

ARTICLE

Measuring income of the aged in household surveys: evidence from linked administrative records

Irena Dushi¹, Brad Trenkamp¹, Charles Adam Bee² and Joshua W. Mitchell³

¹Office of Research Evaluation and Statistics (ORES), US Social Security Administration, Washington, DC, USA, ²Social, Economic, and Housing Statistics Division, US Census Bureau, Washington, DC, USA and ³Center for Economic Studies, US Census Bureau, Washington, DC, USA

Corresponding author: Irena Dushi; Email: irena.dushi@ssa.gov

(Received 23 December 2022; revised 19 December 2023; accepted 23 January 2024)

Abstract

Household survey estimates of retirement income suffer from substantial underreporting which biases downward measures of elderly financial well-being. Using data from both the 2016 Current Population Survey Annual Social and Economic Supplement (CPS ASEC) and the Health and Retirement Study (HRS), matched with administrative records, we examine to what extent underreporting of retirement income affects key statistics: elderly reliance on social security benefits and poverty. We find that retirement income is underreported in both the CPS ASEC and the HRS. Consequently, the relative importance of social security income remains overstated – 53 percent of elderly beneficiaries in the CPS ASEC and 49 percent in the HRS rely on social security for the majority of their incomes compared to 42 percent in the administrative data. The elderly poverty rate is also overstated – 8.8 percent in the CPS ASEC and 7.4 percent in the HRS compared to 6.4 percent in the administrative data.

Keywords: elderly poverty; household surveys; measurement error; retirement income; social security

JEL Codes: H31; H55; I32; J14; J26

1. Introduction

As people transition from working life to retirement, their income sources change. They become less reliant on earnings and more reliant on income from social security, pensions, and personal savings, although the precise mix of these income sources varies across the income distribution (Mitchell *et al.*, 2022). This transition raises questions of retirement income adequacy, which is of primary importance in U.S. policy discussions. Most older Americans are beneficiaries of federal income support programs,¹ and the majority of federal budget expenditures will soon be devoted to supporting Americans aged 65 and over.

To assess the performance of the U.S. retirement system, including the relative importance of each of its components, analysts have long relied on public-use microdata from household surveys. For example, the Social Security Administration (SSA) Office of Research, Evaluation, and Statistics has published biennial statistical series on the income sources of the aged population since 1976.² Two of its most well-known publications are *Income of the Population 55 or Older* and the *Income of*

¹Social security (or old age, survivor, and disability insurance, often referred to as OASDI) is the largest social insurance program in the United States. In 2021, the program paid benefits to 65.2 million Americans, of which 47.3 million were retired workers beneficiaries (Table 5.A1, Social Security Administration, (2022)).

²Although ORES had previously published occasional statistics on income of the aged population, biennial publication began in 1976.

the Aged Chartbook. These publications provide financial well-being estimates for the older population and for various demographic subgroups. Such estimates include total money income, the prevalence and amount of income received from different sources (such as earnings, social security, pensions, and assets), and poverty status. Both publications are based on public data from the Current Population Survey Annual Social and Economic Supplement (CPS ASEC), also known as the ‘March CPS’,³ which is the nation’s official source of income and poverty statistics.

However, an accumulating body of research has cast doubt on the reliability of income measures from household surveys in general (Meyer *et al.*, 2015), and especially retirement income data in the CPS ASEC (Schieber, 1995; Woods, 1996; Iams and Purcell, 2013; Miller and Schieber, 2013; Bee and Mitchell, 2017; Dushi *et al.*, 2017; Brady and Bass, 2021).

Bee and Mitchell, linking the 2013 CPS ASEC microdata at the individual level to Internal Revenue Service (IRS) and SSA administrative records, established that reporting errors were widespread throughout the income distribution. About half of surveyed Americans aged 65 and over who received income from retirement plans (as indicated by IRS administrative data) reported no retirement income at all in the survey, such that the median household income was understated by approximately 30 per cent and the poverty rate was overstated by approximately a quarter (7 versus 9%). While reporting rates of distributions from defined contribution (DC) plans were lower than those from defined benefit (DB) plans, the relative prevalence and size of DB plans among current retirees meant that much of the underreporting could be attributed to DB pensions. In contrast, social security income was reported relatively well in the CPS; however, due to underreporting of retirement and asset income, estimates of beneficiaries’ relative reliance on social security had been overstated.⁴

Some analysts subsequently suggested that the measurement concerns were limited to the CPS and that other household surveys were more reliable. Chen *et al.* (2018) identified the University of Michigan’s Health and Retirement Study (HRS) as a promising candidate, as it focuses solely on households with a head aged 51 years or older, follows a panel of households over several years, and records data on sporadic withdrawals as well as regular payments. Thus, if any other household survey could be expected to better reflect the financial well-being of the elderly, it seems likely it would be the HRS.⁵

This paper’s main contribution is to examine whether previously documented income measurement issues are limited to the CPS or are also found in other household surveys like the HRS. Our findings will inform policymakers in general and SSA’s decision makers in particular: (a) whether the post-redesign CPS is a reliable and accurate data source for continued publication of income statistics for the aged population, (b) whether and to what extent administrative data can enhance the quality of those estimates, and (c) whether the HRS does a better job in measuring income of the aged. More specifically, we address the following questions. Are incomes of the elderly population as underreported in the HRS as they are in the CPS? If so, how does underreporting of retirement income affect key measures of the retirement system’s performance, such as reliance on social security benefits and the poverty rate of older Americans?

To address these questions, we analyze the 2016 wave of the HRS matched with SSA administrative records.⁶ We compare our findings to results derived from the 2016 CPS ASEC linked to administrative records from SSA and IRS. We show that even after the survey redesign, pension and asset income

³Every year, the Census Bureau fields the CPS ASEC in February through April and releases its results in September (U.S. Census Bureau, 2019). Hereafter in this paper, we use ‘CPS ASEC’ and ‘CPS’ interchangeably.

⁴The Census Bureau implemented a full redesign of income questions starting with the 2015 CPS ASEC, in part to collect more retirement income. However, the effects on income estimates appears to be limited, which Bee and Mitchell (2021) and the Census National Experimental Wellbeing Statistics program (Bee *et al.*, 2023) confirmed in subsequent analysis of post-redesign data linked to administrative records.

⁵Another candidate would be the Federal Reserve’s Survey of Consumer Finances (SCF). However, its smaller sample poses a challenge for detailed demographic breakdowns among the elderly.

⁶We focus on the 2016 survey year because at the time of this analysis it was the latest year of publicly available HRS data. The HRS is only linkable to SSA administrative data. Therefore, we cannot directly validate HRS pension and retirement

are still underreported in the CPS. We also find that once we account for underreporting of retirement income by using IRS administrative data, social security is no longer the main source of income for elderly beneficiaries – 53 percent in the survey versus 42 percent in the linked data receive at least half their family income from social security. Furthermore, we find that the poverty rate for the population aged 65 and over is substantially lower in the linked data than in the survey (6.4 versus 8.8%, respectively). Finally, we find that while the HRS does a better job than the CPS ASEC in terms of capturing retirement income, it still falls considerably short of the CPS ASEC when linked to SSA and IRS administrative records. Hence, key statistics such as the relative importance of social security income (49% in the HRS versus 42% in the linked CPS) and the poverty rate (7.4% in the HRS versus 6.4% in the linked CPS) remain overstated. Based on these findings, we conclude that income underreporting is not confined to the CPS ASEC, and surveys linked to a broad range of administrative records remain the most reliable tools to assess the performance of the U.S. retirement system. Therefore, it is our hope that decision makers at both the SSA and the U.S. Census Bureau will facilitate the regular usage of the CPS ASEC linked to administrative data for the continued production of the *Income of the Aged* publications.

The remainder of the paper proceeds as follows. Section 2 provides a brief background of the relevant literature. Section 3 describes the data and methods used for this analysis. Section 4 presents statistics from the different data files. Section 5 provides concluding remarks.

2. Literature review

Extensive literature has shown that household surveys are subject to several types of errors which can lead to inaccurate estimates of income and its distribution, poverty rates, and program participation (Bollinger and David, 2001; Johnson and Moore, 2008; Meyer and Wu, 2018; Meyer and Mittag, 2019; Meyer and Mittag, 2021; Corinth *et al.*, 2022; Celhay *et al.*, 2024). While total survey error comprises many forms of non-sampling error such as coverage error, item non-response, and misreporting, in the last two decades researchers have focused more on the issue of measurement error by taking advantage of survey data linked to administrative data (Meyer and Mittag, 2019).

Historically, the CPS ASEC has been widely used because of its large sample size, public availability, broad array of demographic, geographic and economic data collected, and annual periodicity. However, prior research has shown that different types of income in the CPS ASEC, relative to alternative data,⁷ are subject to reporting errors. Schieber (1995) was the first to show that, even as early as 1990, income from pensions and annuities was underreported in CPS ASEC by as much as one-third when compared to aggregate estimates from the National Income and Product Accounts and the IRS Statistics of Income division (SOI). Since then, as DC plans have become more prevalent among workers, the underreporting of income from pensions and retirement accounts continued to be identified by researchers as a major concern (e.g., Woods, 1996, Rector *et al.*, 1999, Roemer, 2000, Brady and Pierce, 2012, Iams and Purcell, 2013, Miller and Schieber, 2013, Brady and Bass, 2021),⁸ while other analysts identified lower aggregates in interest and dividends relative to benchmarks (e.g., U.S. Bureau of the Census, 1993; Ruser *et al.*, 2004).

income reports against IRS administrative data. Note that this paper is a revised version of our previous work (Dushi and Trenkamp, 2021) with a few additional estimates, particularly for the HRS data.

⁷Some other data sources used to compare income reported in the CPS include income information from: (a) National Income Products Accounts (see, e.g., U.S. Bureau of Census, 1993; Rector *et al.*, 1999; Roemer, 2000; Ruser *et al.*, 2004; Bosworth *et al.*, 2007; Rothbaum, 2015), (b) aggregate IRS Statistics of Income Division (SOI) administrative tax data (see, e.g., Schieber, 1995; Woods, 1996; Sabelhaus and Schrass, 2009; Brady and Pierce, 2012; Miller and Schieber, 2013; Brady and Bass, 2021), (c) SSA administrative records (see, e.g., Roemer, 2002; Sears and Rupp, 2003; Fisher, 2008; Davies and Fisher, 2009; Nicholas and Wiseman, 2009; Dushi and Iams, 2010; Iams and Purcell, 2013; Dushi *et al.*, 2017; Bollinger *et al.*, 2019), (d) matched individuals to IRS administrative records (Bee, 2013; Bee and Mitchell, 2017), or (e) other administrative data (see Meyer *et al.*, 2015).

⁸More recently, Brady and Bass (2021) compared 2010 annual income from the CPS with aggregate data derived from the IRS's SOI administrative tax data and showed that the CPS vastly understates the income of the elderly population. Like Bee

In the absence of validation studies based on individually linked data, retirement researchers were forced to speculate as to the causes of the large and growing aggregate discrepancies. Many pointed to the wording of the CPS retirement income question, which only asked about ‘regular’ distributions. As DC plans became more common among workers and younger cohorts of retirees, evidence that many plan participants withdrew money from these plans in irregular, ad hoc payments implied that CPS would miss these payments (Sabelhaus and Schrass, 2009). It also remained unclear to what extent aggregate discrepancies reflected underreporting only among the affluent or broadly across the distribution.

While the prior studies largely focused on aggregate income comparisons, Bee and Mitchell (2017) took advantage of microdata linkages between the 2013 CPS ASEC and administrative records from SSA and IRS to measure misreporting at the individual level.⁹ They compared amounts reported in the CPS with those reported in the matched administrative records for five types of income: earnings, social security benefits, supplemental security income (SSI) payments, interest and dividends, and ‘retirement income’ (comprised of DB pension benefits and DC account withdrawals apart from roll-overs and conversions). With respect to income sources that are available in the SSA administrative data – earnings, social security benefits, and SSI – the authors reported that among the population aged 65 and over, survey reported earned income¹⁰ and social security benefits track the respective measures in the administrative records very closely (98 and 96%, respectively), whereas survey reported SSI amounts comprise only 73 percent of the amount in the administrative records.¹¹ The authors also found that, among all households headed by an individual aged 65 or older, median household income was approximately 30 percent higher in the administrative records than in the CPS and that the survey-based poverty rate for persons aged 65 or older was overstated by approximately one-quarter.

Bee and Mitchell also showed that while earned income and social security benefits were reported accurately, the difference in estimated overall income was mainly due to underreporting of retirement income (from both DB pensions and DC withdrawals) and, to a lesser extent, of interest and dividends. The discrepancy in median income between survey and administrative data increased from 20 percent in 1990 to 30 percent in 2012. As a result of differential underreporting among income sources, Bee and Mitchell found that the CPS ASEC overstates the relative importance of social security income. Specifically, the proportion of persons aged 65 or older who rely on social security for at least 50 percent of their family income is overstated (55% in the survey versus 42% in the linked data) as is the proportion of those relying on social security for at least 90 percent of their family income (26% for the survey versus 12% for the linked data). These findings clearly quantify the magnitude of misreporting of retirement income and support the Census Bureau’s decision to redesign the income questions.¹² However, in preliminary work, Bee and Mitchell (2021) find that the effect of the redesign on income estimates was limited.

and Mitchell (2017), the authors find that while wages and salary and social security income are measured accurately in the CPS, it missed more than half of retirement income and more than one-third of investment income, while the differences are even large for the older individuals who get a larger share of their income from these two sources.

