ORES Working Paper Series
No. 116

Improving the Measurement of Retirement Income of the Aged Population

Irena Dushi and Brad Trenkamp

Social Security Administration
Office of Retirement and Disability Policy
Office of Research, Evaluation, and Statistics

ORES Working Paper Series

Number 116

Improving the Measurement of Retirement Income of the Aged Population

Irena Dushi and Brad Trenkamp

January 2021

Social Security Administration
Office of Retirement and Disability Policy
Office of Research, Evaluation, and Statistics

The authors are with the Office of Research, Office of Research, Evaluation, and Statistics, Office of Retirement and Disability Policy, Social Security Administration.

Abstract

Research has shown that survey-reported income measures, particularly pension and retirement income, suffer from reporting errors, which lead to biased estimates of income and poverty of the aged population. Two of the Social Security Administration's main publications—*Income of the Population 55 or Older* and the *Income of the Aged Chartbook*—are published biennially and are based exclusively on publicly available data from the U.S. Census Bureau. In this paper, we use data from the Census' 2016 Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC) merged with administrative data—Internal Revenue Service (IRS) tax records and Social Security earnings and benefit records—to examine whether and to what extent using these additional data improves income estimates. We also compare those estimates with public-use data from the 2016 Health and Retirement Study (HRS), which has the reputation of being a reliable source of income measures for the aged population. We find that for the population aged 65 or older, supplementing the CPS ASEC with IRS and Social Security administrative data results in a higher estimate of pension income's share of aggregate income, less estimated reliance on Social Security, and a lower estimated rate of poverty. Furthermore, we find that the HRS provides better estimates of the income of the aged population than the public-use CPS data.

Acknowledgments: We received helpful comments from Lynn Fisher, Bruce Meyer, Joshua Mitchell, Kathleen Romig, Julie Topoleski, and Robert Weathers. We also thank Adam Bee and Joshua Mitchell for their guidance.

Introduction

Measuring the income of the U.S. population overall and of the aged in particular is an important issue for both researchers and policymakers. Accurate measurement of retirement income in national surveys is a challenge. The U.S. Census Bureau's Current Population Survey (CPS) is no exception, and recent research has reinvigorated a debate about the adequacy of its retirement income measures. This has been of unique interest to the Social Security Administration (SSA), which publishes a biennial statistical series on the income sources of the aged population using data from the CPS.

Selected Abbreviations

ASEC	Annual Social and Economic Supplement
CPS	Current Population Survey
HRS	Health and Retirement Study
IRA	individual retirement account
IRS	Internal Revenue Service
PUF	public-use file
SSA	Social Security Administration
SSI	Supplemental Security Income

SSA's Office of Research, Evaluation, and Statistics (ORES) has published statistics on the income of the aged population based on public-use CPS data since 1976.[1] Two of its well-known and commonly cited publications are *Income of the Population 55 or Older* and *Income of the Aged Chartbook*. These publications provide estimates, overall and separately by demographic groups, of the prevalence and the amount of income from different sources (such as earnings, Social Security, pensions, and assets); total money income; the importance of different income sources relative to total income; shares of aggregate income by source; and poverty status; all broken out by various demographic characteristics.

This paper examines the question of whether or not the CPS Annual Social and Economic Supplement (ASEC), also known as the "March CPS,"[2] is a reliable data source for reporting income statistics for the aged population. To do this we compare results taken directly from the CPS ASEC public-use file (PUF) with alternative data files that supplement the survey data with matched administrative records from SSA and the Internal Revenue Service (IRS). We also compare results of the publicly available version of the March CPS with that of the University of Michigan's Health and Retirement Study (HRS).

This paper proceeds as follows. The next section discusses the relevant literature, including recent work that motivated this paper. The third section provides background and describes the data and methods we used for our analysis, focusing on income definitions from the various data files. The fourth section presents statistical results, focusing on comparisons of the distribution of

[1] Although ORES had previously published occasional CPS-based statistics on income of the aged population, biennial publication began in 1976.

[2] Every year, the Census Bureau fields the CPS ASEC in February and releases its results in March. Users have traditionally used the CPS ASEC and March CPS nomenclature interchangeably. Hereafter in this paper, we use "CPS ASEC" and "March CPS" (or simply "CPS") interchangeably.

aggregate income by source, population distribution by income, prevalence of income by source, the relative importance of Social Security, and poverty status. The final section discusses the implications of our results and provides concluding remarks. Appendices provide the income and poverty definitions used in the CPS and the HRS.

Literature Review

For years, there has been an ongoing discussion about the underreporting of retirement income in the CPS, with a wide acknowledgment that although underreporting exists, the CPS is still one of the best sources of information about the income of the U.S. population because of its large sample size, broad array of information collected, and periodicity. However, a 2017 study by Census Bureau economists Adam Bee and Joshua Mitchell renewed interest in that discussion. With the benefit of new linkages to administrative data, they were able to quantify the misreporting problem at the observation level, whereas prior research has focused on comparisons of aggregates across surveys and administrative data.

The studies conducted prior to Bee and Mitchell (2017) did not have the benefit of direct linkages to administrative microdata, and many therefore relied on comparisons of survey and administrative-data aggregates to study income underreporting in the CPS. Several studies used the IRS's Statistics of Income (SOI) series to compare aggregate measures of pension benefits. Using data from the 1990 SOI, Schieber (1995) asserted that the CPS undercounts pension and annuity income by as much as one-third. In a follow-up study, Woods (1996) noted that although there are concerns about comparing the two datasets, Schieber's assessment that the CPS is missing large portions of pension income relative to the SOI data was correct. More recently, Chen, Munnell, and Sanzenbacher (2018) compared results of the CPS and four other national survey datasets with administrative aggregates from the IRS's SOI and Social Security administrative data published in SSA's *Annual Statistical Supplement to the Social Security Bulletin*. They concluded, as prior researchers have, that the CPS misses large portions of retirement income. Interestingly, they found that the other four national surveys were much better than the CPS for capturing retirement income.[3]

Bee and Mitchell (2017) examined the extent and magnitude of measurement and reporting errors for different income sources by comparing the 2013 CPS reports with information from administrative data from SSA and IRS. They compared amounts reported in the CPS and amounts validated by linking the CPS results with the administrative data for five types of

[3] Chen, Munnell, and Sanzenbacher compared the CPS with the Survey of Consumer Finances (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 percent, 94 percent, 97 percent, and 85 percent of retirement income, respectively. However, they found that the CPS accounted for only about 47 percent of the aggregate non–Social Security retirement income based on the SOI data.

income: earnings, Social Security benefits, Supplemental Security Income (SSI) payments, interest and dividends, and "retirement income" (comprising pension benefits and retirement account distributions). The authors found that, among all households headed by an individual aged 65 or older in 2012, median household income was 30 percent higher in the administrative records than in the CPS ($44,400 versus $33,800). As a result, the poverty rate for persons aged 65 or older when estimated using public-use CPS data (9.1 percent) was 2.2 percentage points higher than the estimate using CPS results validated with administrative data (6.9 percent).

