of non-Social Security income across birth cohorts

	Panel I. Changes across birth cohorts at younger ages											
	25th percentile			50th percentile			75th percentile			90th percentile		
Age	1944	1950	Change	1944	1950	Change	1944	1950	Change	1944	1950	Change
58	16,914	15,757	-6.8%	39,188	39,619	1.1%	67,278	69,715	3.6%	106,431	111,271	4.5%
59	15,069	13,973	-7.3%	37,205	37,295	0.2%	65,188	67,423	3.4%	104,569	107,375	2.7%
60	14,610	13,738	-6.0%	37,042	36,936	-0.3%	66,016	67,582	2.4%	107,883	109,226	1.2%
61	14,115	11,708	-17.1%	36,295	34,782	-4.2%	65,228	64,700	-0.8%	108,794	105,942	-2.6%
62	12,815	9,847	-23.2%	34,561	32,528	-5.9%	63,848	62,550	-2.0%	108,527	105,081	-3.2%
63	10,315	7,818	-24.2%	31,516	29,846	-5.3%	61,254	60,189	-1.7%	106,990	102,814	-3.9%
64	7,943	6,563	-17.4%	27,000	28,248	4.6%	54,118	58,348	7.8%	93,868	102,121	8.8%
65	5,640	5,813	3.1%	23,664	26,827	13.4%	49,671	56,987	14.7%	86,610	100,335	15.8%
66	4,870	4,590	-5.8%	22,048	24,162	9.6%	47,138	52,486	11.3%	84,478	94,106	11.4%
67	3,299	3,967	20.3%	19,703	22,547	14.4%	43,157	49,914	15.7%	78,390	91,097	16.2%
Avg			**-8.4%**			**2.8%**			**5.4%**			**5.1%**
	Panel II. Changes across birth cohorts at older ages											
Age	1933	1943	Change	1933	1943	Change	1933	1943	Change	1933	1943	Change
68	4,327	3,538	-18.2%	17,729	18,966	7.0%	36,368	41,185	13.2%	66,456	75,615	13.8%
69	3,743	2,698	-27.9%	16,453	17,644	7.2%	33,685	39,607	17.6%	60,493	74,954	23.9%
70	2,511	2,552	1.7%	15,361	17,733	15.4%	32,545	40,354	24.0%	60,660	76,298	25.8%
71	2,406	2,500	3.9%	15,461	17,776	15.0%	33,617	41,287	22.8%	63,741	79,726	25.1%
72	2,266	2,165	-4.4%	15,328	17,232	12.4%	34,110	40,533	18.8%	66,043	78,392	18.7%
73	2,595	1,750	-32.6%	15,486	16,249	4.9%	34,942	38,618	10.5%	68,227	74,755	9.6%
74	2,566	1,593	-37.9%	15,592	16,161	3.6%	35,873	38,952	8.6%	71,583	76,357	6.7%
Avg			**-16.5%**			**9.4%**			**16.5%**			**17.6%**

Source: Authors' calculations. Non-Social Security income equals pre-tax income less Social Security income. All values are baseline estimates deflated to 2010 dollars obtained by positioning the institutionalized population in the pre-tax income distribution using HRS data. Estimates at the 10th percentile are omitted because they are typically zero or near-zero.

Table 9. Comparisons between changes in the non-Social Security income distribution across birth cohorts in the tax data and survey data

Panel I. Average changes from 1944-1950 birth cohorts at younger fixed ages (58-67)

	Tax data			HRS			SIPP			CPS		
	1944	1950	Avg. % change	1944	1950	Avg. % change	1944	1950	Avg. % change	1944	1950	Avg. % change
p25	9,377	10,559	11.1% [-16.8% to 31.9%]	11,903	11,317	5.8% [-30.8% to 42.7%]	9,958	10,956	11.7% [-19.8% to 74.7%]	9,291	10,013	1.5% [-60.7% to 35.0%]
p50	31,279	30,822	-2.2% [-12.6% to 6.2%]	38,729	32,304	-8.3% [-23.3% to 3.8%]	28,742	28,892	-0.2% [-8.8% to 10.1%]	32,387	30,738	-6.3% [-30.8% to 10.4%]
p75	60,989	58,290	-4.8% [-13.5% to 2.1%]	70,731	62,354	-3.6% [-10.0% to 1.0%]	55,161	54,125	0.1% [-10.2% to 11.6%]	62,616	59,798	-5.1% [-22.4% to 6.8%]
p90	102,937	98,654	-4.4% [-13.9% to 4.1%]	127,401	105,803	-12.7% [-23.2% to 1.1%]	92,975	84,054	-7.8% [-17.5% to 1.8%]	103,044	99,806	-3.2% [-14.1% to 13.4%]
has	0.84	0.85	1.7% [-0.9% to 3.6%]	0.84	0.87	5.6% [2.1% to 9.7%]	0.85	0.88	2.8% [0.7% to 6.5%]	0.83	0.84	1.4% [-1.2% to 4.1%]

