Is Retirement More Unequal Than We Think?

Estimates from the Survey of Income and Program Participation Linked to Administrative Records

Daniel Thompson
United States Census Bureau

May 6, 2021

SEHSD Working Paper 2021-10
SIPP Working Paper 296

CORRESPONDENCE TO:
Daniel Thompson
Program Participation and Income Transfers Branch
Social, Economic, and Housing Statistics Division
U.S. Census Bureau
4600 Silver Hill Road
Washington, DC
daniel.k.thompson@census.gov

ABSTRACT

A growing body of scholarship in the social sciences shows how household surveys – researchers' main source of income and poverty estimates – are affected by nonresponse and underreporting. Linking Wave 1 of the 2014 Survey of Income and Program Participation to Social Security and IRS administrative records, I examine respondents aged 65 and over and the accuracy of their survey responses for earnings, interest, dividends, social insurance, and retirement income. I find that retirement income – income from pensions and individual retirement accounts – is considerably underreported in SIPP. Substituting administrative records of retirement income increases household income 8.7%, 21.5%, and 27.2% at the 10th, 50th, and 90th percentiles, respectively. SIPP generally captures other forms of income well. The results show that income inequality among those aged 65 and over is higher than SIPP estimates suggest.

INTRODUCTION[1,2]

Household income surveys are the source of official income and poverty estimates for the United States, but a growing body of research in the social sciences has shown that reported income in surveys may be biased (Kim and Tamborini 2012a, 2012b; Tamborini and Kim 2013; Meyer, Mok, and Sullivan 2015; Bollinger et al. 2018; Parolin 2019). Whole survey non-response rates are growing, and even when sampled households do respond to surveys, respondents may decline particular items or report inaccurate values. Biases are often non-random, complicating imputations or adjustments. Far from being a narrow methodological concern, these biases affect national estimates of income and poverty. Likewise, these measurement problems are potentially a threat to the validity of research that aims to uncover the correlates and causal drivers of differential income levels and poverty rates.

One finding emerging from this research is that retirement income – income from pensions and individual retirement accounts (IRAs) – is among the most poorly captured forms of income in household surveys (Roemer 2000; Anguelov, Iams, and Purcell 2012; Munnell and Chen 2014). In a much-discussed working paper, Bee and Mitchell (2017) link the Current Population Survey Annual Social and Economic Supplement (CPS ASEC) – one of the main sources of income and poverty estimates for the United States – to administrative records of retirement income, finding that CPS ASEC captures less than half of income from pensions and IRAs. This underreporting biases estimates of total income to a large degree. Bee and Mitchell find that median income among respondents age 65 and over is 30% higher than in CPS ASEC estimates, while poverty rates are 30% lower.

Yet there is not consensus on the extent of retirement income measurement bias. Chen, Munnell and Sanzenbacher (2018) challenged Bee and Mitchell's findings, contending that retirement income bias is

[1] The U.S. Census Bureau reviewed this data product for unauthorized disclosure of confidential information and approved the disclosure avoidance practices applied to this release. CBDRB-FY21-POP001-001.

[2] The estimates in this report (which may be presented in the text, figures, and tables) are based on responses from a sample of the population and may differ from the actual values because of sampling variability and other factors. As a result, apparent differences between the estimates for two or more groups may not be statistically significant. All comparative statements have undergone statistical testing and are statistically significant at the 10 percent significance level, unless otherwise indicated. For further information on the source of the data and accuracy of the estimates, see <www2.census.gov /programs-surveys/sipp/tech-documentation/source-accuracy -statements/2014/sipp-2014-source-and-accuracy-statement.pdf>.

particular to CPS ASEC, and other household surveys perform better. Comparing survey and administrative estimates of total income (rather than linked individual data), Chen et al. (2018) find that other household income surveys capture a greater share of retirement income. Thus it is unclear to what extent the retirement income measurement problem affects other income surveys.

This paper builds on existing research on retirement income measurement in two ways. First, I examine how the Survey of Income and Program Participation (SIPP) compares to restricted-use administrative data from the Social Security Administration (SSA) and the Internal Revenue Service (IRS) linked at the individual level.[3,4] Like CPS ASEC which Bee and Mitchell (2017) examine, SIPP is a large, nationally representative household survey. But in contrast to CPS ASEC, SIPP has more detailed retirement income questions that aim to capture multiple spells of income receipt and lump sum payments, which may generate better estimates of retirement income. Second, the biases that I find in SIPP motivate greater attention to how income measurement and mismeasurement shape patterns of income inequality observed in SIPP. The forms of income that are most mismeasured – pension and IRA income – are among the most unequally distributed forms of income that Americans age 65 and older receive. If estimates of this income are biased, retirement income may be more unequally distributed than we think.