⁹O’Hara *et al.* (2016) compared respondents’ reports in the American Community Survey (ACS) with information from linked IRS records Form 1099-R and found evidence of underreporting of pension income. More specifically, the authors found a similar false-negative rate: about half of Americans aged 65 or older who received a 1099-R would not report that income to the ACS.

¹⁰Extensive research exists on the accuracy of reported earnings in household surveys. For further discussions of previous research see Roemer (2002), Bricker and Engelhardt (2007), and Abowd and Stinson (2013). However, given that most people aged 65 and over do not work, the analyses of accuracy of reported earnings have focused mainly on the working-age population.

¹¹See Koenig (2003), Koenig and Rupp (2004), and Nicholas and Wiseman (2009, 2010), for a discussion of reporting accuracy of social security programs. See also Meyer and Wu (2018) for a detailed review of the literature and their findings of the impact of social insurance and means-tested transfer programs on poverty reduction.

¹²We discuss the redesign of the CPS ASEC in Section 3.1. For further background details on the redesign see Semega and Welniak (2015) and Mitchell and Renwick (2015).

More recently Chen *et al.* (2018) compared reported income from the CPS ASEC and four other national household surveys with aggregates from SOI and SSA's Annual Statistical Supplement to the Social Security Bulletin. They concluded, as prior researchers have, that the CPS ASEC misses large portions of retirement income. Interestingly, they found that the other four national surveys were much better than the CPS in terms of capturing retirement income.¹³ Among the other surveys, the HRS emerges as a particularly attractive option for retirement researchers, as it is specifically designed to measure the well-being of older adults and produces publicly available microdata every 2 years.

Income misreporting in the HRS has also been studied. Retirement income, however, has not typically been the focus of prior work perhaps due to the lack of linkable IRS administrative records. Hurd *et al.* (2003) report that changing the HRS questionnaire to combine asset ownership questions with asset income questions appeared to reduce underreporting of asset income. Alwin *et al.* (2014) exploit the HRS's panel design to assess the reliability of its income data. While they were not able to assess retirement income items due to changes in the questionnaire, they find that measures of veterans' payments and social security are reported more reliably than SSI. Hyde and Harrati (2023) find net underreporting of SSI and disability insurance receipt, relative to linked SSA records.

Some HRS validations have shown that older workers do not accurately report their retirement program participation. Gustman *et al.* (2010), using information from employer plans, find workers are often confused about key plan characteristics. Dushi and Honig (2015), using information from SSA earnings records, find that workers often do not accurately recall whether they contribute to their retirement plans and misreport the amount of their contributions. Gustman *et al.* (2014) find that pairing a cohort of HRS workers in their early 50s to CPS ASEC responses in their late 60s reveals 'disappearing pensions', as over a third of implied annuitized pension wealth seemed to vanish in the CPS ASEC relative to the HRS.

3. Data and methods

The data we use in this paper come from two major household surveys – the 2016 CPS ASEC and the 2016 wave of the HRS – and from two sources of administrative data – SSA data covering earnings, social security benefits, and SSI payments, and IRS data covering interest income (both taxable and tax-exempt), dividends, and retirement income (both DB pensions and DC withdrawals). We compare several income and poverty estimates for the aged population using five alternative data files that we derive consisting of:

- (1) HRS
- (2) CPS ASEC
- (3) HRS data linked with administrative data from SSA (the 'HRS + SSA' data file)
- (4) CPS ASEC data linked with administrative data from SSA (the 'CPS + SSA' data file)
- (5) CPS ASEC data linked with administrative data from both SSA and IRS (the 'CPS + SSA + IRS' data file).

Our focus is on pre-tax family income of persons aged 65 or older.¹⁴ Following SSA's *Income of the Aged* publications, we provide estimates of total family income¹⁵ as well as separate estimates for each

¹³Chen, Munnell, and Sanzenbacher compared the CPS with the SCF, the HRS, the Survey of Income and Program Participation (SIPP), and the Panel Study of Income Dynamics (PSID). When comparing aggregate retirement income (other than social security) from SOI data with the survey results, they found that the SCF, HRS, SIPP, and PSID accounted for 99, 94, 97, and 85 percent of retirement income, respectively. However, they found that the CPS accounted for only about 47 percent of aggregate retirement income according to the SOI data.

¹⁴Note that the unit of analysis in both surveys is respondents aged 65 and over at the time of the survey. If married, their spouse can be of any age.

¹⁵In terms of the family unit, there are potential differences between the two surveys. In the CPS ASEC, a family is a group of two or more people related by birth, marriage, or adoption and residing together; all such people (including

of five income categories: social security benefits, income from pensions and retirement accounts, earnings, asset income (including interests, dividends, and rental income), and other income (including income from veterans' benefits, unemployment and workers' compensation, government transfer programs, and other income).¹⁶ For this analysis, we follow the recommendations of Anguelov *et al.* (2012) and count distributions from retirement accounts as income, including infrequent and periodic withdrawals. This includes payments from both DB and DC plans, as well as traditional and Roth Individual Retirement Account (IRA) withdrawals, but excludes transfers between tax-preferred accounts such as rollovers and conversions. This method was also employed by Bee and Mitchell (2017), and as they pointed out the nature of the administrative data lends itself to counting all withdrawals that permanently leave tax-preferred accounts as income. In both the CPS ASEC and HRS the interviews took place in 2016, but the reference year for all income measures is the previous calendar year, 2015.

3.1 Current population survey annual social and economic supplement

The CPS ASEC surveys a nationally representative sample of the U.S. non-institutionalized population. The survey collects detailed demographic information, including but not limited to age, gender, race/ethnicity, marital status, and household composition. In addition, the CPS ASEC collects information on income from different sources – such as earnings, social security, pensions, assets, and government transfer programs – that each household member aged 15 and over received during the previous calendar year. As such, the CPS ASEC has long been the official source of estimates of household income and poverty.

From the beginning, the main goal of the CPS ASEC was to capture regular cash income received that people could spend (Roemer, 2000). The Census Bureau's *money income* concept is defined as all regular income received before payments of personal income taxes, social security payroll tax, union dues, Medicare deductions, etc.¹⁷ Note that prior to the redesign of the CPS questionnaire, the survey did not ask respondents about certain income sources, and the Census definition of *money income* excluded irregular income such as any lump sum withdrawals from pensions or retirement accounts. As noted above, several studies have been critical of the CPS and have emphasized that income from assets and tax-advantaged retirement accounts (such as 401(k) plans and IRAs) are inadequately measured in the CPS, resulting in estimates that understate the importance of such accounts and consequently overstate the importance of social security benefits (Schieber,

related subfamily members) are considered members of one family. For single persons aged 65 and older living alone, family income will equal to own personal income. For couples living alone family income will be the combined income of the householder and the spouse. In other cases of living arrangements, family income will be the combined income of the person aged 65 or older and any related subfamily members (spouse, child, other) living in the household. The RAND-HRS public data file uses a 'household income' designation which is more restrictive than the CPS ASEC family income definition because it excludes the income of other residing family members besides the spouse (even though the incomes of other residing family members are collected in the survey). Hence, the HRS household income measures used in this study do not include income of other residing family members. More specifically, among married HRS respondents, household income equals the sum of income received from both spouses, whereas among non-married respondents (whether living alone or with others) household income equals their own reported income. Note that the HRS uses 'household' and 'family' income interchangeably. Given that income of other residing family members is not included in the HRS, it is likely that reliance on social security statistics will be biased. The poverty rate, however, is not affected because, to be consistent with Census Bureau definition of poverty, RAND-HRS created an income measure that includes income of all residing members (see Appendix B).

¹⁶See Appendices A and B for definitions of sources of income and poverty measures used in the CPS ASEC and the HRS, respectively. A detailed discussion about the numerous methods researchers have used to determine what counts as income is beyond the scope of this paper. In this paper we use Census's definition of *money income*.

¹⁷See Roemer (2000) and Ruser *et al.* (2004) for a comparison of personal and money income concepts. See also <https://www.census.gov/topics/income-poverty/income/about.html>. Appendix A provides a description of broad and detailed categories of income sources included in the *money income*.

1995; Woods, 1996; Fisher, 2008; Davies and Fisher, 2009; Iams and Purcell, 2013; Miller and Schieber, 2013; Munnell and Chen, 2014).¹⁸

The Census Bureau has been receptive to these studies and fielded a redesigned survey instrument aimed at improving the collection of income from different sources – particularly to capture irregular distributions from retirement accounts and for specific types of households (older and low-income) – by implementing several changes. These included eliminating redundant questions to reduce query fatigue and revising the order of the income questions to target the most likely sources of income. In addition, a ‘dual-pass’ approach was implemented that first asks about sources of income and then about the amounts from each source. Lastly, respondents are now asked about any distributions from DC accounts, which are now included in the Census definition of *money income*.

As noted above, of central interest is whether the underreporting of retirement income remains as prevalent as it was prior to the CPS redesign. Hence, we use data from the 2016 CPS ASEC collected during the second year of full implementation of the redesigned questionnaire. This allows us to compare our findings with Bee and Mitchell (2017), who assessed income reporting in the 2013 CPS ASEC prior to the redesign. We follow their work and link the CPS ASEC with restricted-access administrative records from SSA and IRS using an anonymized unique identifier on a secure Census Bureau server.^{19,20} This linkage allows us to validate data on several income sources by comparing respondents’ survey reports to respective amounts in the administrative records for the same year, 2015.

For respondents that are linkable (approximately 90% of respondents aged 65 or older), we replaced the values of income reported in the CPS ASEC with values from the administrative records. For this group of respondents, administrative records from SSA allow us to validate social security benefits (retirement and disability),²¹ SSI payments, and earnings from employment, resulting in what we call the ‘CPS + SSA’ data file. Adding the IRS administrative records allows us also to validate income from pensions and retirement accounts and income from interests and dividends, resulting in what we call the ‘CPS + SSA + IRS’ file. For the remaining respondents who are not linkable, we used their survey-reported values. It is worth noting here that Bee and Mitchell (2017) limited their analysis sample to linkable respondents and then reweighted the sample using inverse probability weights. Here, we decided not to exclude respondents who cannot be linked and thus we do not reweight the sample. The main reason is we want to maintain a consistent methodology with our analysis of the HRS which, as described below, has a much lower linkage rate to the SSA data and thus would be more sensitive to any reweighting procedure. However, it is worth noting that the 90 percent match rate in the CPS ASEC is high enough that any differences in the main

¹⁸It is worth noting, however, that Bee and Mitchell (2017) demonstrated that retirement income underreporting was also found to be substantial for traditional defined benefit pensions, suggesting that removing language about regular payments would not fully address the issue.

¹⁹The results presented in this paper were approved for release by the Census Bureau’s Disclosure Review Board (CBDRB-FY24-SEHSD003-010). This paper is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed on statistical issues are those of the authors and not necessarily those of the Census Bureau or the IRS.

²⁰We use the same SSA and IRS administrative data files that Bee and Mitchell (2017) used. The administrative data are linked to the CPS respondents using the Personal Identification Key (PIK), a scrambled Social Security Number (SSN), created by the Person Identification Validation System of the U.S. Census Bureau (Wagner and Layne, 2014). Survey respondents are more likely to be assigned a PIK by Census if they participate in any government programs or if they are tax filers. The Census Bureau uses the PIK to link administrative records with survey data. However, CPS respondents have the option to opt out of the linkage with their administrative records. Note that we use the words ‘linked’ and ‘matched’ interchangeably.

²¹Note that social security disability insurance benefits are converted to retirement benefits at full retirement age. Since our samples are comprised of people of 65 and over, the social security benefits are mainly retirement benefits.

results shown here versus restricting to linkable respondents and reweighting are found to be negligible.^{22,23}

3.2 Health and retirement study

The HRS is a comprehensive national longitudinal survey of Americans aged 51 or older. The first HRS interviews took place in 1992, with follow-up interviews conducted every other year since then. The main goal of the HRS is to provide data that allow researchers to examine interactions between social, economic, health, and psychological factors and the retirement decisions of older adults during pre- and post-retirement years. By conducting in-depth interviews, it also provides a broad array of information on topics such as employment, income, wealth, and other characteristics of the population aged 51 or older.²⁴ Another advantage of the HRS is that at the end of the interview it asks respondents for their consent to link their survey information with earnings and benefits information from social security administrative records.²⁵ Furthermore, from its inception the HRS has been more systematic than the CPS in collecting information on pensions, retirement plan account balances, and their distributions. The HRS respondents are first asked whether they have income from different income sources, pensions, and retirement accounts, and then about the amount of income received and frequency (with ‘other’ being one of the frequency options, in addition to the regular ones). If respondents, when asked, do not report the amount of income or wealth, then they are asked follow-up questions about the dollar amount using an ‘unfolding brackets’ approach to identify the range limits of the missing data item.²⁶ For respondents with unfolding brackets, their income is imputed and made available in the public use data file.

Czajka and Denmead (2008) showed that HRS-reported household income amounts in 2002, among people aged 51 or older, were substantially higher (by 20–30%) than amounts reported in the CPS ASEC; and while both samples had similar demographic characteristics, the HRS respondents were less likely to live alone than were their CPS ASEC counterparts. The authors conclude that ‘HRS incomes are higher than those of the Census Bureau surveys but resolving whether this is due to better measurement or over-representation of higher-income families must be left to future research’.

²²It is worth noting that evidence suggests that respondents who do not have a PIK and thus do not have a match to administrative records are less accurate in several parts of the survey (Bollinger and David, 2001; Celhay *et al.*, 2022). Respondents with and without a PIK may systematically differ in terms of their observable or unobservable characteristics as well as direction of misreporting of their income. On the one hand, excluding respondents without a PIK who in our sample are more likely to be of low income, assuming they report income accurately, is likely to lead to a less representative sample in the lower part of the income distribution (if reweighting is not fully adequate) and thus overestimate income and understate poverty rate and reliance on social security. On the other hand, if respondents without a PIK also underreport their income or program participation, including them in the sample will lead to understated income and consequently overstatement of poverty and reliance on social security.