Panel II. Average changes from 1933-1943 birth cohorts at older fixed ages (68-74)

|	1933	1943	Avg. % change	1933	1943	Avg. % change	1933	1943	Avg. % change	1933	1943	Avg. % change
p25	2,916	2,399	-16.5% [-37.9% to 3.9%]	3,569	2,571	-26.9% [-47.5% to 10.3%]	2,525	1,388	-45.5% [-76.3% to -21.2%]	452	400	-11.1% [-80.1% to 86.2%]
p50	15,916	17,395	9.4% [3.6% to 15.4%]	14,361	16,802	15.7% [-4.2% to 30.0%]	11,452	13,209	12.0% [0.2% to 23.3%]	9,322	12,484	37.4% [12.9% to 79.9%]
p75	34,449	40,077	16.5% [8.6% to 24.0%]	32,407	42,632	29.0% [25.3% to 32.8%]	25,398	32,864	27.8% [5.5% to 42.1%]	26,630	35,214	33.4% [13.2% to 42.5%]
p90	65,315	76,585	17.6% [6.7% to 25.8%]	63,369	95,479	54.1% [35.9% to 87.7%]	43,404	62,394	41.2% [17.3% to 91.1%]	57,840	67,263	17.3% [5.2% to 31.8%]
has	0.81	0.79	-1.9% [-3.2% to -0.6%]	0.83	0.81	-2.2% [-4.1% to -0.3%]	0.81	0.77	-5.1% [-8.8% to -1.2%]	0.73	0.73	0.2% [-3.6% to 7.9%]

Source: Authors' calculations. Households are defined as having non-Social Security income if non-Social Security income exceeds $500 in 2010 dollars. Each row displays average non-Social Security income levels (or baseline estimates of non-Social Security income in the tax data) at specified birth cohorts, along with average percent changes between those cohorts, over either ages 58-67 or ages 68-74. Brackets report the minimum and maximum estimates for individual ages. Estimates at the 10th percentile are omitted because they are typically near-zero in all sources.

Table 10. Fraction of households with no non-Social Security income and no IRA balances at age 72, by state

| State	Fraction in 1933	Rank of 1933 fraction	Change from 1933-1943 (pp)	Rank of change
MS	28.1%	1	-2.7%	50
LA	24.0%	2	-0.8%	39
KY	23.2%	3	-0.9%	40
TX	22.3%	4	-1.9%	49
GA	22.1%	5	-0.7%	38
AL	21.9%	6	0.2%	31
TN	21.7%	7	-1.0%	41
AR	21.3%	8	1.0%	19
NM	21.2%	9	-0.7%	37
WV	21.0%	10	-1.1%	42
OK	20.0%	11	0.2%	27
AK	20.0%	12	-5.7%	51
DC	19.9%	13	0.2%	29
SC	19.9%	14	-0.4%	36
CA	19.8%	15	3.0%	2
NC	19.0%	16	-1.4%	45
NY	18.1%	17	0.0%	33
ME	17.3%	18	1.3%	9
FL	16.6%	19	1.4%	8
AZ	15.7%	20	1.1%	13
NV	15.7%	21	3.1%	1
VA	15.1%	22	-1.8%	48
HI	15.0%	23	-0.2%	35
OR	14.8%	24	1.9%	4
ID	14.8%	25	1.1%	11
VT	14.5%	26	-1.4%	46
MT	14.4%	27	0.5%	23
RI	14.4%	28	1.0%	16
MO	14.3%	29	1.4%	7
WY	14.2%	30	-1.1%	43
IL	13.8%	31	0.3%	26
MA	13.8%	32	-0.1%	34
MD	13.5%	33	-1.5%	47
NJ	13.3%	34	0.1%	32
WA	13.2%	35	2.0%	3
CO	13.2%	36	1.4%	6
ND	12.8%	37	-1.3%	44
PA	12.5%	38	0.5%	24
OH	12.4%	39	1.0%	17
IN	12.3%	40	0.9%	20
SD	11.9%	41	0.3%	25
MI	11.8%	42	0.8%	21
UT	11.4%	43	1.1%	12
IA	11.1%	44	0.2%	28
KS	11.1%	45	1.7%	5
NH	11.1%	46	0.5%	22
NE	11.1%	47	1.1%	14
CT	10.8%	48	1.0%	18
DE	10.4%	49	0.2%	30
WI	9.6%	50	1.2%	10
MN	9.4%	51	1.0%	15

Source: Authors' calculations using the IRS tax data. Rows indicate the fraction of households in each birth cohort in the noninstitutionalized population at the indicated age with non-Social Security income of less than $500 in 2010 dollars and zero IRA balances.