The paper is organized as follows: first, I discuss research on retirement income measurement and how it relates to estimation of income, poverty, and inequality. Second, I discuss the survey and administrative data used in the analysis and the linkage of these data. Third, I compare the survey and administrative estimates, finding that replacing SIPP estimates with administrative data on retirement income considerably increases household income at the middle and upper parts of the income distribution. In contrast, SIPP generally captures earnings, interest and dividends, Social Security, and Supplemental Security Income (SSI) well. I conclude by considering the distributional implications of retirement income mismeasurement.

BACKGROUND

Income mismeasurement and its implications for distributional estimates

A growing body of research has shed light on income mismeasurement. Meyer, Mok, and Sullivan (2015) provide an overview of the challenges facing household surveys, including misreporting and increasing item and whole survey non-response, and how these in turn complicate imputation. Meyer et al. find that each of these problems are growing in magnitude, concluding that household surveys are "in crisis". A series of Census Bureau working papers show that estimates of aggregate survey income often fall short of similar estimates from national accounts (Coder and Scoon-Rogers 1996, Roemer 2000, Rothbaum 2015). Bollinger et al. (2018) find that biases in CPS ASEC earnings reporting are concentrated at the tails of the earnings distribution. Kim and Tamborini (2012a, 2012b) and Tamborini and Kim (2013) likewise find bias in the earnings tails of SIPP and that this bias correlates with respondent and survey characteristics.

[3] More information about SIPP is available at <https://www.census.gov/programs-surveys/sipp.html>. The SIPP Data User's Guide is available at <https://www2.census.gov/programs-surveys/sipp/tech-documentation/methodology/2014-SIPP-Panel-Users-Guide.pdf>.

[4] See Westra and Brown (2017) for an analysis of non-response bias in Wave 1 of the 2014 SIPP Panel.

These are not narrow methodological concerns. Instead, these limitations potentially affect widely used distributional estimates to a large degree. The Gini coefficient – the most common summary income distribution measure – is sensitive to values in the tails (Gastwirth 2016), so if there are biases in the tails of the earning distribution, Gini coefficient estimates may be considerably biased (Thompson 2019). Meyer et al. (2015) find that more than half of income from some social programs is omitted from surveys, which should increase measured inequality. Parolin (2019) finds that correcting for underreporting reduces measured poverty.

Retirement income mismeasurement

A subset of this research has concluded that income surveys fail to capture a large share of retirement income, and that this mismeasurement may be increasing with changes to retirement planning (Roemer 2000; Bosworth, Burtless, and Anders 2007; Anguelov, Iams, and Purcell 2012; Iams and Purcell 2013; Munnell and Chen 2014; see also Brady 2020). Much of this concern stems from the fact that CPS ASEC, before its 2013 redesign[5], did not aim to capture irregular, often lump sum distributions from retirement accounts. Comparing total income in CPS ASEC and administrative sources, Anguelov, Iams, and Purcell (2012) find that CPS ASEC retirement income is far lower than administrative estimates. As defined contribution (DC) retirement plans and IRAs have become a larger part of retirement planning, biases stemming from the omission of irregular payments are likely to have increased. DC and IRA distributions are more likely to be irregular than the annuitized benefits that most defined benefit retirement plans provide. This feature of CPS ASEC may contribute to the omission of a considerable share of retirement income. However, even surveys that ask about lump sums may also fail to capture them. Iams and Purcell (2013) compare SIPP retirement income estimates to tax aggregates, concluding that SIPP – which does aim to capture such irregular distributions-- also greatly underestimates aggregate retirement income.

While most research on retirement income measurement compares data sources by summing income across all observations and comparing the totals, Bee and Mitchell (2017) build on this research by linking CPS ASEC with administrative records at the individual level. An advantage of their analysis is that linking individual data can not only examine total income in the two data sources as in earlier studies, but also how bias varies across the income distribution and its correlation with survey, respondent, and household characteristics. Confirming earlier research, Bee and Mitchell find considerable underreporting of retirement income, particularly distributions from IRAs, but Bee and Mitchell also find that this underreporting varies across the distribution.

This variation across the distribution introduces additional concerns. Bee and Mitchell (2017) find that the middle of the income distribution is the most affected by mismeasurement in the pre-redesign CPS ASEC. The authors provide evidence that smaller or more irregular distributions are reported less accurately, while larger or more regular distributions are reported more accurately. The net effect of these biases is that substituting administrative records for survey responses increases incomes,

[5] The CPS ASEC redesign included several changes to the retirement income questions. The redesign differentiates between income from retirement accounts and pensions. Respondents are asked about retirement account ownership, and if they report ownership, they are asked about distributed income from these accounts and whether this distribution was rolled over to another account. The redesigned asset section also distinguishes between asset income inside and outside retirement accounts. These changes were made to better capture irregular income flows from retirement accounts.

particularly at the middle of the income distribution. While this does not notably increase income inequality since middle incomes show the largest changes, it does increase median incomes and reduce poverty rates among the aged.