Bee and Mitchell (2017) also showed that the difference in estimated income is mainly due to underreporting of retirement income (from both defined benefit pensions and defined contribution retirement account withdrawals) and that the discrepancy in median income between survey and administrative data increased from about 20 percent in 1990 to about 30 percent in 2012. This finding reveals that the discrepancy, attributable mainly to CPS' failure to capture many retirement account distributions, arose at a time when retirement accounts and withdrawals from such accounts became more prevalent. Notably, the authors found that about 46 percent of the aged CPS respondents who report no income from retirement accounts actually have such income, according to the administrative records (a false negative type of error). Moreover, according to administrative records, persons with only individual retirement account (IRA) distributions are much less likely to report those distributions (thereby generating a higher false negative rate) than persons with distributions only from employer-sponsored defined contribution plans (false negative rates of 81 percent and 40 percent, respectively). Because of underreporting of retirement income, the CPS overstates the importance of Social Security to total income. When comparing the public-use CPS data with CPS data that have been supplemented with administrative data, the authors found that the proportion of persons aged 65 or older who rely on Social Security for at least 50 percent of their family income differed significantly (55 percent and 42 percent, respectively). Furthermore, the proportion of those relying on Social Security for at least 90 percent of their family income differed even more widely, at 26 percent versus 12 percent, respectively.

Data and Methods

We compare income and poverty statistics for the aged using four alternative data files. We derive the estimates presented here from two major household surveys (the 2016 March CPS and the 2016 wave of the HRS) and administrative data from SSA and IRS. The data files consist of:

1. CPS public-use data (the CPS PUF),
2. HRS public-use data (the HRS PUF),[4]

[4] In the tables and throughout this paper we will refer to the HRS public data as HRS PUF for brevity and consistency with CPS PUF.

3. Restricted CPS data that we have linked with administrative data from SSA (the "CPS+SSA" data file), and
4. Restricted CPS data that we have linked with administrative data from SSA and IRS (the "CPS+SSA+IRS" data file).

The administrative data from SSA provide information on earnings, Social Security benefit receipt, and SSI payment receipt. The IRS administrative data include information on retirement income from sources other than Social Security, including income generated from asset holdings. Restricted-access CPS data were linked to the SSA and IRS data using an anonymized unique identifier on a secure Census Bureau server.[5]

The analysis presented here is primarily concerned with the family income of persons aged 65 or older.[6] Unless otherwise stated—for example, in our analysis of aggregate income—the focus is on family income. A detailed discussion about the numerous methods researchers have devised to determine what counts as income is beyond the scope of this work. For this analysis, we follow the recommendations of Anguelov, Iams, and Purcell (2012) and count all distributions from retirement accounts as income, including infrequent and periodic withdrawals. This includes payments from both defined benefit and defined contribution plans, and traditional and Roth 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.

CPS ASEC

The CPS ASEC is a survey of a nationally representative sample of the U.S. noninstitutionalized population. The survey collects information on income from different sources—such as earnings, Social Security, pensions, assets, and government transfer programs—that the household (each family or household member) received during the prior year. In addition, the CPS collects detailed demographic information, including but not limited to age, gender, race/ethnicity, marital status, and household composition. As such, the CPS has long been a source for national estimates of household income and poverty rates across different population subgroups.

---

[5] The results presented in this paper were approved for release by the Census Bureau's Disclosure Review Board (CBDRB-FY20-018). 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.

[6] A family is a group of two or more people (one of whom is the householder) related by birth, marriage, or adoption and residing together; all such people (including related subfamily members) are considered members of one family. For persons aged 65 or older, we assign the total value of family income at each unit of observation. This value will simply be one's own personal income in the case of single persons aged 65 or older living alone. For others this value will likely be the combined income of the householder and a spouse. In some cases, it will be the combined income of the person aged 65 or older and any related subfamily members (spouse, child, other) living in the household.

As discussed in the previous section, several studies have been critical of the CPS over the years. These studies have emphasized that the CPS inadequately measures income from assets and tax-advantaged retirement accounts (such as 401(k) plans and IRAs), resulting in estimates that understate the importance of such accounts. Consequently, CPS results overstate the contribution of Social Security benefits (Iams and Purcell 2013; Fisher 2008; Davies and Fisher 2009; Miller and Schieber 2013; Munnell and Chen 2014). The Census Bureau has been receptive to these studies, and in 2015 fielded a redesigned survey instrument aimed at improving the collection of income data by implementing a number of 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, the 2015 CPS (2014 reference year) asked separate questions about retirement account withdrawals and distributions and collected information on property income.

Prior to the full implementation of the redesigned questionnaire in 2015, the 2014 March CPS randomly selected ⅜ of the sample to receive the new questionnaire, while the remaining ⅝ of the sample received the traditional CPS questionnaire. Comparisons made using the split-sample design from the 2014 CPS (2013 reference year) indicated that among aged households, the estimated real median income of the redesign respondents was 4.6 percent higher than that of respondents to the traditional questionnaire. In addition, estimates of the prevalence of retirement income other than Social Security were about 50 percent higher using the redesigned versus traditional questionnaire and the aggregate value of that income was about 22 percent higher. Interestingly, both the estimated prevalence and the aggregate value of Social Security income were only about 2 percent greater using the redesigned questionnaire, suggesting that the traditional CPS instrument measured Social Security relatively well (Semega and Welniak 2015).

This study has the advantage of using the 2016 March CPS, the second year of full implementation of the redesigned questionnaire. In addition, we follow Bee and Mitchell (2017) and match the CPS results with administrative data from SSA and IRS.[7] This allows us to compare our findings with their work, which evaluated the CPS prior to the redesign, using matched data up to reference year 2012. Of central interest is the question of whether underreporting of retirement income remained as prevalent in 2015 as it was prior to the CPS

[7] We used the same SSA and IRS data files that Bee and Mitchell used. We matched the administrative data files to the CPS using the Personal Identification Key (PIK). The Census Bureau uses the PIK to link administrative records with survey data. For the 2016 CPS (2015 income reference year), about 90 percent of respondents aged 65 or older were assigned a PIK, and therefore could be matched to administrative data. Bee and Mitchell (2017) limited their sample to survey respondents assigned a PIK and then reweighted the sample. For simplicity, we have opted to keep the full sample of CPS respondents aged 65 or older and to use the survey weights. After testing a number of ways to tabulate our statistics, we decided not to exclude those survey respondents without a PIK and not to reweight the sample. The 90 percent match rate is high enough that any differences were negligible.

redesign in 2012. Although the CPS interviews took place in 2016, the reference year for all income measures (defined in Appendix A) is the previous calendar year (2015).

We validated data on several income sources for CPS respondents by using an identifier that allowed us to link their 2016 CPS responses (reporting income in 2015) to administrative records. For the respondents with an identifier (about 90 percent of respondents aged 65 or older), we replaced the values reported in the CPS with values from the administrative records. For the remaining respondents we used self-reported values from the CPS. For linked respondents, administrative records from SSA allowed us to validate Social Security benefits (retirement and disability), SSI payments, and earnings from employment, resulting in the CPS+SSA data file. IRS administrative records allowed us also to validate income from retirement accounts (defined benefit and defined contribution employer-sponsored plans, and withdrawals from IRAs), and income from interest and dividends, resulting in the CPS+SSA+IRS file.