**Figure 1A. Tax data distribution of pre-tax income by household age**
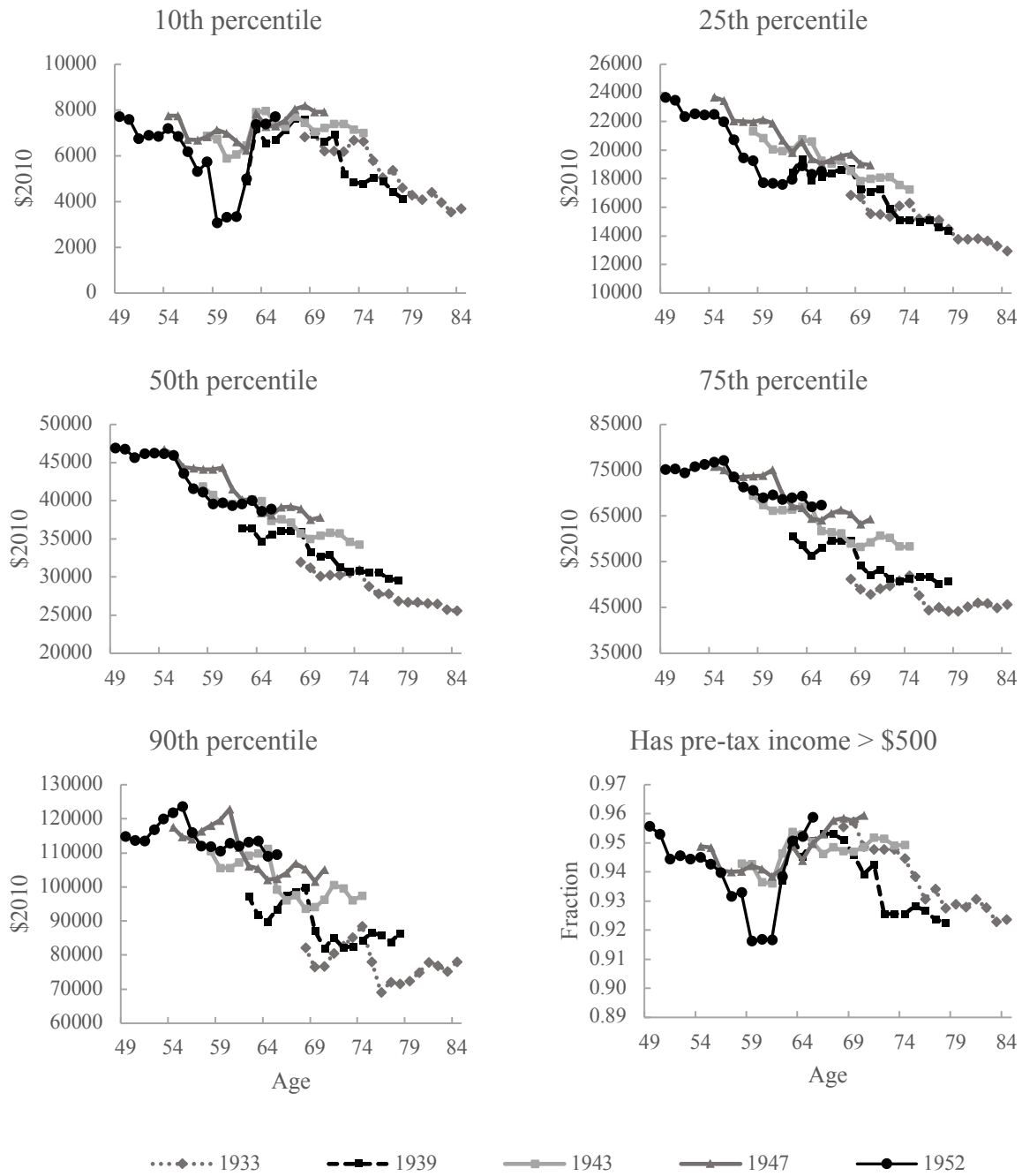
Source: Authors' calculations. Pre-tax income is the sum of (i) wage income; (ii) pension income; (iii) Social Security income; (iv) capital income; and (v) other income. Plotted values are baseline estimates obtained by positioning the institutionalized population in the pre-tax income distribution using HRS data.