Research examining other surveys has been limited. Examining the American Community Survey linked to IRS data, O'Hara, Bee, and Mitchell (2016) find that survey-reported retirement, disability, and survivor income falls short of tax records, but the authors do not examine how this varies across the distribution. Brummet et al. (2017) perform a similar analysis of the Consumer Expenditure Survey, with similar findings. Bee and Mitchell (2017) and Bee and Mitchell (2018) use 1990-2008 SIPP panels linked to administrative data as a benchmark for analyses of CPS ASEC but do not compare the SIPP survey responses to administrative data due to differing reference periods.[6] Comparing aggregates rather than linked individual-level data, Chen, Munnell and Sanzenbacher (2018) examine total retirement income in CPS ASEC, SIPP, the Survey of Consumer Finances (SCF), and the Health and Retirement Study (HRS) in comparison to totals from the IRS Statistics on Income (SOI). They find that SIPP and HRS capture a considerably larger share of aged respondents' total income than CPS ASEC.[7] The authors attribute these differences to CPS ASEC's omission of irregular income flows prior to its redesign. Comparing unlinked individual-level data from HRS, CPS ASEC, SIPP, and IRS records, Choi et al. (2020) find that survey biases vary across the income distribution, again finding that CPS ASEC is the worst performer.

Only CPS ASEC has been used to examine how mismeasurement of retirement income affects different parts of the income distribution and to what extent income inequality changes when administrative data are used instead of survey responses. However, SIPP may perform better than the pre-redesign CPS ASEC, since SIPP collects data on multiple spells and includes a question specifically asking about lump sum retirement account distributions. SIPP may capture more income in aggregate, or perhaps biases affect SIPP differently than CPS ASEC. Discerning the extent and patterns of retirement income mismeasurement in SIPP is the objective of this paper.

DATA

Individually-linked survey and administrative data

This paper links survey responses from SIPP to administrative records, generally following the analytical approach of existing research linking surveys with administrative records. This analysis replicates much of the analysis of Bee and Mitchell (2017) but with SIPP rather than CPS ASEC. SIPP is a nationally representative, longitudinal survey measuring the income, employment, program receipt, and well-being of persons living in the United States. In this analysis, I examine the first wave of the 2014 SIPP panel (reference year 2013[8]), comparing SIPP to corresponding annual data from five administrative data sources.

I benchmark SIPP earnings data to the Social Security Administration's Detailed Earnings Record (DER).[9] The DER contains individual-level records of wage and salary earnings, tips, self-employment earnings,

[6] Before the 2014 redesign, SIPP had a four month reference period.

[7] My replication of Chen et al.'s analysis of total retirement income in SIPP yields substantially lower estimates. The reasons for the differences are unclear.

[8] SIPP interviews take place between March and May, and respondents are asked about the prior calendar year, so the 2014 SIPP covers reference year 2013.

[9] Appendix Table 1 lists the specific SIPP variables and corresponding administrative benchmarks.

and deferred compensation. SSA assembles these data from individual income tax records provided by the IRS. Notably, the DER omits employee payments to health insurance premiums, while SIPP estimates of gross earnings do include these payments.

I compare SIPP interest and dividends data to similar records from IRS Form 1040 records. The IRS records are at the tax-unit-level, which includes up to two adults and up to five dependents. Like Bee and Mitchell (2017), I assign all income to adults and none to dependents. Where there are two adults, I split the 1040 interest and dividend income equally between adults.

I benchmark SIPP data on Social Security income and Supplemental Security Income (SSI) to administrative records from the Social Security Administration on individual-level payments for these programs.

I compare SIPP data on retirement income to extracts from IRS Form 1099-R, "Distributions from Pensions, Annuities, Retirement or Profit-Sharing Plans, IRAs, Insurance Contracts, etc." For the SIPP data, I sum all retirement, disability, survivor, and life insurance income. The 1099-R data has information on two types of distributions, those from pensions and IRAs. Since these categories do not neatly map onto the SIPP categories, I sum the two IRS fields to generate an estimate of total retirement income. Importantly, the IRS 1099-R data omit retirement account distributions that are rolled over into another retirement account, such as when one withdraws funds from a 401(k) (often at retirement or the termination of employment) and immediately deposits these funds into an IRA.