HRS

The HRS is the most comprehensive national longitudinal survey of Americans aged 51 or older.[8] 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 in 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. Another advantage of the HRS is that it asks respondents for their consent to link their survey information with earnings and benefits information from Social Security administrative records. Furthermore, the HRS is more systematic than the CPS in collecting information on pensions, retirement-plan account balances, and their distributions. If HRS 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.

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 percent) than amounts reported in the CPS; and while both samples had similar demographic characteristics, the HRS respondents were less likely to live alone than were their CPS counterparts. The authors conclude that "HRS incomes are higher than those of the Census Bureau surveys, but resolving whether

[8] The 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 HRS data would require a prohibitive amount of a user's time. To make the HRS data easier and 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.

this is due to better measurement or over-representation of higher-income families must be left to future research."

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, as well as additional variables derived from survey reports, which are created in a consistent way across survey years. Specifically, we focus on the sample of people who were aged 65 or older 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, and other sources. For married respondents, the spouse's income from those same sources is measured and available separately. The total household income is the sum of income the respondent received from all sources, and for married couples it includes incomes received by both the respondent and the spouse.[9,10] Although the interviews took place in 2016, the reference year for all income measures is the previous calendar year (2015). See Appendix B for the HRS variable definitions.

In an attempt to complement the HRS PUF and to provide a comparison for the CPS+SSA data file, we also created an HRS+SSA data file. For HRS survey respondents who provided consent, we matched survey reports with information from SSA's restricted earnings and benefits records. For respondents with a match, we replaced survey information on earnings and Social Security benefits with the respective information from the administrative records. However, for our sample of interest, the match rate in 2016 (with income reference year 2015) was only

[9] The RAND documentation data file states that: "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 and others 2020, 33).

[10] Note that different income measures in the RAND HRS file are reported separately for respondents and spouses (if married), and overall for the household. In general, with the exception of other income, the household variables are simply the sum of the respondent and spouse income. Hence, while we use the "household" reference to be consistent with terminology used in the RAND HRS file, we believe that the income variables in HRS are closer to the family income, particularly for respondents who live alone or with only a spouse/partner. For the estimations in this paper, we use the total household income measure (H13ITOT), which as noted in the previous footnote does not include income of other family members, unless reported in other income. To the extent that incomes of other resident family members are not reported, our measure of household income may likely underestimate family income among respondents with resident family members (other than spouse). However, given the age group for our study, it is plausible that it may not be a major issue, since only 22 percent of the HRS sample aged 65 or older live in a family with 3 or more members. In future updates to this paper, we will try to quantify the extent and magnitude of income from other resident members, and determine whether it may affect results.

14 percent.[11] Given the low match rate, the results using the survey-only information (that is, HRS PUF) did not differ from those based on survey results augmented with administrative records (HRS+SSA). Hence, we decided not to report those results here.[12]

Social Security Administrative Data

Previous research has shown that survey respondents may misreport their earnings or Social Security benefits (Dushi, Iams, and Trenkamp 2017; Iams and Purcell 2013; Meyer, Mok, and Sullivan 2015; Bricker and Engelhardt 2007; Bound, Brown, and Mathiowetz 2001; Pedace and Bates 2000; Bollinger 1998; Rodgers, Brown, and Duncan 1993). To account for reporting error, we use Social Security administrative records, which maintain information on annual earnings, Social Security benefits, and SSI payments. More specifically, we match survey data with restricted SSA records for about 90 percent of the sample in the 2016 CPS. The earnings information come from the Detailed Earnings Record file, which indicates the amount of covered or noncovered earnings, as well as self-reported earnings. It also contains detailed information on total compensation, earnings that are subject to Social Security and Medicare tax, and voluntary tax-deferred contributions to retirement accounts. Information about Social Security benefit amounts comes from the Payment History Update System file, which contains information on the net amount of benefits paid to a beneficiary as well as the amount of the Medicare premium paid on the beneficiary's behalf to the Centers for Medicare & Medicaid Services. Hence, the true (or gross) amount of Social Security benefits that a retired beneficiary is entitled to is the sum of net benefits and Medicare premiums. For respondents in our sample with a matched record, the estimated benefit amount in a given year is equal to the sum of monthly benefits received and the Medicare premiums paid. For those without a matched record in our sample, the estimated amount of benefits equals their self-reported amount. Finally, we use the information from the Supplemental Security Records file to obtain information about SSI receipt and payment amounts.

---

[11] Since its inception in 1992, the HRS has asked respondents to provide consent to link the survey information with data from IRS's earnings and SSA's benefits records. Until the 2004 wave, the consent form was retrospective, meaning that the records could be matched for all years prior to the consent year. Thus, for someone who consented in 2004, the record would be matched up to the year 2003. In the 2006 wave, the form was changed to allow prospective consent, with which records could be matched up to the year 2030. In addition, in 2006, the HRS introduced face-to-face (FTF) interviews for half of the sample, with the FTF interview samples to rotate in subsequent waves. The rationale was that FTF interviews would lead to increased consent rates and respondents who provided consent would not be asked again. As a result of FTF interviews, the HRS match rate increased to 65–75 percent. However, in 2012, the consent form changed again, and instead of prospective consent to match earnings and benefits records up to 2030, the match was allowed for up to 6 years after the consent year for IRS earnings records and up to 12 years for SSA benefits records. This new shorter-term prospective match would apply even to those respondents who consented prior to 2012 and consequently all survey respondents would need to be asked again for new consents every 6 to 8 years. Hence, in 2016, half of the sample of any age without a valid consent were asked to provide a consent. Expectedly, these changes affected the consent rate for the overall sample and particularly for the sample aged 65 or older.

[12] Those results are available on request from the authors.

IRS Administrative Data

We match data from two IRS administrative data files to the 2016 March CPS. 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 data from both defined benefit and defined contribution employer-sponsored plans, as well as withdrawals from IRAs. As Bee and Mitchell (2017) note, this file excludes data on direct rollovers, Section 1035 exchanges, and Roth IRA conversions, which is to our advantage, as we want to count only income that permanently leaves tax-preferred accounts.

The second IRS administrative data file we use is composed of data from the IRS Form 1040. From this file, we are able to validate interest and dividend income for CPS respondents who filed an IRS Form 1040 for 2015. For CPS respondents who do not have a 1040 record, either because they did not file one (those with income under a certain amount are not required to file Form 1040) or they did not have interest or dividend income, we use the amount from the CPS. Because we are concerned with the *family income* of persons, we do not worry about splitting income from the Form 1040 for joint filers.[13]

A Brief Word on Income Not Captured by the Administrative Data

For the CPS records that we are able to match, the administrative data files described above capture the bulk of the income that respondents are likely to have. However, there are some income sources that the administrative data files either do not measure or may not measure well in some cases. Income data collected in the CPS that are not included in the administrative data files include unemployment insurance, workers' compensation, public assistance (other than SSI), rents/royalties/estates/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 CPS. Additionally, the Detailed Earnings Record includes only the taxable portion of self-employment earnings and likely misses earnings from the informal labor market.