**Figure 1B. Tax data distribution of after-tax income by household age**
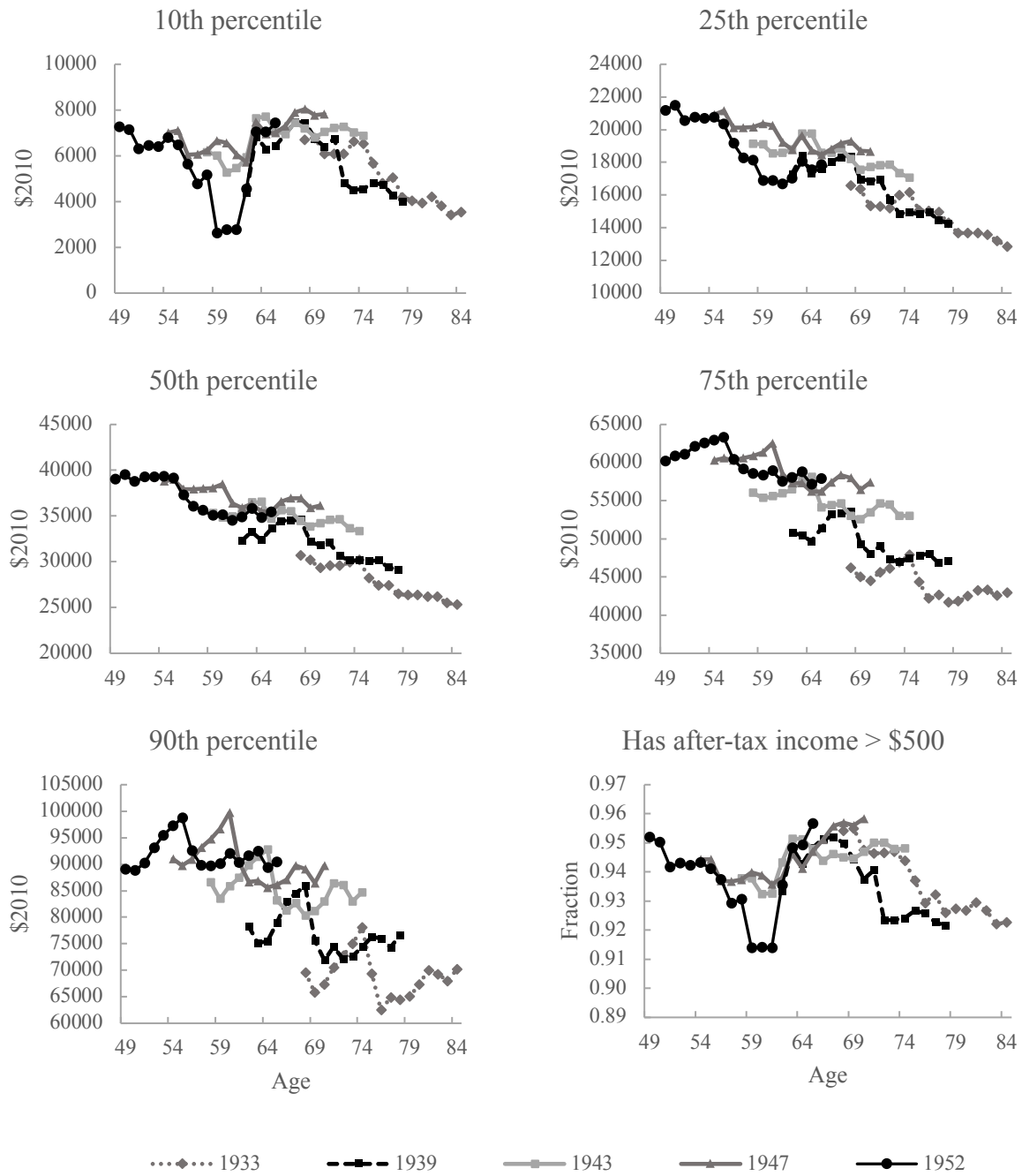
Source: Authors' calculations. After-tax income is pre-tax income, minus reported federal tax liability and estimated FICA taxes, plus any earned income tax credits. Plotted values are baseline estimates obtained by positioning the institutionalized population in the after-tax income distribution using HRS data.