I link the survey and administrative data with a Protected Identification Key (PIK). PIKs are generated by Census Bureau researchers that probabilistically link name, address, and geocoded location data from survey respondents to corresponding records associated with individual Social Security records. Once the survey and Social Security records are linked, it is possible to match survey respondents to a broad set of administrative data such as those examined here.

Not all survey respondents can be assigned a PIK. For Wave 1 of the 2014 SIPP panel, the PIK match rate is 91.5% for the unweighted sample. With inverse probability weighting, I reweight respondents to adjust for potential biases introduced by differential probability of being assigned a PIK. To do so, I estimate a logistic regression where the dependent variable is whether the respondent was assigned a PIK or not, and then I generate predicted probabilities for all respondents using survey and respondent characteristics as covariates. I divide the SIPP weights by the predicted probabilities, thereby scaling the survey weights upward for categories of respondents that are less likely to be assigned a PIK.

Analyses at the individual level examine respondents age 65 or older, while analyses at the household level examine households where the householder[10] is age 65 or older, even if these households include non-householder respondents under age 65. Where non-householders are not assigned a PIK and therefore cannot be linked with administrative records, I use their survey responses for the household-level analyses.

Household level analyses of income inequality often adjust household income for family composition, which accounts for resource sharing among household members. Without scaling, household level

[10] The householder is the person in whose name the housing unit is owned or rented. If there is no such person, any adult member, excluding roomers, boarders, or paid employees, may be specified as the householder. If the house is owned or rented jointly by a married couple, the householder may be either person.

inequality estimates are strongly influenced by variation in household composition. Scaling assumes that additional household members require a less-than-proportional increase of resources. I scale household income by household composition using the approach used in the Supplemental Poverty Measure (SPM) (Short 2015, see also Bee and Mitchell 2017), which assumes that costs to sustain adults are higher than children.[11]

RESULTS

Estimates of total income

Table 1 compares aggregate estimates from the SIPP survey responses to corresponding estimates from linked survey and administrative data. The first set of income estimates are for the SIPP survey responses. These are the amounts reported in SIPP weighted by the SIPP sample weights. The second set reports the SIPP survey responses but only for the SIPP respondents that have been assigned a PIK. For these estimates, I adjust the SIPP weights for the differential probability of receiving a PIK, as described in the previous section. The last set of estimates of total income where the SIPP responses have been replaced by individually-linked administrative records.

The estimates from survey responses in the SIPP sample and the PIK sample are similar; none of the differences reach statistical significance. This provides some assurance that the reweighted PIK sample is similar to the SIPP sample. Comparing the PIK sample and the administrative data, I find that earnings, interest and dividends, social security, and SSI are modestly overreported in the survey relative to administrative data. Retirement income, in contrast, is underreported, and the differences are substantively large. In the administrative data, retirement income is 116% higher than in the survey estimates. In other words, SIPP captures only 46.4% of retirement income recorded in administrative records. This accords with the findings of others who have also found underreporting of retirement income, while other types of income are generally reported more accurately (see, for example, Bee and Mitchell 2017).

Decile and poverty estimates

Table 2A examines how SIPP compares to administrative estimates at the 10th to 90th percentiles, replacing survey responses with administrative data one data source at a time.[12] Consistent with the aggregate estimates, decile income estimates are similar when earnings, interest and dividends, and SSI are replaced with administrative records, while Social Security is somewhat overreported. However, when retirement income is replaced with data from IRS Form 1099-R, much larger changes are

[11] Short's (2015) approach is a three-equation scale.

No children:
scale = $(\text{adults})^{0.5}$

Single parents:
scale = $(\text{adults} + 0.8*\text{first child} + 0.5*\text{other children})^{0.7}$

All other families:
scale = $(\text{adults} + 0.5*\text{children})^{0.7}$

[12] Appendix Table 2 reports estimates of retirement income by race and income in SIPP and administrative data.

observed, especially higher in the income distribution. Income increases 8.7% at the 10th percentile, but 21.5% at the 50th percentile and 27.2% at the 90th percentile.

Because the substitution of survey-reported retirement income with administrative records increases measured incomes to a greater extent at the middle and top of the income distribution than at the bottom of the income distribution, income inequality increases when administrative data are used. A common measure of household income is the 90/10 ratio, which is income at the 90th percentile divided by income at the 10th percentile. This captures the gap between high and low-income households. The 90/10 ratio grows from 6.74 in the survey-reported income to 8.19 in the administrative data, a 21.5% increase.

To get a sense of the magnitude of these changes, it is worth comparing them to the observed increase of the 90/10 household income ratio among all households over time, a concerning distributional trend that has received much attention among researchers. Horowitz, Igielnik, and Kochhar (2020) report that the 90/10 household income ratio increased 29% from 1980 to 2010.