Results

In this section, we present estimates from each of the four data files; the HRS PUF, the CPS PUF, the CPS+SSA file, and the CPS+SSA+IRS file. Table 1 provides a comparison of demographic characteristics between the CPS and HRS samples. Despite a fundamental difference in the survey designs (CPS is cross-sectional and HRS is longitudinal), the two samples of aged respondents exhibit similar demographic characteristics. The two samples differ, however, with respect to the number of persons in the family and the proportions who are Social Security beneficiaries. Similar to findings by Czajka and Denmead (2008) using the 2002 HRS wave,

[13] For additional detail on the IRS data extracts, see Bee and Mitchell (2017).

aged HRS respondents in 2016 are less likely to live alone or in a two-person family than their CPS counterparts are. Furthermore, the proportion of respondents who are Social Security beneficiaries is higher in the HRS than in the CPS.

We begin by discussing aggregate measures of income across the four different data files, looking specifically at how sources of income differ relative to each other and across the files.

Table 1.
Demographic characteristics of survey participants aged 65 or older, by age group and survey, 2016 (in percent)

	HRS					CPS				
		Age					Age			
Characteristic	All	65–69	70–74	75–79	80+	All	65–69	70–74	75–79	80+
Sex										
Men	43.9	45.1	47.4	44.1	38.7	44.6	47.1	45.4	45.8	39.3
Women	56.1	54.9	52.6	55.9	61.3	55.4	52.9	54.6	54.2	60.7
Race/ethnicity										
Non-Hispanic white [a]	80.2	78.3	79.3	79.7	84.0	84.7	83.3	84.5	84.9	86.7
Non-Hispanic black [b]	9.2	10.4	9.3	8.5	7.9	9.1	9.9	9.2	9.2	7.9
Non-Hispanic other [c]	2.7	3.3	2.9	2.6	1.9	4.5	4.9	4.6	4.1	4.0
Hispanic (any race)	7.9	8.0	8.5	9.2	6.2	8.1	8.8	8.0	7.7	7.5
Marital status										
Married	59.4	69.0	65.7	60.1	39.7	56.4	64.6	61.0	57.6	38.8
Nonmarried	40.6	31.0	34.3	40.0	60.3	43.6	35.4	39.0	42.4	61.2
Widowed	23.3	10.1	15.2	24.7	48.2	24.0	10.5	18.7	25.3	48.5
Divorced	12.5	15.1	13.9	11.7	8.2	12.0	15.4	13.3	10.5	6.7
Never married	4.8	5.8	5.2	3.6	3.9	5.2	6.7	4.8	4.4	3.9
Educational attainment										
Less than high school diploma	15.1	9.5	14.4	19.0	20.5	14.6	10.8	13.2	16.0	20.8
High school diploma or equivalent	33.1	28.4	33.8	35.7	37.2	33.2	29.5	32.4	36.1	37.4
Some college	24.2	28.7	23.8	21.8	20.1	16.2	17.0	17.3	16.3	13.9
College degree	27.6	33.4	28.1	23.5	22.2	35.9	42.7	37.1	31.5	28.0
Persons in family [d]										
1	28.1	23.0	24.0	27.8	40.6	32.1	25.6	29.3	31.5	45.0
2	49.6	54.6	54.1	51.0	36.0	52.0	56.1	55.0	54.6	40.9
3 or more	22.3	22.4	21.9	21.2	23.4	15.9	18.3	15.8	13.9	14.1
Social Security beneficiaries	89.7	76.6	96.7	96.4	96.4	82.1	71.4	86.0	87.7	89.7
Total weighted count (thousands)	50,152	17,067	12,177	8,432	12,476	47,550	16,520	11,430	8,420	11,180

SOURCE: Authors' calculations based on HRS (wave 13, 2016) and 2016 CPS ASEC.

NOTES: Estimates are weighted using survey weights.

Rounded components of percentage distributions may not sum to 100.0.

a. Identified in CPS as "white alone."

b. Identified in CPS as "black alone."

c. Identified in CPS as "Asian alone."

d. In the HRS, count is derived from the "family composition" variable in the RAND file that was created to define poverty thresholds and poverty rate.

Aggregate Income

In Table 2, we compare shares of aggregate income for the total U.S. population aged 65 or older, by income source, across the four data files.[14] The first thing to note here is that the results are relatively consistent across the HRS PUF, CPS PUF, and CPS+SSA data files. Comparing the aggregate shares based on the CPS PUF and CPS+SSA files, we see only slight shifts in earnings (from 30.6 percent to 29.4 percent) and Social Security (from 34.9 percent to 35.5 percent)—the two income categories for which administrative data most often replace survey results. The finding suggests that the CPS PUF measures earnings and Social Security benefits rather accurately, and is consistent with the findings of Bee and Mitchell (2017). However, clear differences emerge when comparing these files with the CPS+SSA+IRS file, in which pension income accounts for a much larger share of aggregate income than is reported in the CPS.[15] Specifically, in the CPS PUF, CPS+SSA, and CPS+SSA+IRS files, pension income accounts for 21.7 percent, 21.3 percent, and 35.9 percent of aggregate income, respectively. The 14 percentage-point difference suggests that while the CPS redesign may have somewhat improved the reporting of pension income or retirement account withdrawals, its success in improving the measurement of retirement income has been limited, a finding that is consistent with Bee and Mitchell (2017) and Chen, Munnell, and Sanzenbacher (2018). Particularly noteworthy is that in the CPS+SSA+IRS file, pension income accounts for the largest share of aggregate income,

Table 2.
Percentage distribution of aggregate income among individuals aged 65 or older, by source: Measurements from four alternative data files, 2015

	Survey data (unmatched)		Survey data matched with administrative records [a]	
Income source	HRS PUF	CPS PUF	CPS+SSA	CPS+SSA+IRS
Total	100.0	100.0	100.0	100.0
Earnings	25.6	30.6	29.4	24.7
Social Security	31.7	34.9	35.5	29.9
Pensions [b]	25.0	21.7	21.3	35.9
Asset income	9.6	8.8	8.9	7.5
Other	8.1	4.0	4.9	2.0
Weighted count (thousands)	50,152	47,550	47,550	47,550

SOURCE: Authors' calculations based on HRS (wave 13, 2016), 2016 CPS ASEC, and administrative data from SSA and IRS. Income questions ask respondent about income received in the previous calendar year (2015).

a. Approved for release by the Census Bureau's Disclosure Review Board (CBDRB-FY20-018).

b. Includes IRA withdrawals.

[14] For each income source, we use survey weights to aggregate the income for the total for U.S. population. Then, we calculate each source's share of the total income.

[15] Note that CPS PUF and CPS+SSA files use pension income as reported in the survey, whereas CPS+SSA+IRS replaces the survey-reported pension income with data from IRS records.

whereas in the CPS PUF and the CPS+SSA files, Social Security accounts for the largest share. Interestingly, while the HRS PUF data account for pension income better than the CPS PUF and CPS+SSA files, the share of aggregate income the HRS PUF attributes to pensions (25.0 percent) is still lower (by about 10 percentage points) than that in the CPS+SSA+IRS file. The share of aggregate income attributable to assets is almost the same across the four data files, whereas the "other" income share[16] is lower in all of the CPS data files than in the HRS data, with the lowest share being in the CPS+SSA+IRS file.