Figure 2. Pre-tax income distribution excluding capital gains by age in tax data and survey data, 1941 birth cohort
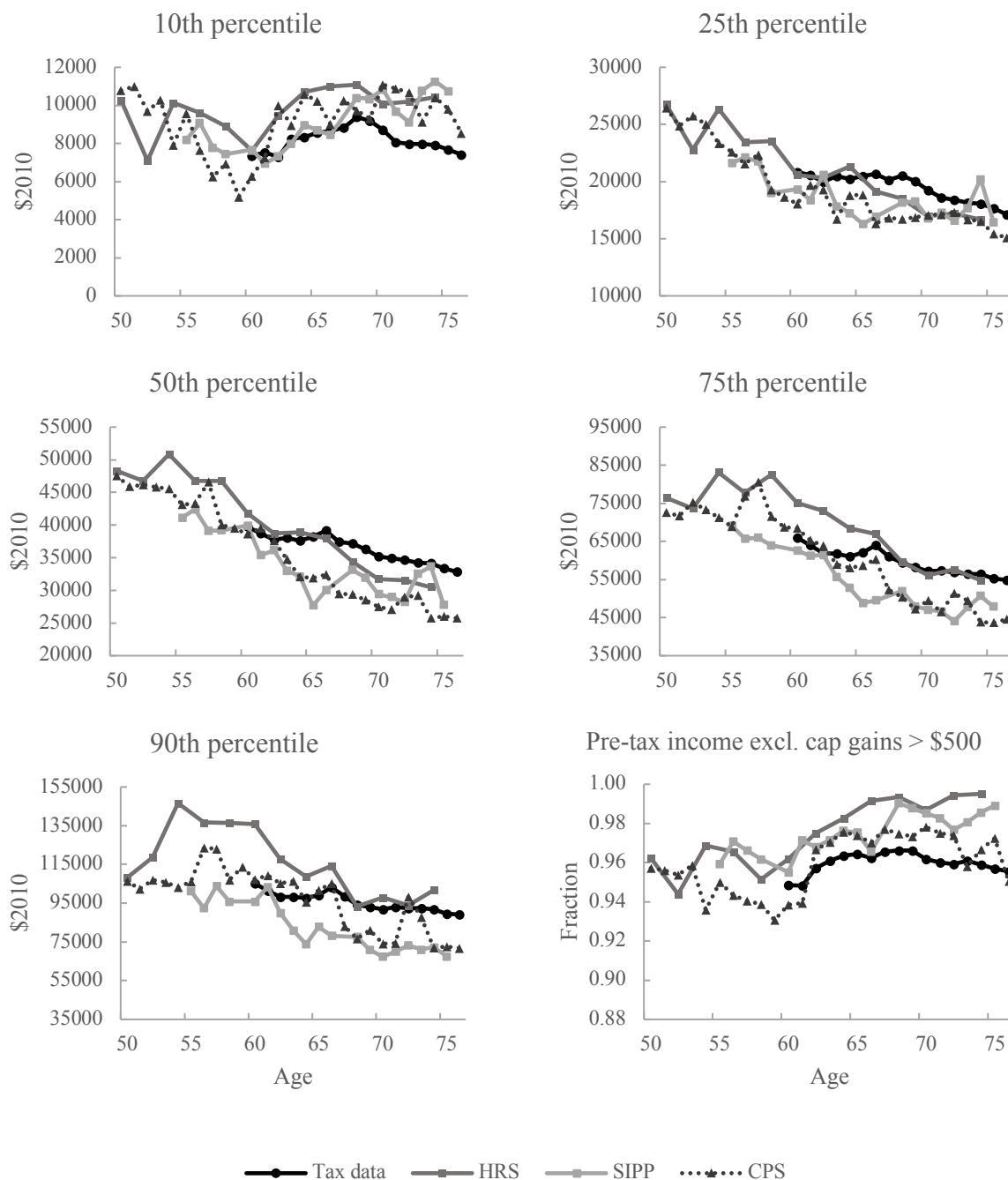
Source: Authors' calculations. Plotted tax data values are baseline estimates of pre-tax income excluding capital gains, where baseline estimates are obtained by positioning the institutionalized population in the pre-tax income distribution using HRS data. Plotted survey data values are corresponding pre-tax income estimates.

Figure 3. IRA balance distribution in the tax data by age for various birth cohorts