²³Table A1 compares demographic characteristics of people with a match (or with a PIK in CPS) versus those without a match, separately in the HRS and CPS. The estimates show that CPS elderly respondents with a match are more likely to be non-Hispanic white, married, to have some college or a college degree, to live in a two-person family, to be social security beneficiaries, and have higher personal income. In contrast, those without a match are more likely to be female, non-white, never married, without a high school degree, in poverty, and in a three-or-more person family. Also, they are less likely to be social security beneficiaries and at the mean have lower self-reported income.

²⁴The HRS survey is conducted by the University of Michigan with support from National Institute on Aging and SSA. The raw data files are available at the HRS website, but compiling even a subset of the extensive amount of the available information would require a prohibitive amount of a user’s time. To make the HRS data more accessible for users, RAND Corporation—through a subcontract from HRS—compiles, maintains, and updates a user-friendly data file, which contains a subset of data with variables that are most widely used by the research community.

²⁵HRS respondents are asked if they agree to the match of their survey information with their SSA earnings and benefits records to be used for research purposes only. HRS respondents still have the option to opt out of the match even after providing the consent by writing to the SSA.

²⁶The CPS ASEC redesign adopted a similar approach.

In this article, we use income information from the RAND-HRS user-friendly data file, which includes information from all interviews conducted from 1992 to 2016,²⁷ standardized across survey years. Specifically, we focus on the sample of people who were aged 65 or older (if married the spouse can be of any age) in the 2016 wave and use only the income variables collected in that wave. For each respondent in the HRS, income measures include earnings, private pensions, social security benefits, income from government welfare programs, capital income, and other sources.²⁸ For married respondents, information about the spouse's income from those same sources is also collected and available as separate variables. Thus, the total household income (H13ITOT is the variable we use) is simply the sum of income received from all sources by the respondent and the spouse (if married or partnered). It is worth noting that while we use the 'household' designation to be consistent with the terminology used in the RAND-HRS file, we believe that it is in fact closer to the Census family income definition, particularly for respondents who live alone or with only a spouse/partner. Given that, as noted above, our measure of total household income does not include income of other residing family members, it is likely that our measure of total family income will be downward biased among respondents with three or more resident family members, and consequently reliance on social security will be biased.²⁹ For the purpose of estimating poverty status, RAND-HRS has created another total household income variable (named HwPOVHHI), which is derived by adjusting the abovementioned variable (HwITOT) in order to conform to the Census definition of family income by adding back the amount of Medicare Part B and/or D premiums (not included in the social security payments) and income of other resident family members, and subtracting food stamps and income of family members in nursing homes. We use this adjusted family income variable for estimating poverty status. However, given this adjustment includes several items, we cannot disentangle the amount of income of other resident members or the source of income (whether earnings, pensions, social security, etc.) and thus we cannot include it in our measures of family income; we acknowledge that our measure of HRS family income is likely understated.³⁰

To complement the HRS and to provide a comparison for the CPS + SSA data file, we also created an HRS + SSA data file. For the HRS survey respondents who consented to have their administrative records linked with HRS data, we matched their survey reports with information from SSA earnings and benefits records using anonymized unique household and person identifier on a secure virtual enclave at the University of Michigan.³¹

²⁷As of the earliest writing of this paper, 2016 was the last wave of HRS data that was publicly available (the final version 2 of data file that includes waves 1992–2016). As of this publication, more recent waves of data are available.

²⁸The RAND documentation data file states: 'We assume that educational assistance and other sources would have been reported as 'other income' in the HRS, but it is likely that at least some assistance from outside the household may not be included in any of the HRS income categories. The HRS total household income, e.g., as calculated in H6ITOT on the RAND HRS Longitudinal File, less food stamps, and including Medicare Part B and/or Part D premiums deducted from Social Security, would seem to be close to the Census definition of income, with the exception of income from resident family members besides the Respondent and spouse. ... Questions ask about the income of resident family members, including the earnings of each and total non-job income of them all. With these questions, we can estimate income of all resident family members, which is not included in HwITOT' (Bugliari *et al.*, 2020, 33).

²⁹In 2016, only about 16 percent of the HRS sample aged 65 or older lived in a family with three or more residing members (derived using variable H13HHRES, the number of residing family members in the household). It is worth noting here that there is another variable in the RAND-HRS user-friendly file called family composition HwPOVFAM, which is derived using household members (including their age, and the number of kids under 18 years old) as reported at the time of the interview. This latter variable is used to define the appropriate income level used to determine the poverty status. According to this latter variable 22 percent of respondents in our sample lived in families with three or more members (including children).

³⁰While the difference between the CPS and HRS regarding inclusion or exclusion of income of residing family members in total family income will likely affect the comparison between the two surveys, it will not affect the comparison within each survey.

³¹Since its inception in 1992, the HRS has asked respondents to sign a consent form to link the survey information with data from SSA earnings and benefits records. HRS respondents must opt in for SSA to allow the match of their records and make them available to HRS for research purposes only. However, respondents can opt out at any time by writing to SSA and asking to be removed from future matches. Consent forms have changed over the years, and the period over which the

For respondents with a match, we replaced survey earnings and social security benefits with their respective values from the administrative records. Out of 42,053 observations that ever entered the HRS sample, 20,912 observations were alive respondents in the 2016 wave, of which 9,986 observations were alive respondents aged 65 and over. From this sample we exclude 407 respondents residing in institutions to be consistent with the CPS ASEC universe, and the remaining 9,579 observations constitute the overall sample for this analysis.^{32,33} Table A2 shows the match rate overall and separately for each administrative data file for our sample of elderly respondents who had a valid consent for the income year 2015.³⁴

3.3 SSA administrative data

Previous research has shown that survey respondents may misreport their earnings, social security benefits, or transfer program participation (Rodgers *et al.*, 1993; Bollinger, 1998; Pedace and Bates, 2000; Bound *et al.*, 2001; Bricker and Engelhardt, 2007; Iams and Purcell, 2013; Meyer *et al.*, 2015; Dushi *et al.*, 2017). To account for potential reporting error, we use social security administrative records, which maintain information on annual earnings, social security (OASDI) benefits, and SSI payments. The earnings information comes from the Detailed Earnings Records file (DER),³⁵ which contains wages and salary payments for each job reported on Form W-2 (defined as the sum of wages, tips and other compensation, and employee tax-deferred contributions to retirement accounts) as well as Medicare-covered self-employment earnings reported on Form 1040 Schedule SE. Information about social security benefit amounts comes from the Payment History Update System (PHUS) file, which contains information separately for each month of the net amounts of benefits paid to a beneficiary as well as the amounts of the Medicare premium paid on the beneficiary's behalf to the Centers for Medicare and Medicaid Services. Hence, the true (or gross) monthly amount of social security benefits that a beneficiary is entitled to is the sum of the net benefit and the Medicare premium.³⁶ Finally, we use the information from the Supplemental Security Records (SSR) file to

consent is valid for a match with SSA earnings and benefits record varies depending on whether the consent form signed was retrospective or prospective and for how long. Due to those changes, it is likely that the match rate will differ across future waves.

³²Note that estimates in Dushi and Trenkamp (2021) do not exclude respondents who reside in institutions, leading to some differences with estimates in this paper.

³³Table A1 also provides the demographic characteristics of HRS respondents with and without a match. Estimates indicate that among HRS elderly respondents, those with a match, compared to those without a match, are more likely to be non-Hispanic white, widowed, to have a high school diploma, to live alone, and to be social security beneficiaries. In contrast, those without a match are more likely non-Hispanic Black and Hispanics, married, to have less than high school diploma, to live in two-person families, and to be in poverty. Interestingly, at the mean, those elderly respondents without a match have higher self-reported total income than those with a match. Across the two surveys, HRS respondents with a match are more likely than their CPS counterparts to be non-Hispanic whites and married, to live in families with three or more members, to be social security beneficiaries, and have higher self-reported total income. Among those without a match, HRS respondents are more likely than CPS respondents to be male, non-Hispanic white, married, to have less than a high school degree or to have some college or college degree, more likely to live in a two-person family, and more likely to be social security beneficiaries. In addition, they are less likely to be in poverty and have higher self-reported income.

³⁴For this study we use the latest version of available restricted administrative records (year 2022 version), which includes respondents who provided consent in any of the survey years prior to 2022.

³⁵IRS regularly provides earnings records to SSA for the purpose of administering its social insurance programs. Strictly speaking earnings records are IRS records, although often we refer to them as SSA records.

³⁶Meyer and Wu (2018) also use the gross amount of social security benefits in their analysis. Iams and Purcell (2013) report that social security income reported in the 2010 CPS closely correspond to the amounts in the social security records, whereas respondents in the 2009 SIPP survey understated their annual social security benefits by about \$1,000 per person on average. Dushi and Iams (2017) and Dushi *et al.* (2017) found that 2012 HRS respondents reported the net amount of annual social security benefits they received and not the gross amount; at the mean the difference in benefits between survey reported and administrative records was \$255 and \$1,270 when compared to the net and gross benefits, respectively, with the latter amount being almost equal to the median Medicare premium. As we will discuss below the under reporting of social security benefits is still the case for 2016 HRS data. Since our goal is to use the most accurate measure of total income and because

obtain information about SSI receipt and payment amounts for respondents with a match.³⁷ In all instances, for respondents who are not linkable to administrative data, we use their survey-reported values.

3.4 IRS administrative data

Available IRS administrative data include information on retirement income from sources other than social security and income generated from asset holdings. We match information from two IRS administrative data files to the 2016 CPS ASEC. Unfortunately, the IRS data are not linkable to the HRS.

The first of these files is composed of data taken from the IRS information return Form 1099-R. The 1099-R data allow us to validate retirement income received from both DB pensions as well as withdrawals from DC employer-sponsored plans and from IRAs. As Bee and Mitchell (2017) note, this file excludes data on direct rollovers, Section 1035 exchanges, and Roth IRA conversions. This is to our advantage, as we want to count only income that permanently leaves tax-preferred accounts.³⁸

The second IRS administrative data file we use is IRS Form 1040. From this file, we validate interest and dividend income for CPS ASEC respondents who filed a return for income year 2015. For CPS respondents who are linkable but did not file a 1040 we use the interest and dividend amounts reported in the survey.

3.5 A brief word on income not captured by administrative data

Administrative data, while generally reported with greater accuracy than survey data, are still not free of errors. Moreover, many types of income are simply not available in the set of administrative data we have access to. For example, income data collected in the CPS ASEC or the HRS that are not available in our administrative data files include unemployment insurance; workers' compensation; public assistance (other than SSI); rents, royalties, estates, and trusts; educational assistance; alimony; child support; and in-kind support from outside the household. For these income sources, we use the values reported in the survey. Additionally, the SSA DER includes only the taxable portion of self-employment earnings and likely misses earnings from the informal labor market. As a result, it is possible that among the self-employed, earnings in administrative records are under-reported. However, given that our focus is people aged 65 and over who are less likely to work, it is plausible that such underreporting is only a modest problem. Furthermore, while SSR contain information about federal SSI benefits, the portion of benefits received from states is incomplete or partially missing (see Meyer *et al.*, 2021). In sum, given our data miss several types of income, it is plausible that even our preferred estimates of income are still understated.

4. Results

In this section, we present estimates from each of the five data files: the HRS, the CPS ASEC, the HRS + SSA, the CPS + SSA, and the CPS + SSA + IRS. Table 1 compares the demographic characteristics of the CPS and the HRS samples. The two samples of aged respondents generally exhibit similar demographic characteristics. The two samples differ, however, with respect to the number of persons in the family and the proportion of those who are social security beneficiaries. Similar to findings by Czajka and Denmead (2008) using the 2002 HRS wave, we find that aged HRS respondents in 2016 are less

Medicare premium deductions are included in the Census money income and poverty definitions, our preferred measure that we use here is the gross amount of social security benefits.

³⁷Note that SSI payments in the SSR file are reported separately for monthly federal payments and monthly federally administered state payments. We sum both measures for all the months in 2015 to derive total federally administered SSI payments (Meyer *et al.*, 2021, use the same approach). It is worth noting that SSI payments will be underestimated to the extent that state SSI payments that are not federally administered are not available in the SSR.

³⁸Bee and Mitchell (2017) in their Appendix Table 9 benchmark the 1099-R data to aggregates reported in Argento *et al.* (2015) and published statistics from SOI. They show that the 1099-R data at the Census Bureau aligns well with tax data aggregates after removing rollovers, conversions, and other types of distributions that should not be counted as income.