Aged Population by Family Income

We now focus on the distribution of persons aged 65 or older by family income level.[17] Table 3 shows the income distribution of this population for each of the four data files. In the two public data files (HRS PUF and CPS PUF) and the CPS+SSA file, the distributions look very similar. However, there are noticeable differences in the distribution from the CPS+SSA+IRS file, for which the general trend is a shift to the higher end of the income distribution. The effect of

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

	Survey data (unmatched)		Survey data matched with administrative records [a]	
Family income ($)	HRS PUF	CPS PUF	CPS+SSA	CPS+SSA+IRS
Total	100.0	100.0	100.0	100.0
Less than 5,000	1.7	2.3	1.3	0.8
5,000–9,999	4.0	3.0	4.2	3.6
10,000–14,999	8.0	7.2	6.9	5.5
15,000–19,999	7.5	7.3	7.6	5.6
20,000–24,999	7.1	7.5	7.2	5.6
25,000–29,999	5.8	6.8	6.6	5.3
30,000–34,999	5.6	6.1	6.6	5.2
35,000–39,999	5.2	5.7	5.8	5.8
40,000–44,999	4.9	4.6	5.2	4.7
45,000–49,999	4.6	4.8	4.4	4.3
50,000–74,999	15.4	15.6	16.2	18.2
75,000–99,999	8.9	10.2	10.3	12.5
100,000 or more	21.4	18.9	17.8	23.0
Weighted count (thousands)	50,152	47,550	47,550	47,550

SOURCE: Authors' calculations based on HRS (wave 13, 2016), 2016 CPS ASEC, and administrative data from SSA and IRS. Income questions ask respondent about income received in the previous calendar year (2015).

NOTE: Rounded components of percentage distributions may not sum to 100.0.

a. Approved for release by the Census Bureau's Disclosure Review Board (CBDRB-FY20-018).

[16] See Appendices for CPS and HRS definitions of the "other income" category.

[17] See Appendix B for the definitions of household (family) income in HRS.

adding income from the Forms 1099-R and 1040 results in a smaller proportion of individuals concentrated in the lowest family-income categories, shares more evenly distributed across the middle income levels, and higher proportions in the highest income categories. Individuals in the lowest three family income categories (below $15,000) comprise 9.9 percent of the aged population in the CPS+SSA+IRS data, compared with 12.5 percent and 13.7 percent in the CPS PUF and HRS PUF, respectively. Furthermore, those in the highest three categories (with family income of $50,000 and higher) comprise 53.7 percent of the population in the CPS+SSA+IRS data compared with 44.7 percent of the population when using the CPS PUF data or 45.7 percent of the population when using the HRS PUF data.

Sources of Income

Table 4 shows, for each of seven income sources, the proportion of people aged 65 or older with family income from that source, broken out by sex. The estimates indicate that the proportion of persons aged 65 or older reporting income from pension and IRA withdrawals in the CPS PUF is 47.2 percent, in contrast with a proportion of 69.0 percent in the CPS+SSA+IRS file. This represents a substantial difference (21.8 percentage points) between the proportion of CPS respondents who have income from retirement accounts and the proportion who report that income in the survey. This finding suggests that not only is the aggregate amount of pension and IRA income greater with the addition of the IRS administrative data (as shown in Table 2), but so is the proportion of persons who have family income from this source.

In addition to the wide variance in the prevalence of retirement income from pensions and IRAs shown in Table 4, some other differences across the four files are worth noting. For instance, the differences in prevalence of Social Security and earnings across the three CPS files is relatively small. This reiterates the earlier assertion that the CPS PUF does reasonably well in measuring Social Security benefits and earnings. Furthermore, note that while there are no major differences across CPS files, the HRS shows a somewhat higher proportion of aged persons receiving income from Social Security, but a lower proportion with family income from earnings. Lastly, the breakdowns by sex show that the similarities and differences across the four files are more or less consistent.

Reliance on Social Security

In Table 5, we present Social Security reliance statistics for each of the four files. The sample in Table 5 is restricted to persons in Social Security beneficiary families (that is, at least one family member, respondent or other, is a beneficiary).[18] The first reliance threshold encompasses

[18] Note that in the HRS, the family income measures include only the income received from the respondents (if single, widowed, divorced, or separated) and from the spouse or partner income (if coupled). Hence, the Social Security beneficiary can be either the respondent, or the spouse, or both.

Table 4.
Percentages of individuals aged 65 or older with family income, by source and sex: Measurements from four alternative data files, 2015

	Survey data (unmatched)		Survey data matched with administrative records [a]	
Income source	HRS PUF	CPS PUF	CPS+SSA	CPS+SSA+IRS
		All		
Social Security	93.1	86.0	89.5	89.5
Asset income	54.3	68.4	68.4	69.2
Pensions [b]	59.2	47.2	47.2	69.0
Earnings	35.6	40.7	45.0	45.0
Veterans' benefits	10.2	5.6	5.6	5.6
Cash public assistance	8.0	4.6	6.0	6.0
Other [c]	7.5	10.8	10.8	10.8
Weighted count (thousands)	50,152	47,550	47,550	47,550
		Men		
Social Security	92.2	84.7	88.7	88.7
Asset income	57.7	70.6	70.6	71.4
Pensions [b]	60.6	48.5	48.5	69.2
Earnings	42.1	44.2	48.6	48.6
Veterans' benefits	12.8	7.5	7.5	7.5
Cash public assistance	6.1	4.0	4.9	4.9
Other [c]	8.6	8.5	8.5	8.5
Weighted count (thousands)	22,020	21,210	21,210	21,210
		Women		
Social Security	93.8	87.0	90.2	90.2
Asset income	51.7	66.6	66.6	67.5
Pensions [b]	58.1	46.2	46.2	68.8
Earnings	30.5	37.8	42.0	42.0
Veterans' benefits	8.2	4.1	4.1	4.1
Cash public assistance	9.4	5.1	6.8	6.8
Other [c]	6.7	12.7	12.7	12.7
Weighted count (thousands)	28,132	26,340	26,340	26,340

SOURCE: Authors' calculations based on HRS (wave 13, 2016), 2016 CPS ASEC, and administrative data from SSA and IRS. Income questions ask respondent about income received in the previous calendar year (2015).

a. Approved for release by the Census Bureau's Disclosure Review Board (CBDRB-FY20-018).

b. Includes IRA withdrawals.

c. Excludes veterans' benefits.

individuals for whom family Social Security income comprises 50 percent or more of total family income. The other two reliance thresholds are 75 percent or more and 90 percent or more. We estimate the proportions of aged persons whose ratio of Social Security income to total family income exceed those thresholds. Comparing the two public-use files, we see that the CPS PUF produces slightly higher reliance figures than the HRS PUF 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 49.8 percent in the HRS PUF compared with 52.5 percent in the CPS

Table 5.
Percentages of individuals aged 65 or older for whom Social Security represents a selected proportion of family income, by sex: Measurements from four alternative data files, 2015

Social Security as a proportion of family income	Survey data (unmatched)		Survey data matched with administrative records [a]	
	HRS PUF	CPS PUF	CPS+SSA	CPS+SSA+IRS
		All		
50% or more	49.8	52.5	51.1	39.9
75% or more	30.5	33.6	32.1	21.2
90% or more	21.2	25.6	23.8	13.8
Weighted count (thousands)	46,687	40,520	42,250	42,250
		Men		
50% or more	44.6	49.0	47.9	37.3
75% or more	25.8	30.1	29.0	18.6
90% or more	17.6	22.5	21.1	12.1
Weighted count (thousands)	20,295	17,800	18,650	18,650
		Women		
50% or more	53.9	55.2	53.7	42.0
75% or more	34.1	36.4	34.6	23.3
90% or more	23.9	28.0	25.9	15.1
Weighted count (thousands)	26,392	22,720	23,600	23,600

SOURCE: Authors' calculations based on HRS (wave 13, 2016), 2016 CPS ASEC, and administrative data from SSA and IRS. Income questions ask respondent about income received in the previous calendar year (2015).

a. Approved for release by the Census Bureau's Disclosure Review Board (CBDRB-FY20-018).