75th percentile

90th percentile

Has IRA balances

····◆···· 1933 ──■── 1939 ──■── 1943 ──▲── 1947 ──●── 1952
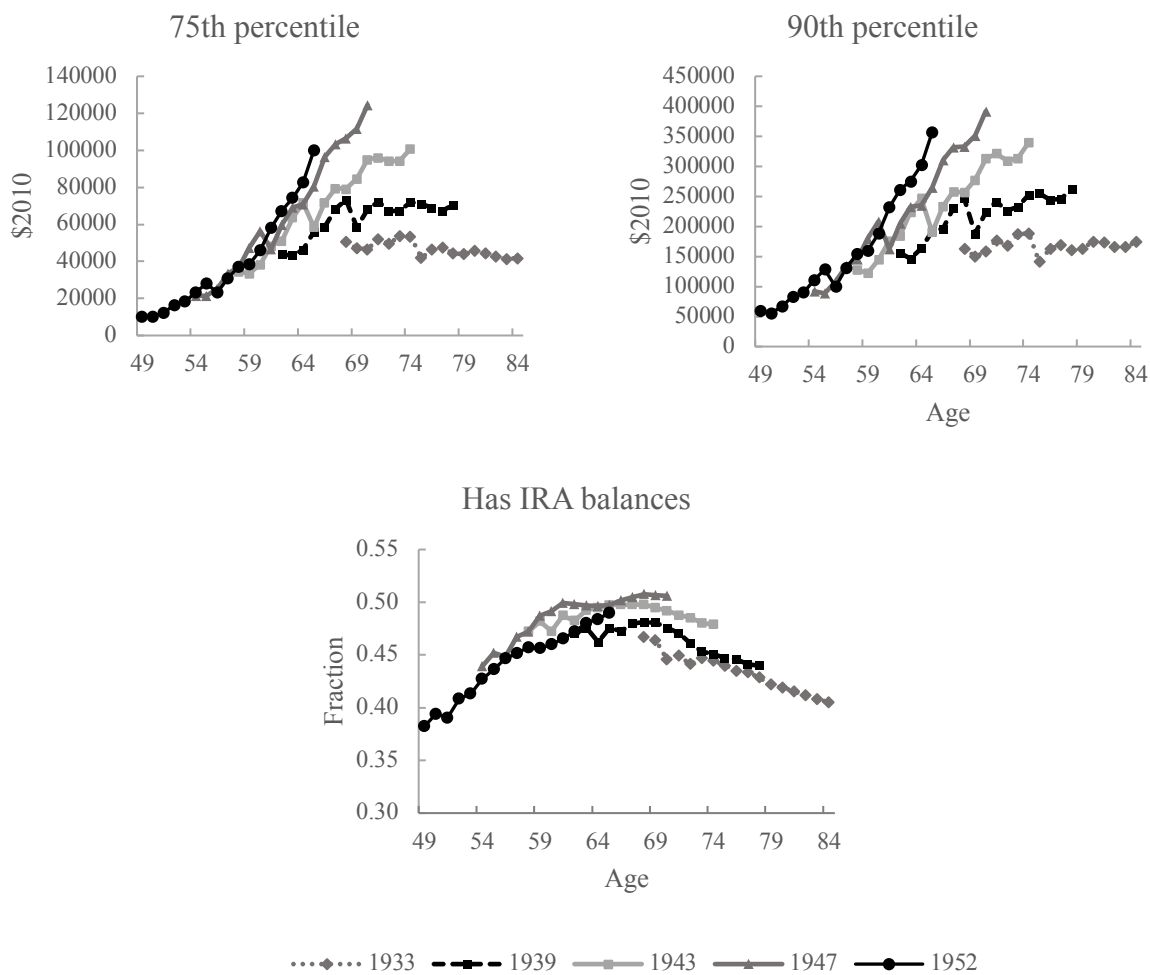
Source: Authors' calculations. Plotted values indicate baseline estimates of IRA balances, where baseline estimates are obtained by positioning the institutionalized population in the IRA balance distribution using HRS data. Households are defined as having IRA balances if IRA balances are positive.

Figure 4A. Fraction of households with no non-Social Security income and no IRA balances in the tax data, by birth cohort
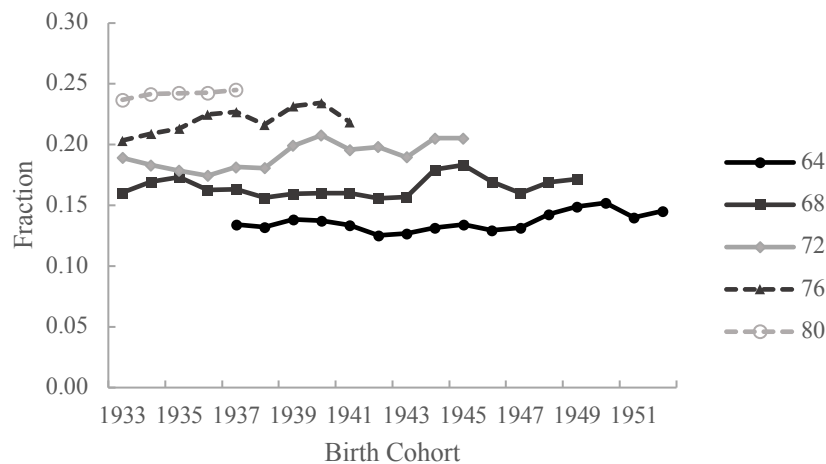