Table 1. Demographic characteristics of respondents aged 65 or older in two surveys, by age group, 2016 (in %)

Characteristic	HRS					CPS				
	All	Age				All	Age			
		65–69	70–74	75–79	80+		65–69	70–74	75–79	80+
<i>Gender</i>										
Men	44.4	45.0	47.3	44.2	40.7	44.6	47.1	45.4	45.8	39.3
Women	55.6	55.0	52.7	55.8	59.3	55.4	52.9	54.6	54.2	60.7
<i>Race/ethnicity</i>										
Non-Hispanic white	79.9	78.3	79.4	79.8	82.9	77.2	75.2	76.9	78.1	79.6
Non-Hispanic Black	9.3	10.4	9.3	8.5	8.3	8.9	9.4	9.1	8.9	7.8
Non-Hispanic other race(s)*	2.7	3.3	2.8	2.6	2.0	5.9	6.5	6.0	5.3	5.1
Hispanic (any race)	8.1	8.0	8.5	9.2	6.8	8.1	8.8	8.0	7.7	7.5
<i>Marital status</i>										
Married	61.1	69.3	66.3	60.8	43.0	57.7	65.8	62.3	59.0	40.2
Non-married										
Widowed	21.7	10.1	15.0	24.2	45.1	24.0	10.5	18.7	25.3	48.5
Divorced/separated	12.4	14.8	13.6	11.4	8.0	13.0	17.0	14.2	11.2	7.4
Never married	4.8	5.8	5.1	3.6	3.9	5.2	6.7	4.8	4.4	3.9
<i>Educational attainment</i>										
Less than high school diploma	14.8	9.5	14.5	18.9	20.4	14.6	10.8	13.2	16.0	20.8
High school diploma or	32.9	28.5	33.5	35.6	37.1	33.2	29.5	32.4	36.1	37.4
Some college	24.3	28.5	23.8	21.7	20.2	24.0	26.5	25.6	23.3	19.4
College degree	28.0	33.5	28.2	23.8	22.3	28.1	33.2	28.8	24.5	22.5
<i>Persons in family</i> [†]										
One	28.3	23.2	24.6	28.1	40.0	32.1	25.6	29.3	31.5	45.0
Two	54.6	57.4	57.8	56.3	45.4	52.0	56.1	55.0	54.6	40.9
Three or more	17.1	19.3	17.5	15.6	14.6	15.9	18.3	15.8	13.9	14.1
<i>Social security beneficiaries</i> [‡]										
Weighted count (in '000)	48,258	16,936	12,014	8,268	11,040	47,550	16,500	11,500	8,397	11,160
Unweighted count	9,579	2,612	1,853	2,291	2,823	23,500	8,388	5,647	4,100	5,366

Notes: Authors' calculations based on 2016 wave of HRS and 2016 CPS ASEC. Estimates are weighted using survey weights. Total percentages may not add to 100 due to rounding. The HRS sample removes respondents in institutions to be consistent with the CPS sample. For brevity, in all the tables we refer to CPS ASEC as CPS. For HRS, the public data come from RAND-HRS public file 2016 version 3 (that includes data from 1992 to 2016).

*For the CPS sample, this category includes respondents identified as 'Asian alone', other race, and those who report multi-race categories. However, the majority of CPS respondents in this category are Asian alone.

[†]We derive these categories using the variable 'h13hhres' in the RAND-HRS data file, which indicates the number of residing family members in the household. There is another variable in this file, named the 'family composition' variable ('h13povfam') that was created in order to define poverty thresholds and poverty rates. We use the latter variable in an earlier version of the paper, but after discussion with RAND researchers, we decided to use the former variable instead to produce this table. It is worth noting that the overall pattern is similar regardless of which variable is used.

[‡]These are survey respondents who report receiving social security OASDI benefits.

likely to live alone than their CPS counterparts (approximately 28 versus 32%, respectively). Furthermore, according to survey reports, the proportion of respondents who are social security beneficiaries is higher in the HRS than in the CPS (approximately 90 versus 82%, respectively).

4.1 Aggregate income

We begin by discussing aggregate income for the population aged 65 and older. Starting first with the CPS, Table 2, panel A reveals that the aggregate total income of the U.S. elderly population is greater when using the administrative records than what is observed in the survey alone. Aggregate total income amounts to \$2,100 billion in the CPS and only slightly higher at \$2,129 billion (or 1.4% higher) in the CPS + SSA file.³⁹ This reflects the modest aggregate boost from substituting administrative record measures of earnings, OASDI benefits, and SSI. However, when examining the CPS + SSA + IRS file, the aggregate income of the elderly population amounts to \$2,520 billion, a substantial 20

³⁹All dollar figures are inflation adjusted using the CPI-U and expressed in 2022 dollars.

Table 2. Aggregate income amounts and shares by source for the U.S. population aged 65 or older: estimates from five alternative data files, 2015

Income source	Survey data (unmatched)		Survey data matched with administrative records*		
	HRS	CPS	HRS + SSA	CPS + SSA	CPS + SSA + IRS
<i>Panel A: Amount of aggregate income</i>					
Total amounts (in \$ billion)	2,934.5	2,100.0	2,852.3	2,129.0	2,520.0
Earnings	947.3	638.4	823.6	608.3	608.3
Social security	791.7	734.9	833.0	790.9	790.9
Pensions [†]	656.1	495.3	656.0	495.3	836.7
Asset income	244.1	178.7	244.1	178.7	228.6
Other	295.2	52.6	295.6	55.4	55.4
<i>Panel B: Share of income</i>					
Total %	100.0	100.0	100.0	100.0	100.0
Earnings	32.3	30.4	28.9	28.6	24.1
Social security	27.0	35.0	29.2	37.1	31.4
Pensions [†]	22.4	23.6	23.0	23.3	33.2
Asset income	8.3	8.5	8.6	8.4	9.1
Other	10.0	2.5	10.3	2.6	2.2
Weighted count (in '000)	48,258	47,550	48,258	47,550	47,550
Unweighted count	9,579	23,500	9,579	23,500	23,500

Notes: Authors' calculations based on the 2016 wave of HRS and 2016 CPS ASEC, matched with administrative data from SSA and IRS. We refer to CPS ASEC as CPS. For HRS, the public data come from RAND-HRS public file 2016 version 3 (that includes data from 1992 to 2016). The HRS sample removes respondents in institutions to be consistent with the CPS sample. Estimates are based on all respondents with or without a match. Records that are not linkable use survey values throughout. Estimates are weighted using survey weights. SSA administrative data contain wage and salary earnings from Form W-2, self-employment earnings from Form 1040 Schedule SE, OASDI benefits, and SSI benefits. IRS administrative data contain DB pension income and DC withdrawals (excluding rollovers and conversions) from Form 1099-R and interest and dividends from Form 1040. Income reference year is 2015, the calendar year prior to the interview. Dollar amounts are in billions of 2022 dollars, inflation adjusted using CPI-U. Total percentages may not add to 100 due to rounding.

*Reported estimates for CPS data are approved for release by the Census Bureau's Disclosure Review Board (CBDRB-FY24-SEHSD003-010). Analysis of the HRS data is performed at the HRS data enclave at the University of Michigan, and these estimates are approved for release by HRS's Disclosure Review Board in November 2023.

[†]Pension income includes DC withdrawals but excludes rollovers and conversions.

percent higher than the CPS alone. This reflects that the survey greatly underestimates retirement income and to a lesser extent asset income.

Another important finding is that the amount of aggregate income in the HRS is substantially higher than in the CPS. This pattern continues to hold with the linked data as aggregate income in the HRS + SSA data file is 13 percent higher compared to the CPS + SSA + IRS file (\$2,852 versus \$2,520 billion, respectively). With the important exception of pension income, incomes from all other sources are higher in the HRS + SSA data file than in the CPS + SSA + IRS file.⁴⁰ The largest proportional difference between the two data files is the amount of 'Other' income,⁴¹ which is 5.3 times higher in the HRS than in CPS. Note that, in the HRS + SSA file, earnings, social security benefits, and asset income are also higher (by approximately 35, 5, and 8%, respectively), whereas pension income is lower (by 28%) than in the CPS + SSA + IRS file.

In terms of each source's share of aggregate income, Table 2, panel B reveals that there are few differences between the CPS and CPS + SSA files. That is, when supplementing with the SSA administrative data, we observe only slight changes in the shares of earnings (from 30 to 29%) and social security (from 35 to 37%). This finding is consistent with the findings of Bee and Mitchell (2017) who found only small aggregate earnings and social security benefit differences between survey and administrative measures. Also of note, the 'Other' share of aggregate income changes modestly – the aggregate dollar amount increased by \$3 billion because of higher amounts of SSI found in the administrative data.

⁴⁰Comparing income estimates in HRS + SSA file with the two CPS matched files is somewhat imperfect. The lower match rate to the HRS means that a larger proportion of records continue to use survey reports rather than administrative values.

⁴¹See Appendices A and B, for CPS and HRS definitions of the 'Other' income category.

Pension income accounts for a much larger share of aggregate income in the administrative data than in the survey (23.6% in the CPS versus 33.2% in the CPS + SSA + IRS).⁴² The 10-percentage-point difference suggests that while the CPS redesign may have somewhat improved the reporting of pensions and retirement account withdrawals, its success in improving the measurement of retirement income has been limited relative to IRS data, a finding that is consistent with Bee and Mitchell (2017) and Chen *et al.* (2018). Particularly noteworthy is that in the CPS + SSA + IRS file, pension income accounts for the largest share of aggregate income (33%), followed by social security benefits (31%). In contrast, when only examining CPS or CPS + SSA data, social security accounts for the largest share (37%) with pensions only the third highest share (23%). Finally, even after the survey redesign, it is noteworthy that the asset income share is still understated in the CPS, although to a lesser extent than the pension income share.

Interestingly, in the HRS data, pension income accounts for almost the same share of aggregate income as in the CPS and CPS + SSA files (23%). Put another way, the HRS pension share is still substantially lower than that in the CPS + SSA + IRS file (22 versus 33%). It is worth emphasizing that, the aggregate dollar amount of HRS pension income is about 32 percent larger than the CPS amount (\$656 versus \$495 billion); however, this is still approximately 22 percent lower than the amount in the CPS + SSA + IRS file (\$835.7 billion).

4.2 Aged population by family income

Table 3, panel A shows the distribution of persons aged 65 or older by family income level. In general, in both surveys, the family income distribution is somewhat similar between the standalone survey and the survey matched with the SSA data file. However, there are noticeable differences in comparison to the CPS + SSA + IRS file, for which the general pattern is a shift in the distribution toward higher incomes. The effect of supplementing with SSA data and particularly with IRS data results in a decrease in proportion of individuals concentrated in the lowest part of the income distribution, with their shares more evenly distributed across the middle and upper parts of the income distribution. Respondents with annual family income below \$20,000 comprise about 10 percent of the aged population in the CPS + SSA + IRS file, compared with about 14 and 20 percent in the CPS and HRS, respectively. The same proportions are slightly lower in the CPS + SSA and HRS + SSA files (approximately 13 and 18%, respectively), suggesting that the combination of earnings, social security benefits, and SSI is underreported in the lower part of the distribution. Furthermore, elderly respondents with annual family incomes of \$75,000 or higher comprise about 46 percent of the population in the CPS + SSA + IRS data compared with 37 and 31 percent of the population in the CPS and HRS, respectively. These proportions are almost the same when we substitute information in the survey with SSA administrative data suggesting accurate reporting of those income sources in the upper part of the income distribution.

In addition, panel B of Table 3 shows family income at select percentiles. It is notable that despite large differences in aggregate income, at the 10th and 25th percentiles annual family income is almost the same in both HRS and CPS. Even at the median, family income is only modestly higher in the HRS (by approximately \$3,000). This pattern continues to hold when comparing the HRS + SSA and CPS + SSA files. There is a much greater difference between the HRS and CPS in the upper half of the distribution. At both the 75th and 90th percentiles, family income is higher in the HRS than in the CPS, and slightly higher in the standalone surveys compared to the surveys linked with SSA administrative data. Finally, relative to the CPS, the CPS + SSA + IRS file shows substantially higher family income at all percentiles. At the median, for example, total family income in the CPS is underreported by approximately 25 percent (\$54,840 versus \$68,650). Compared to the CPS + SSA + IRS, the HRS and HRS + SSA also fall short at the median by 18 and 14 percent, respectively.

⁴²Note that CPS and CPS + SSA files use pension income as reported in the survey, whereas CPS + SSA + IRS replaces the survey-reported pension income with information from the IRS administrative records.

Table 3. Percentage distribution of the population aged 65 or older, by family annual income: estimates from five alternative data files, 2015

Family income (\$)	Survey data (unmatched)		Survey data matched with administrative records*		
	HRS	CPS	HRS + SSA	CPS + SSA	CPS + SSA + IRS
<i>Panel A: Percentage distribution</i>					
Total	100.0	100.0	100.0	100.0	100.0
Less than 9,999	5.3	3.3	4.5	2.2	1.6
10,000–19,999	15.0	10.9	13.4	10.3	8.6
20,000–29,999	12.8	12.1	12.7	11.6	8.9
30,000–39,999	10.7	10.7	10.9	10.5	8.5
40,000–49,999	9.3	9.4	9.7	10.2	8.7
50,000–74,999	15.6	16.6	17.0	17.6	18.0
75,000–99,999	9.4	10.7	10.4	11.7	13.2
100,000+	21.9	26.3	21.4	25.8	32.6
<i>Panel B: Percentiles of family income (in 2022 \$)</i>					
10th	16,274	16,570	17,527	17,740	19,940
25th	28,683	28,950	31,053	30,810	37,270
50th	57,806	54,840	59,922	56,130	68,650
75th	111,124	103,400	109,318	102,200	119,000
90th	196,525	172,100	191,810	170,200	189,100
Mean	99,049	80,950	97,848	81,710	94,650
Weighted count (in '000)	48,258	47,550	48,258	47,550	47,550
Unweighted count	9,579	23,500	9,579	23,500	23,500

Notes: Authors' calculations based on the 2016 wave of HRS and 2016 CPS ASEC, matched with administrative data from SSA and IRS. We refer to CPS ASEC as CPS. For HRS, the public data come from RAND-HRS public file 2016 version 3 (that includes data from 1992 to 2016). The HRS sample removes respondents in institutions to be consistent with the CPS sample. Estimates are based on all respondents with or without a match. Records that are not linkable use survey values throughout. Estimates are weighted using survey weights. SSA administrative data contain wage and salary earnings from Form W-2, self-employment earnings from Form 1040 Schedule SE, OASDI benefits, and SSI benefits. IRS administrative data contain DB pension income and DC withdrawals (excluding rollovers and conversions) from Form 1099-R and interest and dividends from Form 1040. Income reference year is 2015, the calendar year prior to the interview. Dollar amounts are inflation adjusted using CPI-U. For disclosure purposes, in both surveys, reported percentiles are linearly interpolated across intervals of \$2,500. Total percentages may not add to 100 due to rounding.