PUF. Similarly, the proportions at the other end of the spectrum (90 percent or more of family income from Social Security) range from 21.2 percent in the HRS PUF to 25.6 percent in the CPS PUF. This difference could arise because the HRS asks respondents for the net amount of Social Security benefits and thus excludes the Medicare Part B and/or Part D premiums that are deducted from the benefits, whereas the CPS asks for the gross Social Security benefit amount, including the Medicare premium (see Dushi, Iams, and Trenkamp 2017).[19]

Expanding the comparisons to include the CPS+SSA file reveals similar patterns. The three files yield relatively similar results, with only slight differences between them. For example, the proportions of persons aged 65 or older who rely on Social Security for 75 percent or more of their total family income are 30.5 percent, 33.6 percent, and 32.1 percent, respectively, in the HRS PUF, CPS PUF, and CPS+SSA files. The highest value (from the CPS PUF) and the lowest (from the HRS PUF) vary by only 3.1 percentage points.

[19] The reliance rates may also be higher in the CPS than the HRS because the CPS defines family income as including income of family members in addition to the spouse, increasing the reported amount of family income received from Social Security. In future updates, we will provide estimates of mean and median Social Security income at the respondent and family level for both the HRS and CPS samples.

Shifting attention to the CPS+SSA+IRS columns of Table 5 shows lower apparent reliance on Social Security when incorporating the IRS administrative data. 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 PUF. The proportion is only slightly lower (23.8 percent, a 1.8 percentage point difference) in the CPS+SSA file. This is not surprising, as the income sources replaced in the CPS+SSA file (earnings, Social Security, and SSI) are relatively accurate in the CPS PUF; and in the case of SSI, the prevalence is low. However, that proportion is substantially lower—13.8 percent—in the CPS+SSA+IRS file. The marginal difference between the CPS+SSA file (23.8 percent) and the CPS+SSA+IRS file (13.8 percent) is 10 percentage points, which is attributable to a correction in retirement income from the IRS administrative data. The same pattern is evident for individuals who rely on Social Security for at least 75 percent or at least 50 percent of family income, and for both men and women across all three reliance-threshold categories.

Poverty Rates

As with the other metrics, there are clear differences in poverty rates resulting from the additional retirement income captured by incorporating the IRS administrative data, which the public-use CPS data miss. Table 6 shows that the poverty rate for all persons aged 65 or older is 7.1 percent in the CPS+SSA+IRS file, compared with 8.7 percent in the CPS+SSA file and 8.8 percent in the CPS PUF. Interestingly, the result in the HRS PUF tracks that of the CPS+SSA+IRS file closely, with a slightly lower poverty rate of 6.6 percent.[20,21]

A closer look at the three CPS files shows that poverty rates of the CPS PUF and CPS+SSA files track each other closely for most race/ethnicity categories (Table 6) and age groups (Table 7). Expanding the comparison to include the CPS+SSA+IRS file, Tables 6 and 7 show poverty rates

[20] It is worth emphasizing that the different poverty rates in the HRS and the CPS may be due to several reasons that affect family income and consequently poverty rate. First, as Table 1 shows, HRS respondents are more likely than CPS respondents to live in a family with three or more members and are more likely to be Social Security beneficiaries. Second, HRS respondents are more likely to be in families with income of $50,000 or more (45.7 percent versus 44.7 percent; see Table 3). Both of these factors would plausibly lead to higher total family income and lower poverty rates in the HRS. Third, the total family income variable in the RAND-HRS data file does not include the income of other residing family members, although only 22 percent of HRS respondents aged 65 or older live in a family with three or more members and are affected by this exclusion. Lastly, the family income measure in Table 3 does not include Medicare Part B and/or Part D premiums deducted from Social Security benefits, whereas these amounts are included in the family income variable used for calculating poverty rate in Tables 6 and 7. The addition of Medicare premiums would lead to higher income and thus an upward shift in income distribution and a lower poverty rate. See Appendix B for a detailed discussion of the HRS total income and poverty measures.

[21] Also, as noted above, Czajka and Denmead (2008) found that HRS-reported household income in 2002 among people aged 51 or older was 20–30 percent higher than that in the CPS, with HRS respondents being less likely to live alone than their CPS counterparts. This, if still the case in the 2016 wave of the HRS, would help explain the poverty rate being lower in the HRS than the CPS. In future updates of this paper, we will examine whether higher-income families are overrepresented in the HRS survey relative to the CPS, along with other factors that could contribute to the difference in poverty rates observed in the two surveys.

Table 6.
Percentages of individuals aged 65 or older in or near poverty, by sex, race/ethnicity, and marital status: Measurements from four alternative data files, 2015

Characteristic	Survey data (unmatched)		Survey data matched with administrative records [a]	
	HRS PUF	CPS PUF	CPS+SSA	CPS+SSA+IRS
		In poverty		
All	6.6	8.8	8.7	7.1
Sex				
Men	4.4	7.0	6.9	5.6
Women	8.4	10.3	10.2	8.3
Race/ethnicity				
Non-Hispanic white [b]	3.5	7.5	7.5	6.0
Non-Hispanic black [c]	20.6	18.4	16.0	12.9
Non-Hispanic other [d]	8.6	11.8	14.3	13.6
Hispanic (any race)	22.0	17.5	18.0	16.5
Marital status				
Married	2.9	4.4	4.1	3.0
Nonmarried	12.1	14.6	14.7	12.4
		In or near poverty		
All	10.6	13.8	13.8	11.1
Sex				
Men	8.0	10.6	10.8	8.8
Women	12.8	16.4	16.3	12.9
Race/ethnicity				
Non-Hispanic white [b]	6.5	12.1	12.2	9.5
Non-Hispanic black [c]	29.8	26.1	23.9	19.5
Non-Hispanic other [d]	10.6	17.3	20.1	19.3
Hispanic (any race)	31.4	25.7	26.4	24.2
Marital status				
Married	4.8	6.9	6.5	4.8
Nonmarried	19.2	22.7	23.3	19.2
Weighted count (thousands)	50,152	47,550	47,550	47,550

SOURCE: Authors' calculations based on HRS (wave 13, 2016), 2016 CPS ASEC, and administrative data from SSA and IRS.