Figure 4B. Fraction of households with no non-Social Security income in the tax data and survey data, age 72
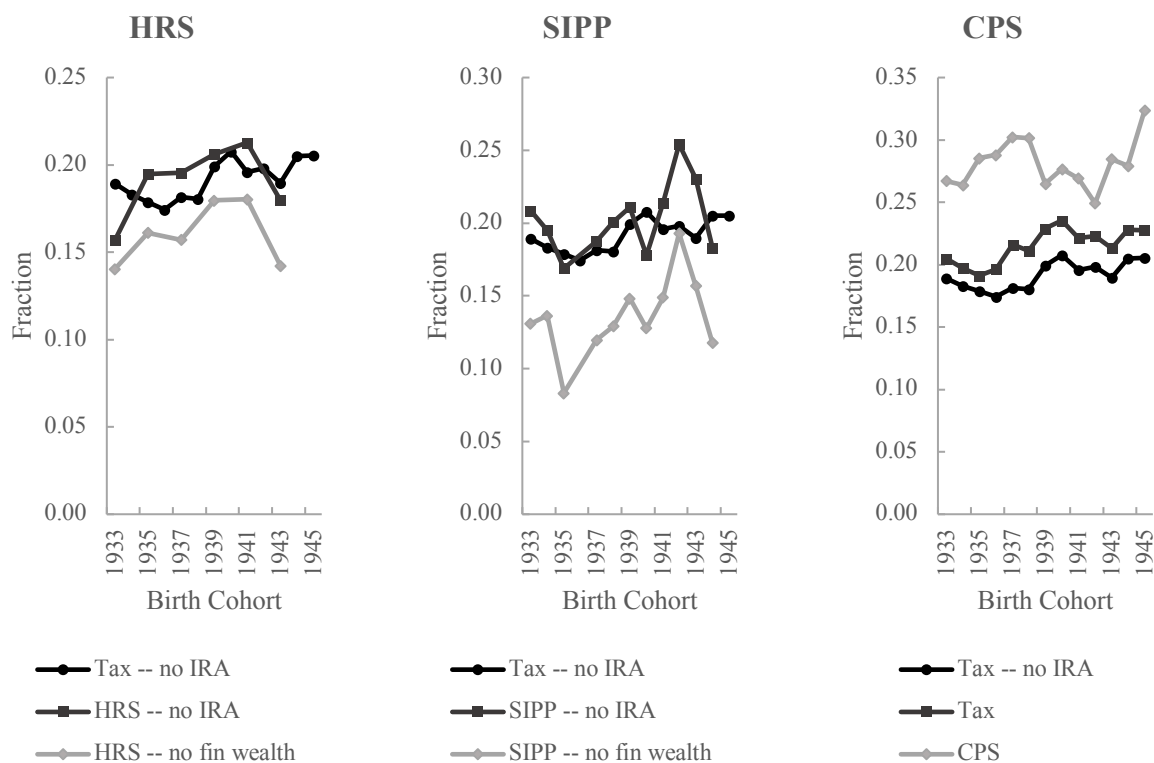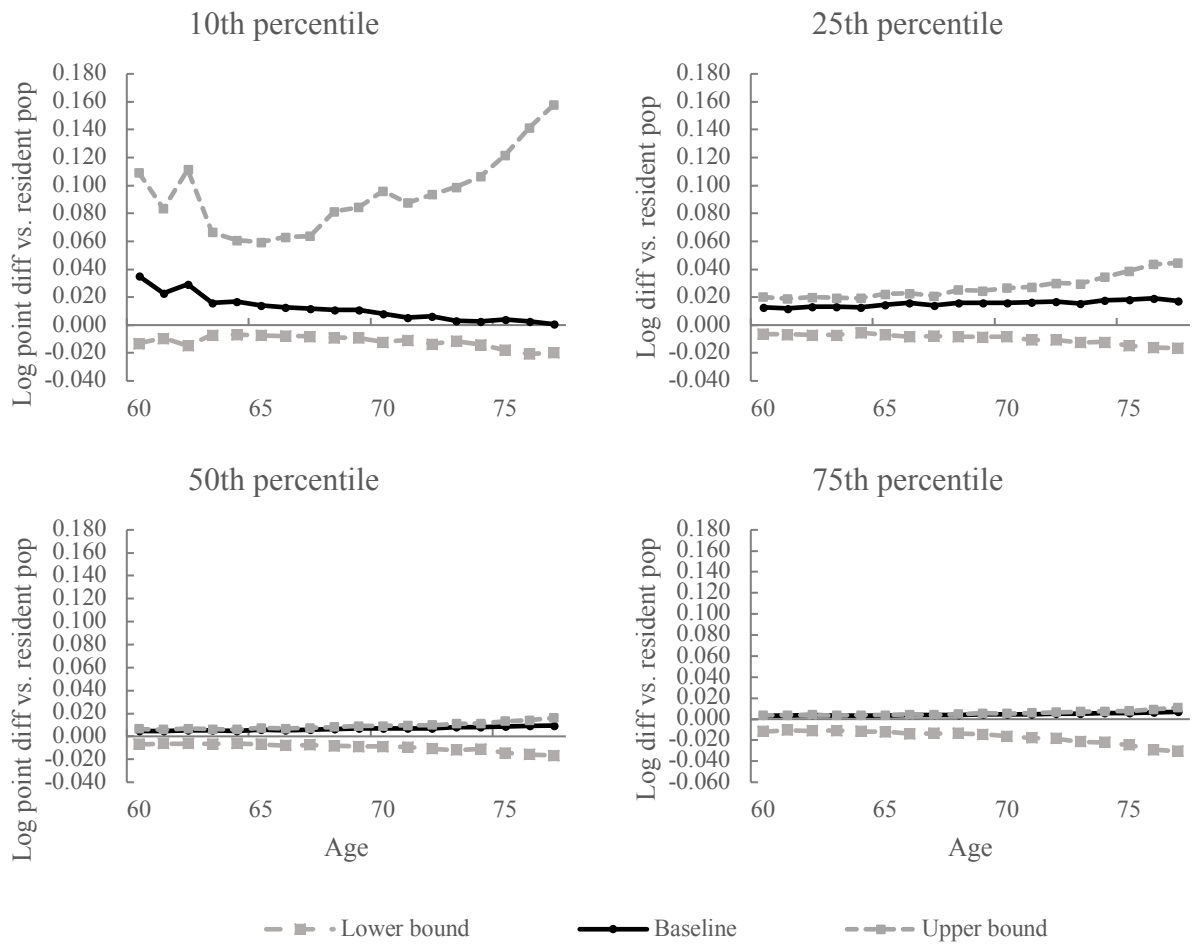