The poverty rate also changes as survey responses are substituted with administrative data, but in opposing directions (see Table 2B). Substituting earnings and retirement data reduces the poverty rate, while substituting Social Security data increases the poverty rate, since this income is overreported in survey data. The combined effect of these changes on the poverty rate is negligible, as the last column shows.

Table 3A is similar to Table 2A, though while Table 2A replaces SIPP estimates with administrative data sources one-by-one, Table 3A replaces them cumulatively. This table shows that when survey responses for earnings, interest and dividends, Social Security, and SSI are replaced with administrative records, the total changes are small in comparison to the changes observed when replacing reported retirement income with administrative records. Median income increases 16.8% when retirement income is replaced with administrative records but falls 4.3% when all other administrative data sources are used.

These patterns observed in SIPP diverge somewhat from what Bee and Mitchell (2017) found in CPS ASEC. They found that substituting administrative data on retirement income increases incomes at the middle of the income distribution to a greater degree than the tails. Bee and Mitchell find that that CPS ASEC incomes increase 12.7%, 30.4%, and 14.5% at the 10th, 50th and 90th percentiles, respectively. The middle of the income distribution moves toward the top in Bee and Mitchell's analysis of CPS ASEC. In contrast, I find a nearly monotonic increase in SIPP where incomes increase 8.7%, 21.5%, and 27.2% at the 10th, 50th and 90th percentiles, respectively, when substituting administrative data on retirement. The middle and top of the income distribution pull away from the bottom.

Patterns of retirement income misreporting

Table 4A compares how income estimates change when replacing false positives, false negatives, and true positives. False positives are when there is reported income in the survey but no administrative data, while false negatives are cases in which there is no survey income reported when a non-zero administrative record exists. In contrast, true positives are those cases in which there is both survey reported income and a non-zero administrative record. Examining true positives sheds light on how reported amounts diverge from administrative records for true recipients that report income in SIPP.

There are small changes when replacing false positives. At the 80th percentile, for example, income differences are not statistically significant when false positives are removed. However, substantively large changes occur when either false negatives or true positives are replaced. Income increases 12.3% and 12.0% when false negatives and true positives are replaced, respectively. These results suggest that not only is retirement income often unreported in SIPP, but even when it is reported, the amount is often incorrect.

In Table 5, I examine how income reporting correlates with respondent and household characteristics. In Model 1, I present a linear probability model of retirement income reporting in SIPP among true positives in administrative records. The dependent variable is coded as 1 where a respondent reports retirement income in SIPP and 0 where the respondent does not. I find that wholly imputed interviews, marital separation, and 1099-Rs with only IRA income are the largest negative correlates, while interestingly, imputed retirement income is positively correlated with retirement income reporting, conditional on survey and respondent characteristics. Yet the most striking finding that emerges is that nonreporting of retirement income is only weakly correlated with respondent characteristics such as age and education. This suggests that misreporting is not confined to respondents from specific social categories and instead is generalized across the sample.

In Table 5 Model 2, I examine true positive responses. For respondents with an administrative record of retirement income who report any retirement income in the survey, I estimate a linear probability model of that survey response coming within 75-125% of the administrative value[13]. Most covariates are only weakly correlated with the dependent variable. Respondents with imputed retirement income and 1099-Rs with either just IRA income, or both pension and IRA income, are more likely to have differences greater than 25% between reported income and 1099-Rs. Like the previous model, I find that the accuracy of reporting is at most weakly correlated with respondent characteristics.

DISCUSSION

Recent research has shown that retirement income from pensions and IRAs is considerably underreported in household income surveys. My analysis of SIPP linked with administrative records finds a similar degree of underreporting in aggregate. SIPP captures just 46.4% of retirement income against a tax record benchmark. However, I find a somewhat different pattern of underreporting as bias increases monotonically at higher levels of household income. As measured by administrative records, inequality is higher than what is found in SIPP survey responses.

While the analysis sheds light on the extent of biases from retirement income underreporting, the exact causes of retirement income misreporting remain unclear. Regression analyses show that some respondent and household characteristics are associated with the accuracy of retirement income reporting. Among the strongest correlates of misreporting is receipt of IRA income, as captured by 1099-R records. This accords with Bee and Mitchell's findings that IRA distributions, which are more likely to be irregular, lump sum amounts, are less accurately reported (2017). As employer-sponsored – and especially defined-benefit – retirement accounts become a smaller part of retirement planning and IRAs take a larger role (Munnell and Chen 2017), this misreporting may be a growing problem. Retirement income may be increasingly difficult to measure in the future.

[13] The results are robust to alternative specifications.