NOTES: "In poverty" = with income at or below 100% of the federal poverty guideline; "in or near poverty" = with income at or below 125% of the federal poverty guideline.

Poverty measures are based on family income (in 2015 dollars) and 2015 Census Bureau poverty thresholds corresponding to family size and composition.

Estimates are weighted using survey weights.

a. Approved for release by the Census Bureau's Disclosure Review Board (CBDRB-FY20-018).

b. Identified in CPS as "white alone."

c. Identified in CPS as "black alone."

d. Identified in CPS as "Asian alone."

Table 7.
Percentages of individuals aged 65 or older in or near poverty, by age group: Measurements from four alternative data files, 2015

	Survey data (unmatched)		Survey data matched with administrative records [a]	
Age group	HRS PUF	CPS PUF	CPS+SSA	CPS+SSA+IRS
		In poverty		
65–69	6.9	8.1	8.1	6.8
70–74	5.0	7.8	7.0	5.9
75–79	6.2	8.2	8.7	7.3
80 or older	8.3	11.4	11.3	8.5
		In or near poverty		
65–69	10.6	11.9	12.0	10.1
70–74	9.0	12.1	11.6	9.7
75–79	10.2	13.6	13.8	10.8
80 or older	13.0	18.5	18.9	14.1
Weighted count (thousands)	50,152	47,550	47,550	47,550

SOURCE: Authors' calculations based on HRS (wave 13, 2016), 2016 CPS ASEC, and administrative data from SSA and IRS.

NOTES: "In poverty" = with income at or below 100% of the federal poverty guideline; "in or near poverty" = with income at or below 125% of the federal poverty guideline.

Poverty measures are based on family income (in 2015 dollars) and 2015 Census Bureau poverty thresholds corresponding to family size and composition.

Estimates are weighted using survey weights.

a. Approved for release by the Census Bureau's Disclosure Review Board (CBDRB-FY20-018).

that are lower than those in the other two CPS files for nearly all race/ethnicity categories and age groups. The only exception is the Asian ("non-Hispanic other") race/ethnicity category, for which the poverty rate is 11.8 percent in the CPS PUF, 14.3 percent in the CPS+SSA file, and 13.6 percent in the CPS+SSA+IRS file.

Summary and Follow-Up Work

We have supplemented March 2016 CPS data with administrative data from SSA and IRS and compared the matched data files with the original public-use data to reveal that, despite the redesigned CPS questionnaire (first fully implemented in the 2015 survey), underreporting of retirement income continues to be an issue with the public-use data. The analysis presented here confirms earlier research that shows retirement income from sources other than Social Security to be significantly underreported in the CPS. While the HRS is better than the public-use CPS in estimating the income of the aged population, we find that it still produces lower figures than those generated by the CPS data supplemented with SSA and IRS administrative data.

The CPS PUF reports Social Security as the primary source of aggregate income among the population aged 65 or older; however, our CPS+SSA+IRS data file reports that pensions (including IRA withdrawals) are the largest source of aggregate income. Furthermore, the proportion

of persons aged 65 or older relying on Social Security for 90 percent or more of their family income is reported as 25.6 percent in the public-use CPS but is 13.8 percent if the CPS data are supplemented with the full set of administrative data (CPS+SSA+IRS). For comparison, the HRS public-use data report the same statistic as 21.2 percent. Finally, 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 percent to 7.1 percent. Meanwhile, the HRS produced a somewhat lower poverty rate estimate of 6.6 percent.

The results presented here provide strong evidence in support of supplementing survey data with administrative data to describe the income of older Americans. Bee and Mitchell (2017) shed light on pervasive underreporting of income from retirement pensions in the CPS prior to the questionnaire redesign, and our work confirms that it continues to be an issue. Thus, the CPS ASEC PUF data understate the retirement security and well-being of the aged population. The SSA publications *Income of the Population 55 or Older* and the *Income of the Aged Chartbook* reach both the media and policymakers. Through 2016, those publications used the CPS ASEC PUF; they have since been suspended, pending analysis of the redesigned CPS questionnaire and potential alternative data sources. We find that survey data supplemented with administrative data are needed to ensure that these publications provide the public with reliable information.

Of course, moving from exclusively public-use data to blended survey and administrative data is not without risks and challenges. Of utmost concern when using administrative data is the avoidance of any disclosure that could potentially lead to the identification of individuals. It is worth noting that SSA, the Census Bureau, and IRS take the protection of personally identifiable information very seriously. However, the measures necessary to protect administrative data present challenges to the timely release of statistics. For a statistical publication released on a regular basis, this could be an important consideration. In this case, matching the CPS with administrative data from SSA and IRS requires interagency cooperation and lengthy disclosure review processes that would inevitably result in delays.

In forthcoming work we plan to make several additions and improvements. We will extend the analysis on aggregate income to include total dollar amounts by each income source for each data file. This should add some much-needed context to the current presentation of the distribution of aggregate income by source. Additionally, we plan to include a more detailed look at the population distributions by income level across the four data files. This will include statistics on mean and median income within each income quintile. Lastly, we plan to conduct sensitivity analyses including investigating whether respondents with matched administrative records differ systemically from unmatched respondents, and reweighting the matched CPS sample.

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 nonfarm 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. Nonfarm 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 income**: 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**: is the sum of Social Security benefits and public and private pensions.
  - —**Social Security**: Social Security includes retired-worker benefits, dependents' 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.

- **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 Aid to Families with Dependent Children (AFDC, ADC), temporary assistance to needy families (TANF), general assistance, and emergency assistance.
- **Other income**: is total income minus earnings, Social Security, pensions, asset income, and cash public assistance; included are unemployment and workers' 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 detect 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 excludes capital gains and noncash 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).

The poverty thresholds are increased each year by the same percentage as the annual average CPI.

Appendix B: Definitions of HRS Income Variables and Poverty Rate

Earnings (individual-level variable)

For each HRS respondent, total survey-reported earnings is 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 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 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 each 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, pensions, 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 pension (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.[22] 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. We recognize that this derived variable may not exactly track the pension income in the CPS PUF, but it may more closely compare with the pension income variable in the CPS+SSA+IRS file (which includes IRA withdrawals).

[22] It 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.

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 estimating the poverty rate in Tables 6 and 7, 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 noncash 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 22 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 deductions;[23] 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 noncash 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—less food stamps, and including Medicare Part B and/or Part D premiums deducted from Social

[23] Medicare Part B and/or Part D premiums are added if the respondent reports that they were deducted from Social Security payments.

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. 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 total household income. More specifically, total household income, for poverty calculation purposes, is equal to:

(Total household income – food stamps)
+ (Medicare Part B and/or Part D premiums in instances when the respondent had deducted these amounts from reported Social Security benefits)
+ (income of non-core resident family members)
– (income of any core HRS nursing home residents, including earnings, pensions, Social Security, SSI, unemployment and workers compensation, and government transfer income)

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.[24]

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 percent 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 percent of the poverty threshold) then the respondent is defined as being in or near poverty.

References

Anguelov, Chris E., Howard M. Iams, and Patrick J. Purcell. 2012. "Shifting Income Sources of the Aged." *Social Security Bulletin* 72(3): 59–68.

Bee, C. Adam, and Joshua W. Mitchell. 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.