Source: Authors' calculations. Unless otherwise noted, plotted values typically indicate baseline estimates of the fraction of households in each birth cohort at the indicated age with non-Social Security income of less than $500 and zero IRA balances. Values without the "no IRA" suffix drop the restriction that households have no IRA balances. Survey data with the "no fin wealth" suffix indicate that households are excluded if they have financial wealth exceeding $3,500 in 2010 dollars. No asset screen is applied for CPS data.

Figure 5. Map of the fraction of households with no non-Social Security income and no IRA balances in the 1943 birth cohort observed at age 72
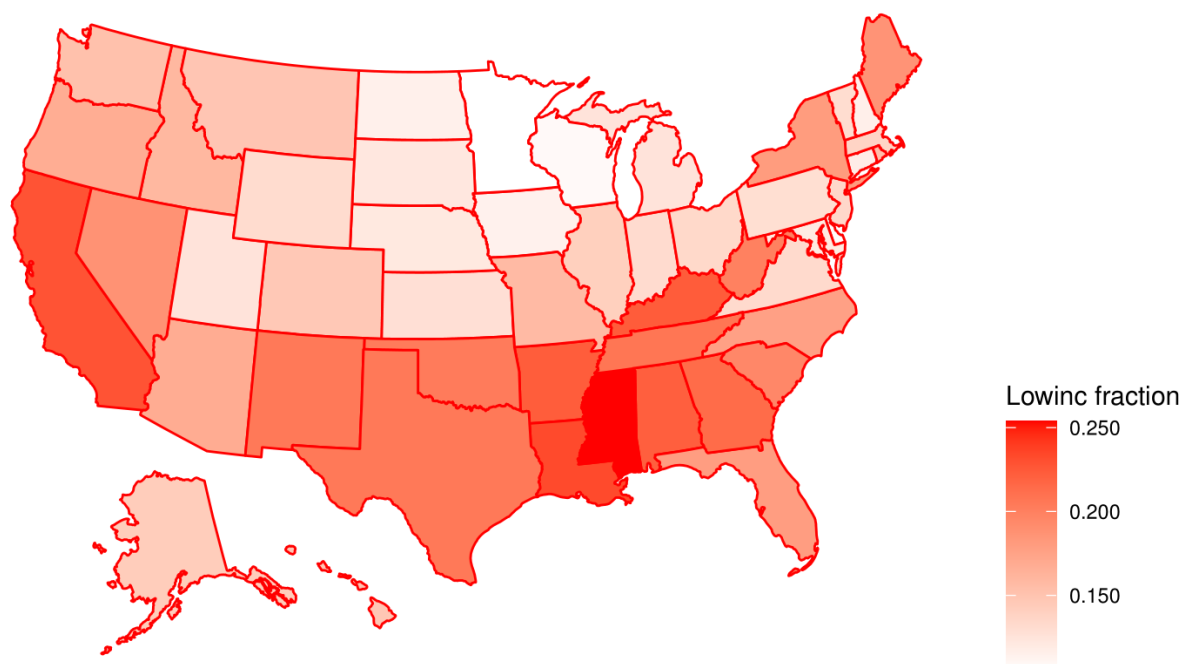
Source: Authors' calculations. Color intensity indicates the fraction of the noninstitutionalized population in the 1943 birth cohort at age 72 with non-Social Security income of less than $500 in 2010 dollars and zero IRA balances.

Appendix Figure A1. Tax data bounds and baseline estimates for pre-tax income, 1941 birth cohort

Source: Authors' calculations. Pre-tax income is the sum of (i) wage income; (ii) pension income; (iii) Social Security income; (iv) capital income; and (v) other income. Plotted values indicate the log point difference between the percentile estimate for the US civilian noninstitutionalized population and the estimate for the US resident population. Lower and upper bounds are obtained by assuming that institutionalized tax units lie at the top or bottom of the pre-tax income distribution, respectively. Baseline estimates are obtained by estimating the institutionalized population's position within the resident population distribution using data from the HRS.