*Reported estimates for CPS data are approved for release by the Census Bureau's Disclosure Review Board (CBDRB-FY24-SEHSD003-010). Analysis of the HRS data is performed at the HRS data enclave at the University of Michigan, and these estimates are approved for release by HRS's Disclosure Review Board in November 2023.

4.3 Sources of income

Table 4 shows the proportion of people aged 65 or older with any family income from each of seven income sources, overall and separately by gender. The estimates indicate that a substantially lower proportion of people aged 65 or older reported income from pensions and retirement account withdrawals in the CPS compared to the CPS + SSA + IRS (48.5 versus 68.4%). In the HRS approximately 60 percent of elderly respondents have family income from pensions or retirement accounts, suggesting that the HRS does a better but still imperfect job with retirement income reporting at the extensive margin.

Besides retirement income, there are a few other statistics from Table 4 that are worth highlighting. First, differences in the prevalence of social security and earnings across the CPS files are relatively small. This reflects the fact that the CPS does reasonably well in measuring social security benefits and earnings at the extensive margin. Second, the HRS shows a somewhat higher proportion of aged persons receiving social security benefits, but a lower proportion with earnings, despite our previous finding of a much higher HRS earnings aggregate.⁴³ Similarly, a higher proportion of the aged

⁴³Table A3, panel A, shows separately for social security benefits, SSI income, and earnings, the proportion of HRS survey respondents who reported their benefits or earnings either accurately or inaccurately according to the administrative records (the latter, either a false-positive or a false-negative error). Thus, with respect to social security benefits, the proportion who inaccurately report their benefits or earnings is rather small (except for the SSI false-negative error). In terms of reported amounts of social security benefits and SSI payments, survey reported amounts comprise 94 percent of their benefits in administrative records. With respect to earnings, the vast majority (83%) of respondents report consistently with administrative records either a positive or zero amount. However, among the sample with matched records, at the mean, respondents

Table 4. Percentage of individuals aged 65 or older with family income from each source, by gender: estimates from five alternative data files, 2015

Income sources	Survey data (unmatched)		Survey data matched with administrative records*		
	HRS	CPS	HRS + SSA	CPS + SSA	CPS + SSA + IRS
All					
Social security	93.1	86.0	93.2	89.6	89.6
Asset income	58.1	68.4	58.1	68.4	65.6
Pensions [†]	59.8	48.5	59.8	48.5	68.4
Earnings	36.7	40.7	39.6	45.0	45.0
Veterans' benefits [‡]	10.1	5.6	10.1	5.6	5.6
Cash public assistance [§]	9.5	4.6	9.9	6.2	6.2
Other	7.7	5.7	7.7	5.7	5.7
Weighted count (in '000)	48,258	47,550	48,258	47,550	47,550
Unweighted count	9,579	23,500	9,579	23,500	23,500
Men					
Social security	92.1	84.7	92.3	88.7	88.7
Asset income	61.9	70.6	61.9	70.6	67.3
Pensions [†]	61.1	49.5	61.1	49.5	68.7
Earnings	43.0	44.2	45.8	48.6	48.6
Veterans' benefits [‡]	12.7	7.5	12.7	7.5	7.5
Cash public assistance [§]	7.7	4.0	7.9	5.1	5.1
Other	8.8	5.4	8.8	5.4	5.4
Weighted count (in '000)	21,447	21,210	21,447	21,210	21,210
Unweighted count	3,928	10,500	3,928	10,500	10,500
Women					
Social security	93.9	87.0	94.0	90.3	90.3
Asset income	55.1	66.6	55.1	66.6	64.3
Pensions [†]	58.7	47.7	58.7	47.7	68.1
Earnings	31.7	37.8	34.6	42.0	42.0
Veterans' benefits [‡]	8.1	4.1	8.1	4.1	4.1
Cash public assistance [§]	11.0	5.1	11.5	7.0	7.0
Other	6.8	5.9	6.8	5.9	5.9
Weighted count (in '000)	26,811	26,340	26,811	26,340	26,340
Unweighted count	5,651	13,000	5,651	13,000	13,000

Notes: Authors' calculations based on the 2016 wave of HRS and 2016 CPS ASEC, matched with administrative data from SSA and IRS. We refer to CPS ASEC as CPS. For HRS, the public data come from RAND-HRS public file 2016 version 3 (that includes data from 1992 to 2016). The HRS sample removes respondents in institutions to be consistent with the CPS sample. Estimates are based on all respondents with or without a match. Records that are not linkable use survey values throughout. Estimates are weighted using survey weights. SSA administrative data contain wage and salary earnings from Form W-2, self-employment earnings from Form 1040 Schedule SE, OASDI benefits, and SSI benefits. IRS administrative data contain DB pension income and DC withdrawals (excluding rollovers and conversions) from Form 1099-R and interest and dividends from Form 1040. Income reference year is 2015, the calendar year prior to the interview.

*Reported estimates for CPS data are approved for release by the Census Bureau's Disclosure Review Board (CBDRB- FY24-SEHSD003-010). Analysis of the HRS data is performed at the HRS data enclave at the University of Michigan, and these estimates are approved for release by HRS's Disclosure Review Board in November 2023.

†Pension income includes DC withdrawals but excludes rollovers and conversions.

‡Information on veterans' benefits is not available in the administrative data and thus the percentages in the CPS do not change across the alternative data files.

§Cash public assistance includes SSI, which is available in the SSA administrative data. Therefore the percentages change when we supplement with SSA data.

||This income category is not available in administrative data and thus the percentages remain the same across alternative data files.

population in the HRS than in CPS report receiving income from veterans' benefits and public assistance (which includes SSI payments). The increase in the proportion with earnings and public assistance in the HRS + SSA file compared to the HRS file again indicates extensive margin underreporting of income from these sources. Lastly, the breakdown by gender shows that the similarities and differences across the five files are relatively consistent. Still, it is apparent that, across all data files, women are more likely than men to have family income from social security and from public assistance. In

reported earnings that were 41 percent higher than the 'true' value in their restricted administrative records (\$57,934 versus \$41,181). We could not provide a similar table for the CPS data.

Table 5. Percentage of individuals aged 65 or older for whom social security income represents a selected proportion of family income, by gender: estimates from five alternative data files, 2015*

Social security as a % of family income is:	Survey data (unmatched)		Survey data matched with administrative records [†]		
	HRS	CPS	HRS + SSA	CPS + SSA	CPS + SSA + IRS
All					
50% or more	49.1	52.5	51.1	52.7	41.7
75% or more	29.6	33.6	29.6	33.0	21.3
90% or more	20.5	25.6	19.9	24.0	13.7
Weighted count (in '000)	44,917	40,520	44,975	42,300	42,270
Unweighted count	8,987	20,000	9,026	21,000	20,900
Men					
50% or more	43.8	49.0	46.1	49.5	39.0
75% or more	25.1	30.1	25.0	29.9	18.8
90% or more	17.2	22.5	16.6	21.3	12.2
Weighted count (in '000)	19,755	17,800	19,786	18,670	18,650
Unweighted count	3,674	8,800	3,682	9,300	9,300
Women					
50% or more	53.1	55.2	55.1	55.2	43.8
75% or more	33.1	36.4	33.2	35.4	23.2
90% or more	23.0	28.0	22.6	26.1	14.9
Weighted count (in '000)	25,162	22,720	25,189	23,630	23,620
Unweighted count	5,313	11,200	5,344	11,700	11,700

Notes: Authors' calculations based on the 2016 wave of HRS and 2016 CPS ASEC, matched with administrative data from SSA and IRS. We refer to CPS ASEC as CPS. For HRS, the public data come from RAND-HRS public file 2016 version 3 (that includes data from 1992 to 2016). The HRS sample removes respondents in institutions to be consistent with the CPS sample. Estimates are based on all respondents with or without a match. Records that are not linkable use survey values throughout. Estimates are weighted using survey weights. SSA administrative data contain wage and salary earnings from Form W-2, self-employment earnings from Form 1040 Schedule SE, OASDI benefits, and SSI benefits. IRS administrative data contain DB pension income and DC withdrawals (excluding rollovers and conversions) from Form 1099-R and interest and dividends from Form 1040. Income reference year is 2015, the calendar year prior to the interview.

*In this table we restrict the sample to respondents in social security beneficiary families for whom family earnings and family asset income are non-negative, and family total income and family social security benefits are greater than zero.

†Reported estimates for CPS data are approved for release by the Census Bureau's Disclosure Review Board (CBDRB- FY24-SEHSD003-010). Analysis of the HRS data is performed at the HRS data enclave at the University of Michigan, and these estimates are approved for release by HRS's Disclosure Review Board in November 2023.

contrast, women are less likely than men to have family income from assets, pensions or retirement accounts, earnings, and veterans' benefits.

4.4 Reliance on social security

We next explore the consequences of income misreporting on social security's measured share of total income across families. Table 5 presents social security reliance statistics for each of the five data files. Paralleling *Income of the Aged* publications, we restrict the sample to persons in social security beneficiary families (i.e., at least one family member, respondent, or other, is a beneficiary), for whom family earnings and family asset income are non-negative, and family total income and family social security benefits are greater than zero.⁴⁴ We use three indicators of reliance. The first reliance threshold encompasses individuals for whom family social security income comprises 50 percent or more of their total family income. The other two reliance thresholds are 75 percent or more and 90 percent or more.

We observe that the CPS produces slightly higher reliance figures than the HRS across the board. For example, the proportion of all persons aged 65 or older relying on social security for 50 percent or more of their family income is 52.6 percent in the CPS compared with 49.1 percent in the HRS. It is worth noting that one reason why the HRS might show lower reliance figures than the CPS is that only

⁴⁴Note that in the HRS the family income measures include only the income received from the respondent (if single, widowed, divorced, or separated) and the spouse or partner (if coupled). Hence, the social security beneficiary can be either the respondent, or the spouse, or both. In the CPS, a beneficiary can be any residing member of the family.

the net amount of social security benefits are collected in the HRS, and the benefit amounts deducted for Medicare Parts B and D premiums are excluded (see Dushi *et al.*, 2017).

The four files without IRS administrative data all yield similar results at all reliance thresholds as well as when looking separately by gender. However, when examining the CPS + SSA + IRS file, reliance on social security is considerably lower. Most starkly, the proportion of persons aged 65 or older who rely on social security for 90 percent or more of their family income is 25.6 percent in the CPS and 20.5 percent in the HRS but only 13.7 percent in the CPS + SSA + IRS sample. The difference is mainly attributable to better reporting of retirement income in IRS administrative data. In sum, our findings suggest that income from social security is not as large a share of household budgets as previously thought.

We have also examined the robustness of these reliance patterns by limiting our analysis to the subset of beneficiaries that can be linked to administrative data. Following Bee and Mitchell (2017), we first run a logistic regression to model the probability of SSN assignment as a function of survey demographic, geographic, and economic characteristics and estimate the propensity score. We next construct new weights by multiplying the survey weights by the inverse of the estimated propensity score. We then compute the same statistics for the linked subsample using the newly constructed weights. In both surveys, we find that respondents are only 1–2 percentage points less likely to be reliant on social security than the overall sample, across all reliance indicators and gender. For example, in the CPS + SSA + IRS file, among respondents with a linkage, 40.4 percent of them rely on social security benefits for at least half their family income compared to 41.7 percent in the same file among all respondents (with or without linkages). Respective figures in the HRS + SSA file are 50.5 percent for the matched subsample compared to 51.1 percent for the whole sample.

The CPS reliance patterns for the linked subsample can also be directly compared to previous work by Bee and Mitchell to assess the impact of the survey redesign. Prior to the redesign, Bee and Mitchell found that in 2012, 55.5 percent of survey respondents relied on social security for 50 percent or more of their family incomes while the linked data showed a lower estimate of 42.2 percent. Post-redesign, we find that in 2015, the survey number is 52.5 percent while the linked number is 41.7 percent. Thus, the survey only showed a small decrease post-redesign, and the linked data continue to show a significantly lower reliance on social security.

4.5 Poverty rates

Finally, we examine differences in estimated poverty rates across the five files. Table 6 shows that the poverty rate for all persons aged 65 or older is 8.8 percent in the CPS, 7.7 percent in the CPS + SSA file, and only 6.4 percent in the CPS + SSA + IRS file. Interestingly, the HRS shows a lower poverty rate than the CPS at 7.4 percent, but the HRS + SSA shows a still lower poverty rate of 6.5 percent which is very close to the CPS + SSA + IRS estimate. In both the CPS and the HRS, the standalone surveys considerably overestimate the poverty rate relative to the linked administrative data, indicating income at the bottom of the distribution is underreported. The general finding that poverty rate is lower in the HRS than in the CPS, is not surprising given that we already demonstrated that family incomes are generally higher in the HRS in Table 3.⁴⁵

A closer look across the three CPS files shows that poverty rates in the CPS + SSA file are lower than those in the CPS file across all demographic subgroups. However, in the CPS + SSA + IRS file, poverty rates are substantially lower than those in the other two CPS files for nearly all race/ethnicity categories

⁴⁵It is worth noting that the lower poverty rates in the HRS than the CPS may be due to several factors that affect family income. First, as Table 1 shows, HRS respondents are more likely than CPS respondents to live in a family with two or more residing members and are more likely to be social security beneficiaries. Second, HRS respondents are more likely than their CPS counterparts to be in families with higher income throughout the distribution except for those in the 10th and 25th percentiles (Table 3, panel B, columns HRS and CPS). Both factors would plausibly lead to lower poverty rates in the HRS. Third, while the total family income variable in the RAND-HRS data file does not include the income of other residing family members, only 17 percent of HRS respondents aged 65 or older live in a family with three or more members and thus are affected by such exclusion (see Appendix B for a detailed discussion of the HRS total income and poverty measures).