Some researchers have suggested that retirement income underreporting is confined to CPS ASEC, and that it is largely an artifact of the omission of irregular income flows from the survey instrument before the 2013 CPS ASEC redesign (Chen et al. 2018). However, my analysis also finds considerable underreporting in SIPP, which does include a question about lump sum retirement account distributions. Differences between CPS ASEC and other surveys may not be a smoking gun.

Two characteristics of SIPP facilitate further study. One is that SIPP is a longitudinal survey, so individual-level variation in distributions and reporting can be examined over multiple years. Another is that audio recordings of some SIPP survey items are available. Using computer-assisted recorded interviewing (CARI), a sample of SIPP items are recorded during a subset of interviews for quality control purposes.[14] Examining these recordings may shed light on whether field representatives depart from interview scripts or respondents are confused by survey questions.

Improvements in survey methodology may attenuate misreporting in future surveys. The 2021 SIPP includes new content on employer-sponsored retirement accounts and IRAs, in which questions about account balances are immediately followed by questions about distributions. Asking about distributions at the same time as balances may improve reporting. Since SIPP is a longitudinal survey, differences in reporting in the new survey can be compared for the same respondents, potentially allowing researchers to examine whether these changes improve misreporting.

ACKNOWLEDGMENTS

I thank Adam Bee, Jennifer Bernard, Katherine Giefer, Mark Klee, Lindsay Monte, Josh Mitchell, Trudi Renwick, Jonathan Rothbaum, Rachel Shattuck, Mahdi Sundukchi, Christopher Tamborini, and Lewis Warren for their helpful advice and feedback on this project. I also thank Holly Fee and the staff in the Census Bureau's Center for Optimization and Data Science for technical assistance.

REFERENCES

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

Bee, Adam, and Joshua Mitchell. 2017. *Do Older Americans Have More Income Than We Think? SEHSD Working Paper*. 2017–39. United States Census Bureau.

Bee, C. Adam, and Joshua Mitchell. 2016. *The Hidden Resources of Women Working Longer: Evidence from Linked Survey-Administrative Data*. w22970. National Bureau of Economic Research.

Bollinger, Christopher R., Barry T. Hirsch, Charles M. Hokayem, and James P. Ziliak. 2018. "Trouble in the Tails? What We Know about Earnings Nonresponse 30 Years after Lillard, Smith, and Welch." *Journal of Political Economy* 127(5):2143–85. doi: 10.1086/701807.

Bosworth, Barry P., Gary Burtless, and Sarah Anders. 2007. *Capital Income Flows and the Relative Well-Being of America's Aged Population*. *Working Paper*. 2007–21. Center for Retirement Research.

Brady, David. 2020. "The Crucial Role of Measurement in the Study of Racial Inequality." *ASA IPM*

[14] Respondents may opt out of interview recording.

Newsletter 5(4):2–3.

Brummet, Quentin, Denise Flanagan-Doyle, Joshua Mitchell, John Voorheis, Laura Erhard, and Brett McBride. 2018. *Investigating the Use of Administrative Records in the Consumer Expenditure Survey*. 2018–01. U.S. Census Bureau.

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*. 2018–14. Center for Retirement Research.

Choi, James, Lucas Goodman, Justin Katz, and Shanthi P. Ramnath. 2020. "The Evolution of Late-Life Income and Assets: Measurement in IRS Tax Data and Three Household Surveys."

Coder, John, and Lydia Scoon-Rogers. 1996. *Evaluating the Quality of Income Data Collected in the Annual Supplement to the March Current Population Survey and the Survey of Income and Program Participation*. United States Census Bureau.

Gastwirth, Joseph L. 2017. "Is the Gini Index of Inequality Overly Sensitive to Changes in the Middle of the Income Distribution?" *Statistics and Public Policy* 4(1):1–11. doi: 10.1080/2330443X.2017.1360813.

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.

Kim, ChangHwan, and Christopher R. Tamborini. 2012a. "Do Survey Data Estimate Earnings Inequality Correctly? Measurement Errors Among Black and White Male Workers*." *Social Forces* 90(4):1157–81. doi: 10.1093/sf/sor042.

Kim, ChangHwan, and Christopher R. Tamborini. 2012b. "Response Error in Earnings: An Analysis of the Survey of Income and Program Participation Matched With Administrative Data." *Sociological Methods & Research*. doi: 10.1177/0049124112460371.

Meyer, Bruce D., Wallace K. C. Mok, and James X. Sullivan. 2015. "Household Surveys in Crisis." *Journal of Economic Perspectives* 29(4):199–226. doi: 10.1257/jep.29.4.199.