Bollinger, Christopher R. 1998. "Measurement Error in the Current Population Survey: A Nonparametric Look." *Journal of Labor Economics* 16(3): 576–594.

Bound, John, Charles Brown, and Nancy Mathiowetz. 2001. "Measurement Error in Survey Data." In *Handbook of Econometrics, Volume 5,* edited by James J. Heckman and Edward Leamer (3705–3843). Amsterdam: Elsevier Science BV.

Bricker, Jesse, and Gary V. Engelhardt. 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.

[24] Note that the terms household income and family income are used interchangeably in the HRS.

Bugliari, Della, Joanna Carroll, Orla Hayden, Jessica Hayes, Michael Hurd, Adam Karabatakis, Regan Main, Joyce Marks, Colleen McCullough, Erik Meijer, Michael Moldoff, Philip Pantoja, Susan Rohwedder, and Patricia St. Clair. 2020. *RAND HRS Longitudinal File 2016 (V2) Documentation.* Santa Monica, CA: RAND Center for the Study of Aging.

Census Bureau. 1982. *Characteristics of the Population Below the Poverty Level: 1980.* Current Population Reports Series P-60, No. 133. Washington, DC: Government Printing Office.

———. 2020. *Current Population Survey 2020 Annual Social and Economic (ASEC) Supplement Technical Documentation.* Washington, DC: Census Bureau. https://www.census.gov/data/datasets/time-series/demo/cps/cps-asec.html.

Chen, Anqi, Alicia H. Munnell, and Geoffrey T. Sanzenbacher. 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.

Czajka, John L., and Gabrielle Denmead. 2008. *Income Data for Policy Analysis: A Comparative Assessment of Eight Surveys. Final Report.* Washington, DC: Mathematica Policy Research.

Davies, Paul S., and T. Lynn Fisher. 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, Irena, Howard M. Iams, and Brad Trenkamp. 2017. "The Importance of Social Security Benefits to the Income of the Aged Population." *Social Security Bulletin* 77(2): 1–12.

Fisher, T. Lynn. 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.

Iams, Howard M., and Patrick J. Purcell. 2013. "The Impact of Retirement Account Distributions on Measures of Family Income." *Social Security Bulletin* 73(2): 77–84.

Meyer, Bruce D., Wallace K. C. Mok, and James X. Sullivan. 2015. "The Under-Reporting of Transfers in Household Surveys: Its Nature and Consequences." Chicago, IL: University of Chicago. https://harris.uchicago.edu/files/underreporting.pdf.

Miller, Billie Jean, and Sylvester J. Schieber. 2013. "Employer Plans, IRAs, and Retirement Income Provision: Making a Molehill Out of a Mountain." *Insider* (October). London: Willis Towers Watson.

Munnell, Alicia H., and Anqi Chen. 2014. "Do Census Data Understate Retirement Income?" Issue in Brief No. 14-19. Chestnut Hill, MA: Center for Retirement Research at Boston College.

Pedace, Roberto, and Nancy Bates. 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.

Rodgers, Willard L., Charles Brown, and Greg J. Duncan. 1993. "Errors in Survey Reports of Earnings, Hours Worked, and Hourly Wages." *Journal of the American Statistical Association* 88(424): 1208–1218.

Schieber, Sylvester J. 1995. "Why Do Pension Benefits Seem So Small?" *Benefits Quarterly* 11(4): 57–70.

Semega, Jessica L., and Edward Welniak, Jr. 2015. "The Effects of the Changes to the Current Population Survey Annual Social and Economic Supplement on Estimates of Income." Paper presented at the proceedings of the 2015 Allied Social Science Association Research Conference, Boston, MA (January 3–5).

Woods, John R. 1996. "Pension Benefits Among the Aged: Conflicting Measures, Unequal Distributions." *Social Security Bulletin* 59(3): 3–30.

ORES Working Paper Series
Recent Working Papers

Number	Date	Title	Author(s)
116	January 2021	Improving the Measurement of Retirement Income of the Aged Population	Irena Dushi and Brad Trenkamp
115	September 2020	Working and Claiming Behavior at Social Security's Early Eligibility Age Among Men by Lifetime Earnings Decile	Hilary Waldron
114	August 2020	Trends in Working and Claiming Behavior at Social Security's Early Eligibility Age by Sex	Hilary Waldron
113	June 2018	Correlation Patterns between Primary and Secondary Diagnosis Codes in the Social Security Disability Programs	Javier Meseguer
112	October 2015	A Multidisciplinary Review of Research on the Distributional Effects of Raising Social Security's Early Entitlement Age	Hilary Waldron
111	August 2008	The Out-of-Sample Performance of Stochastic Methods in Forecasting Age-Specific Mortality Rates	Javier Meseguer
110	December 2007	Cohort-Specific Measures of Lifetime Social Security Taxes and Benefits	Dean R. Leimer
109	October 2007	Social Security as a Retirement Resource for Near-Retirees, by Race and Ethnicity, Nativity, Benefit Type, and Disability Status	Benjamin Bridges, Sharmila Choudhury
108	October 2007	Trends in Mortality Differentials and Life Expectancy for Male Social Security–Covered Workers, by Average Relative Earnings	Hilary Waldron
107	June 2006	New Evidence on Earnings and Benefit Claims Following Changes in the Retirement Earnings Test in 2000	Jae G. Song, Joyce Manchester
106	May 2005	Social Security as a Retirement Resource for Near-Retirees	Benjamin Bridges, Sharmila Choudhury
105	May 2004	Heterogeneity in Health and Mortality Risk Among Early Retiree Men	Hilary Waldron

ORES Working Paper Series
Recent Working Papers

Number	Date	Title	Author(s)
104	April 2004	Using Predisability Earnings as an Instrument for Disability Determination Outcomes	Jae G. Song
103	March 2004	Young Widow(er)s, Social Security, and Marriage	Michael J. Brien, Stacy Dickert-Conlin, David A. Weaver
102	February 2004	Historical Redistribution Under the Social Security Old-Age and Survivors Insurance and Disability Insurance Programs	Dean R. Leimer
101	April 2003	Historical Redistribution Under the Social Security Old-Age and Survivors Insurance Program	Dean R. Leimer
99	December 2002	Mortality Differentials by Race	Hilary Waldron
98	August 2002	Social Security Disability Programs: Assessing the Variation in Allowance Rates	Alexander Strand
97	July 2002	Do Early Retirees Die Early? Evidence from Three Independent Data Sets	Hilary Waldron
96	June 2002	Social Security Benefit Reporting in the Survey of Income and Program Participation and in Social Security Administrative Records	Janice A. Olson
95	April 2002	Racial and Ethnic Differences in Wealth Holdings and Portfolio Choices	Sharmila Choudhury
94	September 2001	Income Growth and Future Poverty Rates of the Aged	Seyda G. Wentworth, David Pattison
93	August 2001	Links Between Early Retirement and Mortality	Hilary Waldron
92	June 2001	The Widow(er)'s Limit Provision of Social Security	David A. Weaver
91	June 2001	Methods in Modeling Income in the Near Term (MINT I)	Barbara A. Butrica, Howard M. Iams, James H. Moore, Mikki D. Waid