Table 6. Percentage of individuals aged 65 or older in poverty or near poverty, by selected characteristics: estimates from five alternative data files, 2015

Characteristic	Survey data (unmatched)		Survey data matched with administrative records*		
	HRS	CPS	HRS + SSA	CPS + SSA	CPS + SSA + IRS
<i>In poverty (below poverty threshold)</i>					
All	7.4	8.8	6.5	7.7	6.4
<i>Gender</i>					
Men	5.1	7.0	4.4	6.2	5.1
Women	9.3	10.3	8.2	9.0	7.4
<i>Race/ethnicity</i>					
Non-Hispanic white	4.0	6.6	3.2	5.6	4.4
Non-Hispanic Black	21.2	18.2	18.8	14.4	11.8
Non-Hispanic other race(s) [†]	8.9	12.4	8.0	12.7	11.9
Hispanic (any race)	25.0	17.5	23.5	16.6	15.9
<i>Marital status</i>					
Married	3.1	4.8	2.5	3.8	2.7
Non-married	14.3	14.3	12.7	13.1	11.4
<i>Age</i>					
65–69	7.5	8.1	6.4	7.5	6.4
70–74	5.5	7.8	4.9	6.3	5.5
75–79	7.0	8.2	6.2	7.7	6.5
80 or older	9.8	11.4	8.4	9.5	7.1
<i>In or near poverty (below 125% of poverty threshold)</i>					
All	11.0	13.8	9.8	12.1	9.9
<i>Gender</i>					
Men	8.4	10.6	7.5	9.4	7.7
Women	13.1	16.4	11.7	14.3	11.6
<i>Race/ethnicity</i>					
Non-Hispanic white	6.6	10.8	5.6	9.2	7.0
Non-Hispanic Black	29.2	26.1	27.1	21.8	18.3
Non-Hispanic other race(s) [†]	10.6	18.2	2.3	18.7	16.7
Hispanic (any race)	33.6	25.7	31.4	24.1	22.9
<i>Marital status</i>					
Married	4.9	7.3	4.2	5.8	4.2
Non-married	20.7	22.6	18.7	20.7	17.7
<i>Age</i>					
65–69	10.3	11.9	9.3	11.0	9.4
70–74	8.9	12.1	8.0	10.1	8.6
75–79	10.6	13.6	9.3	11.8	9.7
80 or older	14.8	18.5	13.1	16.1	12.1
Weighted count (in '000)	48,258	47,550	48,258	47,550	47,550
Unweighted count	9,579	23,500	9,579	23,500	23,500

Notes: Authors' calculations based on the 2016 wave of HRS and 2016 CPS ASEC, matched with administrative data from SSA and IRS. We refer to CPS ASEC as CPS. For HRS, the public data come from RAND-HRS public file 2016 version 3 (that includes data from 1992 to 2016). The HRS sample removes respondents in institutions to be consistent with the CPS sample. Estimates are based on all respondents with or without a match. Records that are not linkable use survey values throughout. Estimates are weighted using survey weights. SSA administrative data contain wage and salary earnings from Form W-2, self-employment earnings from Form 1040 Schedule SE, OASDI benefits, and SSI benefits. IRS administrative data contain DB pension income and DC withdrawals (excluding rollovers and conversions) from Form 1099-R and interest and dividends from Form 1040. Poverty measures are based on family income in 2015 and the Census Bureau poverty thresholds corresponding to family size and composition. For the HRS, we use the poverty threshold provided in the RAND-HRS file. HRS total family income is similar to the Census Bureau definition as it includes income the other residing family members (see Appendix B for detailed discussion of poverty measures). To the HRS measure we add IRA withdrawals to be consistent with CPS definition. 'In poverty' = with income below 100% of the poverty threshold; 'near poverty' = with income below 125% of the poverty threshold.

*Reported estimates for CPS data are approved for release by the Census Bureau's Disclosure Review Board (CBDRB-FY24- SEHSD003-010). Analysis of the HRS data is performed at the HRS data enclave at the University of Michigan, and these estimates are approved for release by HRS's Disclosure Review Board in November 2023.

†For the CPS sample, this category includes respondents identified as 'Asian alone', other race, and those who report multi-race categories. However, the majority of CPS respondents in this category are Asian alone.

and age groups. While the overall poverty rate is almost the same in the HRS + SSA and CPS + SSA + IRS files, there are some differences across demographic subgroups. For example, the poverty rate is much higher among non-Hispanic Black and Hispanic elderly in the HRS + SSA file (18.8 and

23.5%, respectively) than in the CPS + SSA + IRS file (11.8 and 15.9%, respectively). In contrast, poverty rates in the two surveys are relatively similar across age, gender, and marital status categories.

We can once again compare our CPS results to those of Bee and Mitchell. For 2012, they reported a CPS poverty rate for 65 or over of 9.1 percent versus the linked data rate of 6.9 percent, whereas we find for 2015 a survey poverty rate of 8.8 percent versus the linked data rate of 6.4 percent. Evidently, the CPS redesign has had little effect on poverty estimates for those 65 or over.

5. Conclusions

Previous research on the aged has shown that survey-reported income, particularly retirement income, is subject to substantial underreporting, leading to downwardly biased estimates of financial well-being. In this paper, we examine whether previously documented income measurement issues are confined to the CPS ASEC or whether they also extend to the well-regarded HRS. Our findings show that total family income, particularly pension and asset income, continue to be significantly underreported in the redesigned CPS ASEC. While the HRS performs somewhat better than the CPS ASEC, we find that it still produces lower estimates of retirement income than those generated by the CPS ASEC linked with administrative data.

We also find that according to the CPS ASEC the primary source of aggregate income among the aged is social security, while the administrative records indicate that retirement income is the largest source of aggregate income. Furthermore, social security reliance indicators are also upwardly biased. Thus, the proportion of people aged 65 and over who rely on social security for at least half their family income decreases from 52.5 percent in the CPS ASEC to 41.7 percent when using administrative data. Similarly, the proportion of the elderly who rely on social security for 90 percent or more of their family income is considerably lower in the administrative data (13.7%) compared to the survey (25.6%). For comparison, 20.5 percent of elderly in the HRS rely on social security for 90 percent or more of their family income, which is lower than the CPS ASEC but still well above the CPS + SSA + IRS file.

Finally, mismeasurement of income leads to an upwardly biased estimate of the poverty rate for those 65 and over. Thus, compared to the CPS ASEC, supplementing the CPS with SSA and IRS administrative data resulted in a reduction in the estimated poverty rate among people aged 65 or older from 8.8 to 6.4 percent. In addition, the HRS indicates somewhat lower estimates of poverty rate than the CPS (7.4% in the survey data and 6.5% in the HRS + SSA). These lower rates are in part due to HRS elderly being less likely to live alone and reporting higher family incomes.

Based on these findings, we conclude that declining data quality is not confined to the CPS ASEC, and surveys linked to administrative records remain the most reliable sources for information about the performance of the U.S. retirement system. Therefore, it is our hope that decision makers at both the SSA and the Census Bureau will facilitate the regular usage of linked survey-administrative data for the continued production of the *Income of the Aged* publications.

Acknowledgments. Any opinions and conclusions expressed herein are those of the authors and do not reflect the views of the Social Security Administration or the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product (Disclosure Review Board [DRB] approval number CBDRB-FY24-SEHSD003-010). Similarly, the HRS has reviewed and approved for release the data product provided in this paper. We thank the Special Issue Editor and three anonymous referees for helpful comments.

References

- Abowd JM and Stinson MH (2013) Estimating measurement error in annual job earnings: a comparison of survey and administrative data. *Review of Economics and Statistics* 95(5), 1451–1467.
- Alwin DF, Zeiser K and Gensimore D (2014) Reliability of self-reports of financial data in surveys: results from the health and retirement study. *Sociological Methods and Research* 43(1), 98–136. <https://doi.org/10.1177/0049124113507908>
- Anguelov CE, Iams HM and Purcell PJ (2012) Shifting income sources of the aged. *Social Security Bulletin* 72(3), 59–68.
- Argento R, Bryant VL and Sabelhaus J (2015) Early withdrawals from retirement accounts during the great recession. *Contemporary Economic Policy* 33(1), 1–16.

- Bee CA** (2013) An evaluation of retirement income in the CPS ASEC using Form 1099-R Microdata. Working paper prepared for the 2014 Meetings of the Society of Labor Economists. Available at <http://www.sole-jole.org/14467.pdf>
- Bee CA and Mitchell JW** (2017) Do older Americans have more income than we think? SEHSD Working Paper No. 2017-39. Washington, DC: Census Bureau, Social, Economic, and Housing Statistics Division. Available at <https://www.census.gov/library/working-papers/2017/demo/SEHSD-WP2017-39.html>
- Bee CA and Mitchell JW** (2021) Do older Americans still have more income than we think? Evidence from the CPS ASEC Redesign. Working paper presented at American Economic Association – Society of Government Economists Annual Meeting, January 3.
- Bee CA, Mitchell J, Mittag N, Rothbaum J, Sanders C, Schmidt L and Unrath M** (2023) National experimental well-being statistics. U.S. Census Bureau Working Paper No. CES-23-04. Available at <https://www.census.gov/data/experimental-data-products/national-experimental-wellbeing-statistics.html>
- Bollinger CR** (1998) Measurement error in the current population survey: a nonparametric look. *Journal of Labor Economics* 16(3), 576–594.
- Bollinger CR and David MH** (2001) Estimation with response error and nonresponse: food-stamp participation in the SIPP. *Journal of Business & Economic Statistics* 19(2), 129–141.
- Bollinger CR, Hirsch BT, Hokayem CM and Ziliak JP** (2019) Trouble in the tails? What we know about earnings nonresponse 30 years after Lillard, Smith, and Welch. *Journal of Political Economy* 127(5), 2143–2185. <https://doi.org/10.1086/701807>
- Bosworth BP, Burtless G and Anders SE** (2007) Capital income flows and the relative well-being of American's aged population. Working Paper No. 2007-21. Chestnut Hill, MA: Center for Retirement Research at Boston College.
- Bound J, Brown C and Mathiowetz N** (2001) Measurement error in survey data. In Heckman JJ and Leamer E (eds), *Handbook of Econometrics*, Volume 5. Amsterdam: Elsevier Science BV, pp. 3705–3843.
- Brady PJ and Bass S** (2021) Comparing the current population survey to income tax data. Statistics of Income Working Paper. Washington, DC: Internal Revenue Service. Available at <https://www.irs.gov/pub/irs-soi/21r1pcomparingcpstoincomeaxdata.pdf>
- Brady PJ and Pierce K** (2012) The promise and potential pitfalls of using administrative tax data: A case study. Unpublished paper, Investment Company Institute (April).
- Bricker J and Engelhardt GV** (2007) Measurement error in earnings data in the Health and Retirement Study. Working Paper No. 2007-16. Chestnut Hill, MA: Center for Retirement Research at Boston College.
- Bugliari D, Carroll J, Hayden O, Hayes J, Hurd M, Karabatakis A, Main R, Marks J, McCullough C, Meijer E, Moldoff M, Pantoja P, Rohwedder S and Clair PS** (2020) *RAND HRS Longitudinal File 2016 (V2) Documentation*. Santa Monica, CA: RAND Center for the Study of Aging. April.
- Celhay P, Meyer BD and Mittag N** (2022) Stigma in Welfare Programs. NBER Working Paper No. 30307. doi: 10.3386/w30307
- Celhay P, Meyer BD and Mittag N** (2024) What leads to measurement errors? Evidence from reports of program participation in three surveys. *Journal of Econometrics* 238(2), 1–19. <https://doi.org/10.1016/j.jeconom.2023.105581>
- Chen A, Munnell AH and Sanzenbacher GT** (2018) How much income do retirees actually have? Evaluating the evidence from five national datasets. Working Paper No. 2018-14. Chestnut Hill, MA: Center for Retirement Research at Boston College.
- Corinth K, Meyer BD and Wu D** (2022) The change in poverty from 1995 to 2016 among single-parent families. *American Economic Association Papers and Proceedings* 112, 345–350.
- Czajka JL and Denmead G** (2008) *Income data for policy analysis: A comparative assessment of eight surveys. Final Report*. Washington, DC: Mathematica Policy Research.
- Davies PS and Fisher TL** (2009) Measurement issues associated with using survey data matched with administrative data from the social security administration. *Social Security Bulletin* 69(2), 1–12.
- Dushi I and Honig M** (2015) How much do respondents in the health and retirement study know about their contributions to tax-deferred contribution plans? A cross-cohort comparison. *Journal of Pension Economics and Finance* 14(3), 203–239. doi: 10.1017/S1474747214000237
- Dushi I and Iams HM** (2010) The impact of response error on participation rates and contributions to defined contribution pension plans. *Social Security Bulletin* 70(1), 45–60.
- Dushi I and Iams HM** (2017) Reporting accuracy of social security benefits and its implications in the health and retirement study. *Journal of Economics and Social Measurement* 42, 271–297. <https://doi.org/10.3233/JEM-180449>
- Dushi I and Trenkamp B** (2021) Improving the measurement of retirement income of the aged population. Office of Research, Evaluation, and Statistics Working Paper No. 116.
- Dushi I, Iams HM and Trenkamp B** (2017) The importance of social security benefits to the income of the aged population. *Social Security Bulletin* 77(2), 1–12.
- Fisher TL** (2008) The impact of survey choice on measuring the relative importance of social security benefits to the elderly. *Social Security Bulletin* 67(2), 55–64.
- Gustman AL, Steinmeier TL and Tabatabai N** (2010) *Pensions in the Health and Retirement Study*. Cambridge, MA: Harvard University Press.