Munnell, Alicia H., and Anqi Chen. 2014. *Do Census Data Understate Retirement Income? Issue Brief*. 14–19. Center for Retirement Research.

Munnell, Alicia H., and Anqi Chen. 2017. *401(k)/IRA Holdings in 2016: An Update from the SCF. Working Paper* 17–18. Center for Retirement Research.

O'Hara, Amy, Adam Bee, and Joshua Mitchell. 2016. *Preliminary Research for Replacing or Supplementing the Income Question on the American Community Survey with Administrative Records*. 16–7. United States Census Bureau.

Parolin, Zachary. 2019. "The Effect of Benefit Underreporting on Estimates of Poverty in the United States." *Social Indicators Research* 144(2):869–98. doi: 10.1007/s11205-018-02053-0.

Roemer, Marc I. 2000. *Assessing the Quality of the March Current Population Survey and the Survey of*

Income and Program Participation Income Estimates, 1990 - 1996. United States Census Bureau.

Rothbaum, Jonathan L. 2015. *Comparing Income Aggregates: How Do the CPS and ACS Match the National Income and Product Accounts, 2007-2012*. *SEHSD Working Paper*. 2015–01. United States Census Bureau.

Short, Kathleen. 2015. *The Supplemental Poverty Measure: 2014*. P60-254. U.S Census Bureau.

Tamborini, Christopher R., and ChangHwan Kim. 2013. "Are Proxy Interviews Associated with Biased Earnings Reports? Marital Status and Gender Effects of Proxy." *Social Science Research* 42(2):499–512. doi: 10.1016/j.ssresearch.2012.11.004.

Thompson, Daniel. 2019. *Labor Taxation and the Uneven Growth of Wage Inequality*. Johns Hopkins University.

Westra, Ashley, and Faith Nwaoha-Brown. 2017. *Nonresponse Bias Analysis for Wave 1 of the 2014 Survey of Income and Program Participation (SIPP)*. U.S. Census Bureau.

Table 1: Total income by source (in millions USD), reference year 2013

	1		2		3	
Variable	**SIPP**	**MOE**	**SIPP PIK sample**	**MOE**	**Admin**	**MOE**
Earnings	502,700	51,240	502,400	54,700	473,000	38,660
Interest and dividends	106,500	17,520	109,000	18,070	102,500	18,090
Social Security	598,800	8,870	601,000	9,583	551,400	9,623
SSI	11,710	1,378	11,750	1,364	10,460	1,483
Retirement	292,800	12,540	294,000	13,010	634,000	31,010
Other	188,000	31,640	190,900	32,360	190,900	32,360

Source: Survey of Income and Program Participation, 2014 Panel, Wave 1; SSA Detailed Earnings Record; SSA OASDI payment records; SSI payment records; IRS 1040; IRS 1099-R.

Notes: MOE is the 90% margin of error. Estimates are equivalence-adjusted, household-year level for households where the household head is age 65 or older among the civilian, non-institutionalized population of the United States. SIPP estimates are survey-reported and imputed responses using the SIPP weights. SIPP PIK sample estimates are survey-reported and imputed estimates for those respondents assigned a Protected Identification Key (PIK). For the PIK sample, the SIPP weights are adjusted for the differential probability of being assigned a PIK. The Admin sample replaces survey-reported income estimates with administrative records of income from SSA and IRS for PIK sample respondents. See text for definitions of income from earnings, interest and dividends, Social Security, SSI, and retirement.

Table 2A: Income estimates replacing survey data with administrative records one income type at a time, reference year 2013

	Total income	Percentage differences						Total income
Measure	SIPP PIK sample	Replace earnings	Replace interest and dividends	Replace Social Security	Replace SSI	Replace retirement	Replace all	Replace all income types
P10	11,890	1.6%	0.0%	-6.8%	-0.5%	8.7%	0.2%	11,910
P20	16,770	0.9%	-0.1%	-5.1%	-0.6%	12.5%	6.9%	17,920
P30	21,460	0.0%	0.0%	-5.5%	-0.3%	15.5%	9.8%	23,570
P40	25,930	1.1%	-0.1%	-5.0%	-0.2%	19.8%	14.2%	29,600
P50	30,980	0.7%	-0.3%	-4.3%	-0.3%	21.5%	16.8%	36,170
P60	37,640	0.5%	-0.1%	-4.3%	-0.1%	20.5%	15.3%	43,410
P70	45,700	-0.1%	-0.2%	-4.2%	-0.1%	22.6%	18.2%	54,040
P80	57,310	-0.2%	-0.5%	-4.4%	0.0%	24.1%	19.8%	68,660
P90	80,140	-1.1%	-0.9%	-3.1%	0.0%	27.2%	21.7%	97,570
P90/P10	6.74	-2.7%	-0.9%	3.9%	0.5%	17.0%	21.5%	8.19

Table 2B: Poverty estimates replacing survey data with administrative records one income type at a time, reference year 2013

	Poverty rate						
Measure	SIPP PIK sample	Replace earnings	Replace interest and dividends	Replace Social Security	Replace SSI	Replace retirement	Replace all
Poverty rate	8.30%	7.86%	8.30%	9.59%	8.31%	6.79%	8.39%

Source: Survey of Income and Program Participation, 2014 Panel, Wave 1; SSA Detailed Earnings Record; SSA OASDI payment records; SSI payment records; IRS 1040; IRS 1099-R.