- Gustman AL, Steinmeier TL, Thomas L and Tabatabai N** (2014) Mismeasurement of pensions before and after retirement: the mystery of the disappearing pensions with implications for the importance of social security as a source of retirement support. *Journal of Pension Economics and Finance* **13**(1), 1–26. <https://doi.org/10.1017/S147474721300017>
- Hurd M, Thomas Juster F and Smith JP** (2003) Enhancing the quality of data on income: recent innovations from the HRS. *Journal of Human Resources* **38**(3), 758–772. <https://doi.org/10.2307/1558775>
- Hyde JS and Harrati A** (2023) The alignment between self-reported and administrative measures of disability program application and benefit receipt in the health and retirement study. *Social Security Bulletin* **83**(4), 19–51.
- Iams HM and Purcell PJ** (2013) The impact of retirement account distributions on measures of family income. *Social Security Bulletin* **73**(2), 77–84.
- Johnson BW and Moore K** (2008) Differences in income estimates derived from survey and tax data. Statistics of Income Working Paper. Washington, DC: Internal Revenue Service. Available at <https://www.irs.gov/pub/irs-soi/08rjjohnson.pdf>
- Koenig ML** (2003) An assessment of the Current Population Survey and the Survey of Income and Program Participation using Social Security Administrative data. Paper presented at the 2003 Federal Committee on Statistical Methodology Research Conference, Arlington, VA, November 17–19.
- Koenig M and Rupp K** (2004) SSI recipients in households and families with multiple recipients: prevalence and poverty outcomes. *Social Security Bulletin* **65**(2), 14–27.
- Meyer BD and Mittag N** (2019) Using linked survey and administrative data to better measure income: implications for poverty, program effectiveness, and holes in the safety net. *American Economic Journal: Applied Economics* **11**(2), 176–204. <https://doi.org/10.1257/app.20170478>
- Meyer BD and Mittag N** (2021) An empirical total survey error decomposition using data combination. *Journal of Econometrics* **224**(2), 286–305. <https://doi.org/10.1016/j.jeconom.2020.03.026>
- Meyer BD and Wu D** (2018) The poverty reduction of social security and means-tested transfers. *Industrial and Labor Relations Review* **71**(5), 1106–1153.
- Meyer BD, Mok WKC and Sullivan JX** (2015) The under-reporting of transfers in household surveys: Its nature and consequences. Chicago, IL: University of Chicago. Available at <https://harris.uchicago.edu/files/underreporting.pdf>
- Meyer BD, Wu D, Moors V and Medalia C** (2021) The use and misuse of income data and extreme poverty in the United States. *Journal of Labor Economics* **39**(S1), S5–S58. <https://doi.org/10.1086/711227>
- Miller BJ and Schieber SJ** (2013) Employer plans, IRAs, and retirement income provision: Making a Molehill out of a mountain. *Insider* (October). London: Willis Towers Watson.
- Mitchell J and Renwick T** (2015) A comparison of official poverty estimates in the redesigned Current Population Survey Annual Social and Economic Supplement, SEHSD Working Paper No. 2014–35, January. Available at <https://www.census.gov/content/dam/Census/library/working-papers/2014/demo/SEHSD-WP2014-35.pdf>
- Mitchell OS, Clark R and Lusardi A** (2022) Income trajectories in later life: longitudinal evidence from the health and retirement study. *Journal of the Economics of Aging* **22**. <https://doi.org/10.1016/j.jeoa.2022.100371>
- Munnell AH and Chen A** (2014) Do Census Data understate retirement income? Issue in Brief No. 14–19. Chestnut Hill, MA: Center for Retirement Research at Boston College.
- Nicholas J and Wiseman M** (2009) Elderly poverty and supplemental security income. *Social Security Bulletin* **69**(1), 45–73. <https://www.ssa.gov/policy/docs/ssb/v69n1/v69n1p45.html>
- Nicholas J and Wiseman M** (2010) Elderly poverty and supplemental security income, 2002–2005. *Social Security Bulletin* **70**(2), 1–29. <https://www.ssa.gov/policy/docs/ssb/v70n2/v70n2p1.html>
- O'Hara A, Bee A and Mitchell J** (2016) Preliminary research for replacing or supplementing the income question on the American Community Survey with Administrative Records. American Community Survey Research and Evaluation Report Memorandum Series No. ACS16-RER-6, U.S. Census Bureau.
- Pedace R and Bates N** (2000) Using administrative records to assess earnings reporting error in the survey of income and program participation. *Journal of Economic and Social Measurement* **26**(3–4), 173–192.
- Rector RE, Johnson KA and Youssef SE** (1999) The extent of material hardship and poverty in the United States. *Review of Social Economy* **57**(3), 351–387. <http://www.jstor.org/stable/29770022>
- Rodgers WL, Brown C and Duncan GJ** (1993) Errors in survey reports of earnings, hours worked, and hourly wages. *Journal of the American Statistical Association* **88**(424), 1208–1218.
- Roemer M** (2000) Assessing the quality of the March Current Population Survey and the Survey of Income and Program Participation Income Estimates, 1990–1996. Staff Papers on Income, Housing and Household Economic Statistics Division. Washington, DC: Census Bureau.
- Roemer M** (2002) Using Administrative Earnings Records to Assess Wage Data Quality in the March Current Population Survey and the Survey of Income and Program Participation. Technical Paper TP-2002-22, LEHD. Washington, DC: Census Bureau.
- Rothbaum J** (2015) Comparing income aggregates: How do the CPS and ACS match the national income and product accounts, 2007–2012. U.S. Census Bureau SEHSD Working Paper No. 2015-01.
- Ruser J, Pilot A and Nelson C** (2004) Alternative measures of household income: BEA personal income, CPS money income, and beyond. Federal Economic Statistics Advisory Committee Working Paper, November. Available at <https://www.bea.gov/about/pdf/AlternativemeasuresHHincomeFESAC121404.pdf>

- Sabelhaus J and Schrass D** (2009) The evolving role of IRAs in U.S. retirement planning. *Investment Company Institute Research Perspective* 15(3), 1–32. <http://www.ici.org/per15-03.pdf>
- Schieber SJ** (1995) Why do pension benefits seem so small? *Benefits Quarterly* 11(4), 57–70.
- Sears J and Rupp K** (2003) Exploring social security payment history matched with the survey of income and program participation. Paper presented at the Federal Committee on Statistical Methodology Research Conference, Washington, DC, November 18.
- Semega JL and Welniak E Jr** (2015) The effects of changes to the Current Population Survey Annual Social and Economic Supplement on Estimates of Income. Paper presented at the 2015 Allied Social Science Association Meetings. Available at <https://www.census.gov/content/dam/Census/library/working-papers/2015/DEMO/ASSA-Income-CPSASEC-Red.pdf>
- Social Security Administration** (2022) *Annual Statistical Supplement to the Social Security Bulletin*. Available at <https://www.ssa.gov/policy/docs/statcomps/supplement/2022/supplement22.pdf>, 1–492.
- U.S. Census Bureau** (1993) Money income of households, families, and persons in the United States: 1992. Current Population Reports Series P60-184. Washington, DC: U.S. Government Printing Office.
- U.S. Census Bureau** (2019) Current population survey: Design and methodology, Technical Paper 77, Washington, DC, October. Available at <https://www.census.gov/programs-surveys/cps/methodology/CPS-Tech-Paper-77.pdf>
- U.S. Census Bureau** (2020) Current population survey 2020 Annual Social and Economic (ASEC) Supplement Technical Documentation. Washington, DC: Census Bureau. Available at <https://www.census.gov/data/datasets/time-series/demo/cps/cps-asec.html>
- Wagner D and Layne M** (2014) The Person Identification Validation System (PVS): Applying the Center for Administrative Records and Research and Applications' (CARRA) Record Linkage Software. U.S. Census Bureau Center for Administrative Records Research and Applications Report Series No. 2014-01.
- Woods JR** (1996) Pension benefits among the aged: conflicting measures, unequal distributions. *Social Security Bulletin* 59(3), 3–30.

Appendix A: Definitions of income and poverty in CPS

Income

- **Earnings:** Includes the following:
 - **Wages and salaries:** Money wages or salary is defined as total money earnings received for work performed as an employee during the income year. It includes wages, salary, Armed Forces pay, commissions, tips, piece-rate payments, and cash bonuses earned, before deductions are made for taxes, bonds, pensions, union dues, and so forth. Earnings for self-employed persons in incorporated businesses are considered wage and salary.
 - **Self-employment:** Income from self-employment is the combined income from farm and non-farm self-employment. Farm self-employment is net money income (gross receipts minus operating expenses) from the operation of a farm by a person on their own account, as an owner, as a renter, or as a sharecropper. Non-farm self-employment is net money income (gross receipts minus expenses) from one's own business, professional enterprise, or partnership.
- **Asset income:** Includes the following:
 - **Interest:** Interest includes payments people receive (or have credited to their accounts) from bonds, treasury notes, IRAs, certificates of deposit, interest-bearing savings and checking accounts, and all other investments that pay interest.
 - **Dividends:** Dividends include income people receive from stock holdings and mutual fund shares. The CPS does not include capital gains from the sale of stock holdings as income.
 - **Rents, royalties, and estates and trusts:** Include net income people receive from the rental of a house, store, or other property, receipts from boarders or lodgers, net royalty income, and periodic payments from estate or trust funds.
- **Retirement benefits:** It is the sum of social security benefits and public and private pensions, as well as private survivor and disability payments.
 - **Social security:** Social security includes retired-worker benefits, dependent or survivor benefits, and disability benefits made by SSA prior to deductions for medical insurance and railroad retirement insurance checks from the U.S. Government. Medicare reimbursements are not included.

- **Pensions:** Includes the following:
 - **Employer pensions:** Employer pensions include pensions from Railroad Retirement, government employee pensions, and private pensions and annuities.
 - **Government employee pensions:** Government employee pensions include payments from federal government (civil service), military, and state or local governments.
 - **Private pensions and annuities:** Private pensions and annuities include payments from companies or unions, annuities or paid-up insurance policies, IRAs, Keogh, or 401(k) payments.
 - **Private survivor and disability payments:** These are included with other private retirement payments to follow the payment reported in Form 1099-R. Survivor benefits include widow's pensions, estates, trusts, annuities, etc. Disability income is income received as a result of health problems (other than social security disability benefits), such as worker's compensation, disability insurance, and black lung payments.
- **Cash public assistance:** Includes the following:
 - **Supplemental security income:** Includes federal, state, and local welfare agency payments to low-income people who are 65 years old or older, or people of any age who are blind or disabled.
 - **Other public assistance:** Includes cash public assistance payments low-income people receive, such as temporary assistance to needy families, general assistance, and emergency assistance.
- **Other income:** It is the total income minus earnings, social security, pensions, asset income, and cash public assistance; included are unemployment compensation, veterans' payments, and personal contributions.
 - **Unemployment compensation:** Includes payments the respondent received from government unemployment agencies or private companies during periods of unemployment and any strike benefits the respondent received from union funds.
 - **Workers' compensation:** Includes payments people receive periodically from public or private insurance companies for injuries received at work.
 - **Veterans' payments:** Include payments disabled members of the armed forces or survivors of deceased veterans receive periodically from the Department of Veterans Affairs for education and on-the-job training and means-tested assistance to veterans.
 - **Personal contributions:** Include child support, alimony, and financial assistance from friends and relatives.

For additional details on income definitions in the CPS ASEC, see Census Bureau (2020, 7.3–7.5).

Poverty rate

Following the Office of Management and Budget's (OMB's) Directive 14, the Census Bureau uses a set of money income thresholds that vary by family size and composition to determine who is poor. If a family's total income is less than that family's threshold, then that family, and every individual in it, is considered poor. The poverty thresholds do not vary geographically, but they are updated annually for inflation with the consumer price index (CPI). The official poverty definition counts money income before taxes and thus excludes capital gains and non-cash benefits (such as public housing, Medicaid, and Supplemental Nutrition Assistance Program benefits).

Poverty statistics are based on a definition developed by SSA's Mollie Orshansky in 1964 and revised in 1969 and 1981 by interagency committees. This definition was established as the official definition of poverty for statistical use in all Executive departments in 1969 (in Bureau of the Budget Circular No. A-46) and was reconfirmed in OMB Statistical Policy Directive No. 14. For further details, see the section 'Changes in the Definition of Poverty' in Census Bureau (1982).

Table A1. Demographic characteristics of survey respondents aged 65 or older, by match status to administrative data, in 2016 (in %)

Characteristic	HRS			CPS		
	All	Match status		All	Match status*	
		Yes	No		Yes	No
<i>Gender</i>						
Men	44.4	44.2	45.9	44.6	44.8	42.9
Women	55.6	55.8	54.1	55.4	55.2	57.1
<i>Race/ethnicity</i>						
Non-Hispanic white	79.9	81.3	71.2	77.2	78.2	67.3
Non-Hispanic Black	9.3	8.7	12.8	8.9	8.6	11.2
Non-Hispanic other [†]	2.7	2.5	4.1	5.9	5.4	9.9
Hispanic (any race)	8.1	7.5	11.9	8.1	7.7	11.6
<i>Marital status</i>						
Married	61.1	60.7	63.5	57.7	58.0	55.2
Non-married	39.9	39.3	36.5	42.3	42.0	44.8
Widowed	21.7	22.3	18.1	24.0	24.2	23.0
Divorced/separated	12.4	12.3	12.7	13.0	13.2	11.8
Never married	4.8	4.7	5.7	5.2	4.7	10.0
<i>Educational attainment</i>						
Less than high school diploma	14.8	14.2	18.9	14.6	14.4	16.8
High school diploma or equivalent	32.9	33.5	29.1	33.2	33.1	34.5
Some college	24.3	24.2	24.6	24.1	24.3	22.1
College degree	28.0	28.1	27.5	28.1	28.3	26.5
<i>Persons in family[‡]</i>						
One	28.3	28.6	26.2	32.1	32.0	32.7
Two	54.6	54.2	57.2	52.0	53.1	41.4
Three or more	17.2	17.2	16.6	15.9	14.9	25.8
<i>Social security beneficiaries[§]</i>	90.3	91.0	86.1	82.1	84.1	63.8
<i>Poverty rate (using survey measures)</i>	7.7	7.2	10.9	8.8	8.4	13.1
<i>Total person income including IRA withdrawals (survey reported)</i>						
Mean	60,810	60,293	64,070	44,160	44,530	40,770
Weighted count (in '000)	48,258	41,648	6,610	47,550	42,870	4,680
Unweighted count	9,579	8,309	1,270	23,500	21,500	2,400

Notes: Authors' calculations based on the 2016 wave of HRS and 2016 CPS ASEC, matched with administrative data from SSA and IRS. We refer to CPS ASEC as CPS. For HRS, the public data come from RAND-HRS public file 2016 version 3 (that includes data from 1992 to 2016). The HRS sample removes respondents in institutions to be consistent with the CPS sample. Estimates are weighted using survey weights.

*In the CPS, those shown with a match (or linkage) are those for whom there is a Personal Identification Key (PIK) identifier. Reported estimates for CPS data are approved for release by the Census Bureau's Disclosure Review Board (CBDRB- FY24-SEHSD003-010). Analysis of the HRS data is performed at the HRS data enclave at the University of Michigan, and these estimates are approved for release by HRS's Disclosure Review Board in November 2023.

†For the CPS sample, this category includes respondents identified as 'Asian alone', other race, and those who report multi-race categories. However, the majority of CPS respondents in this category are Asian alone.

‡We derive these categories using the variable 'h13hhres' in the RAND-HRS data file, which indicates the number of residing family members in the household. There is another variable in this file, named the 'family composition' variable ('h13povfam') that was created in order to define poverty thresholds and the poverty rate. We used this latter variable in an earlier version of the paper, but after discussion with RAND researchers we decided to use the former variable instead to produce this table. It is worth noting that the overall pattern is similar regardless of which variable is used.

§These are survey respondents who report receiving social security OASDI benefits.

Table A2. Match rates across different administrative data files, among HRS respondents aged 65+, in 2016

	All*		With a positive amount [†]	
	N	Match rate (as % of total)	N	Match rate (as % of total)
Total sample	9,579	100	6,846	71.5
With a matched record in the administrative files:				
Benefit records (PHUS)	6,804	71.0	6,314	65.9
SSI records (SSR)	1,148	12.0	52	0.5
Summary earnings records (MEF)	6,521	68.1	1,117	11.7
Detailed earnings records (DER)	1,196	12.5	1,185	12.4
Matched with at least in one of the above files	8,309	86.7	6,600	68.9

Source: Authors' calculations based on the 2016 wave of HRS, using RAND-HRS public data file version 3 (including data from 1992 to 2016), matched with the SSA administrative records. Sample excludes respondents in institutions. Reported number of observations and match rates are unweighted. Analysis of the HRS data is performed at the HRS data enclave at the University of Michigan, and these estimates are approved for release by HRS's Disclosure Review Board in November 2023.

*The overall sample consists of respondents aged 65 and over in the 2016 wave of the HRS. The samples corresponding to each data file are respondents with a valid consent who have a non-missing administrative record for income year 2015. Match rates are proportions of the sample out of the overall sample ($N = 9,579$).

†These samples are comprised of respondents with a matched record for whom the income amount reported in the administrative record for year 2015 is greater than zero. Match rates are proportions of the sample out of the overall sample ($N = 9,579$).

Table A3. Reporting accuracy among HRS respondents aged 65 and over with matched records for income year 2015

	Income category (restricted data file)		
	Social security benefits (PHUS)	SSI (SSR)	Earnings (DER)
<i>Income reporting accuracy (%)</i>			
1 - True negative	7.8	76.4	0.4
2 - False positive	1.2	2.9	0
3 - False negative	2	5.5	17
4 - True positive	89	15.3	82.6
Total HRS (2 + 4)	90.2	18.2	82.6
Total Admin (3 + 4)	91	20.7	99.6
% False negative 3/(3 + 4)	2.2	26.6	17
Weighted count (in '000)	34,775	5,001	7,378
Unweighted count	6,804	1,148	1,196
<i>Mean amount of income (\$) among the overall sample</i>			
1 - Survey reported	16,407	225	16,976
Weighted count (in '000)	48,258	48,258	48,258
Unweighted count	9,579	9,579	9,579
<i>Mean amount of income (\$) among those with a match</i>			
1 - Survey reported (matched sample)	16,604	1,073	57,934
2 - Admin records (matched sample)	17,644	1,146	41,181
3 - \$ difference (1 - 2)	-1,041	-73	16,753
4 - % of admin (1/2)	0.94	0.94	1.41
Weighted count (in '000)	34,775	5,001	7,378
Unweighted count	6,804	1,148	1,196
<i>Mean amount of income among those with true positive reports</i>			
1 - Survey report (matched sample)	18,291	5,732	70,035
2 - Admin record (matched sample)	19,446	5,740	46,270
3 - \$ difference (1 - 2)	-1,154	-9	23,765
4 - % of admin (1/2)	0.94	1	1.51
Weighted count (in '000)	30,965	764	6,095
Unweighted count	6,160	196	931

Notes: Authors' calculations based on the 2016 wave of HRS, matched with SSA administrative records. Overall sample is comprised of all respondents with or without a match. Sample excludes respondents in institutions and thus it is consistent with the CPS sample. Income reference year is 2015, the calendar year prior to the interview. Estimates are weighted using survey weights. Dollar amounts are in 2022 dollars, inflation adjusted using CPI-U. The HRS public data come from RAND-HRS public file 2016 version 3, which includes data from 1992 to 2016. Analysis of the HRS restricted data is performed at the HRS data enclave at the University of Michigan, and these estimates are approved for release by HRS's Disclosure Review Board in November 2023.

Appendix B: Definitions of HRS income variables and poverty rate

Earnings (individual-level variable)

For each HRS respondent, total survey-reported earnings equal the sum of reported wages, self-employment income, and business and farm income. For couples, the spouse's earnings, defined in the same way, are also included. In the RAND-HRS file, self-employment income and household business and farm income are included in household capital income. Therefore, to be consistent with CPS, we subtract self-employment income and household business and farm income from the capital income category and add them to the earnings category.

Social security benefits (individual-level variable)

For each respondent, the self-reported amount of social security benefits is defined as the sum of retired-worker benefits, dependent or survivor benefits, and disability benefits. For married couples, the spouse's social security benefits (if any), defined the same way, are also included. Thus, the household's total social security benefits variable is the sum of benefits received by both respondent and spouse.

Asset income (household-level variable)

Asset income in the HRS is the household capital income, which aggregates several other variables reported in the survey. It includes business or farm income, self-employment earnings, business income, gross rent, dividend and interest income, trust funds and royalties, and other asset income. To be consistent with CPS definition, as noted above, we subtract business or farm income and self-employment earnings from the asset income variable and include them instead in the earnings category. For couples, the amount for this variable is divided by two and assigned to the total income for each spouse.

Cash public assistance (individual-level variable)

In the RAND-HRS public data, income from public programs is an aggregate variable called government transfers, which is the sum of self-reported amounts of veterans' benefits, welfare, and food stamps. For this study, to make our 'cash public assistance' variable (Table 4) consistent with the CPS data, we subtract the amount of veterans' benefits and add the amount of self-reported SSI payments received. The HRS government transfer variable and its components are available separately for the respondent and for the spouse of a married respondent. We create a household-level variable, which is equal to either the respondent's amount (if not married) or the sum of respondent's and spouse's amounts (if married).

Other income (household-level variable)

In the HRS data, 'other income' includes alimony; lump sums from insurance, trust funds, and inheritances; and income from miscellaneous other sources. By contrast, in the CPS data, the variable 'other income' includes unemployment and workers' compensation, veterans' benefits, personal contributions (such as child support, alimony, and financial assistance), and income from miscellaneous other sources. To make the HRS and CPS variables consistent, we define 'other income' as the combined household amounts of veterans' benefits; unemployment and workers' compensation; alimony; lump sums from insurance, pension, and inheritance; and income from miscellaneous other sources.

Pension income (individual-level variable)

In the HRS, the pension income variable includes self-reported regular income received from all pensions and annuities; if the respondent is married, the spouse's pension income is similarly defined. While the HRS question asks about different types of pensions (such as, veterans' benefits, retirement or other pensions, annuities, IRA distributions, stocks and bonds, and other), this pension income variable is created by RAND and does not include veterans' benefits.⁴⁶ The pension income variable also omits withdrawals or distributions from IRA accounts. However, a separate variable is available in the RAND-HRS data file and is called 'IRA withdrawals in the last calendar year'. Hence, we create a household-level variable, which is the sum of IRA withdrawals and income from pensions and annuities received by the respondent and, if married, also by the spouse.

Total household income (respondent and spouse only)

In the HRS, total household income is calculated as the sum of the respondent's and the spouse's earnings, pensions and annuities, SSI payments, social security disability and retirement benefits, unemployment and workers' compensation, other government transfers, household capital income, and other income. This is the variable we use in Tables 3–5. For

⁴⁶It is worth remembering that for the sake of comparability with CPS, we subtracted veterans' benefits from the HRS government transfers variable and added them to other income.

estimating the poverty rate in Table 6, we use the family income variable created by RAND for calculating poverty rate (see below). The difference between total household income and family income is that the latter includes the amount deducted from social security benefits for Medicare Part B and/or Part D premiums and it excludes non-cash benefits (such as food stamps) and capital gains and losses. Therefore, it is likely that using the latter measure may result in a lower poverty rate than using the total household income measure. However, it is also worth noting that only 16 percent of the HRS sample aged 65 or older live in a family with three or more members (Table 1).

Poverty rate

According to the RAND-HRS data documentation, HRS poverty measures follow the methods and definitions that the Census Bureau applies to CPS data to derive the national poverty rate. The poverty threshold that applies to an HRS family is determined by using poverty threshold levels defined annually by the Census Bureau for each family composition type. The two key variables for applying these methods to HRS families are income and family composition.

Family composition is determined by the number of resident family members, the number of those aged under 18, and the age of the head of household in one- or two-member households. People living in institutions, such as nursing homes and college dormitories, are not included when counting resident family members.

Family income includes before-tax incomes from earnings, unemployment insurance, and worker's compensation; SSI, public assistance, and veterans' benefits; social security income before deductions⁴⁷; pension and retirement income; interest, dividends, rents, royalties, and income from estates and trusts; education assistance; alimony and child support; assistance from outside the household; other sources; and income of all resident family members. Income does not include non-cash benefits such as Supplemental Nutrition Assistance Program benefits (food stamps) and capital gains and losses.

Education assistance and other sources are assumed to have been reported as 'other income' in the HRS, but it is likely that at least some assistance from outside the household may not be included in any of the HRS income categories. The HRS total household income – excluding food stamps and including Medicare Part B and/or Part D premiums deducted from social security – would seem to be close to the Census definition of income, except for income from resident family members besides the respondent and spouse. Survey questions ask about the income of resident family members, including the earnings of each and the total non-job income of them all. With these questions, the income of all resident family members can be estimated but is not included in the total household income. More specifically, total household income, for poverty calculation purposes, is equal to:

$$\begin{aligned}
 & \text{(Total household income – food stamps)} \\
 & + \text{(Medicare Part B and/or Part D premiums in instances when the} \\
 & \text{respondent had deducted these amounts from reported social security benefits)} \\
 & + \text{(income of non-core resident family members)} \\
 & - \text{(income of any core HRS nursing home residents, including earnings,} \\
 & \text{pensions, social security, SSI, unemployment and workers'} \\
 & \text{compensation, and government transfer income)}
 \end{aligned}$$

Family composition is defined based on household members reported at the time of the interview. Then, after the income and poverty threshold are determined as described above, the HRS family income is compared with the appropriate poverty threshold for the last calendar year. If household income for the last calendar year is below the poverty threshold, then the household is defined as being in poverty.⁴⁸

Another variable available in the RAND-HRS file is the ratio of household income to the poverty threshold. We use this variable to construct the poverty measures used in this study. If the ratio of household income to the poverty threshold is equal to or less than 1 (at or below 100% of the poverty threshold) then the respondent is defined as being in poverty. If the ratio of household income to the poverty threshold is equal to or less than 1.25 (at or below 125% of the poverty threshold) then the respondent is defined as being in or near poverty.

⁴⁷Medicare Part B and/or Part D premiums are added if the respondent reports that they were deducted from social security payments.

⁴⁸Note that the terms household income and family income are used interchangeably in the HRS.