Table 4A: Income and poverty estimates replacing retirement survey data with false positives, false negatives, and true positives from administrative data, reference year 2013

	Total income	Percentage differences			
Measure	SIPP PIK sample	Replace false negatives	Replace true positives	Replace false positives	Replace all retirement income
P10	11,890	8.3%	0.4%	-0.2%	8.7%
P20	16,770	9.7%	1.1%	-0.2%	12.5%
P30	21,460	11.1%	2.6%	-0.2%	15.5%
P40	25,930	13.3%	4.9%	-0.1%	19.8%
P50	30,980	13.6%	6.9%	-0.1%	21.5%
P60	37,640	11.2%	8.6%	0.0%	20.5%
P70	45,700	10.8%	9.5%	0.0%	22.6%
P80	57,310	12.3%	12.0%	-0.1%	24.1%
P90	80,140	13.7%	15.2%	0.0%	27.2%
P90/P10	6.74	4.9%	14.7%	0.2%	17.0%

Table 4B: Income and poverty estimates replacing retirement survey data with false positives, false negatives, and true positives from administrative data, reference year 2013

	Poverty rate				
	SIPP PIK sample	Replace false negatives	Replace true positives	Replace false positives	Replace all retirement income
Poverty rate	8.30%	6.85%	8.12%	8.39%	6.79%

Source: Survey of Income and Program Participation, 2014 Panel, Wave 1; SSA Detailed Earnings Record; SSA OASDI payment records; SSI payment records; IRS 1040; IRS 1099-R.

Notes: Estimates are equivalence-adjusted, household-year level for households where the household head is age 65 or older among the civilian, non-institutionalized population of the United States. SIPP PIK sample estimates are survey-reported and imputed estimates for those respondents assigned a Protected Identification Key (PIK). For the PIK sample, the SIPP weights are adjusted for the differential probability of being assigned a PIK. The Admin sample replaces survey-reported income estimates with administrative records of income from SSA and IRS for PIK sample respondents. See text for definitions of income from earnings, interest and dividends, Social Security, SSI, and retirement.

Table 5: Linear probability models of retirement income reporting accuracy, reference year 2013

	Model 1: Any survey-reported retirement income conditional on IRS Form 1099-R retirement income receipt			Model 2: Survey-reported retirement income within 75-125% of IRS Form 1099-R retirement income.		
Variable	Coefficient	Standard Error		Coefficient	Standard Error	
Householding: Non-householder omitted						
Householder	0.035	0.016	*	-0.053	0.022	*
Sex: Male omitted						
Female	-0.013	0.019		-0.002	0.026	
Household size: Single person household omitted						
2 persons	-0.061	0.039		0.044	0.058	
3+ persons	-0.068	0.044		0.030	0.064	
Race: White omitted						
Black	-0.043	0.024		-0.011	0.032	
Asian	-0.154	0.044	*	-0.125	0.068	
Other	-0.009	0.058		0.112	0.080	
Hispanic origin: Non-Hispanic omitted						
Hispanic	-0.025	0.030		0.038	0.042	
Reporter: Self-report omitted						
Proxy	-0.058	0.021	*	-0.029	0.030	
Imputed interview	-0.339	0.047	*	-0.014	0.088	
Tenure status: Owner omitted						
Rent	0.023	0.022		0.014	0.029	
Occupy without rent	-0.023	0.046		0.092	0.071	
Marital status: Married with spouse present omitted						
Married with spouse absent	-0.146	0.087		-0.118	0.130	
Widowed	-0.169	0.066	*	-0.048	0.110	
Divorced	-0.185	0.067	*	-0.013	0.111	
Separated	-0.279	0.097	*	0.070	0.164	
Never married	-0.136	0.069	*	0.004	0.116	
Age: 65-74 omitted						
75-84	0.024	0.014		0.032	0.019	
85+	-0.004	0.020		-0.032	0.028	
Citizenship: Native born omitted						
Citizen, foreign-born	0.010	0.027		0.067	0.039	
Non-citizen	-0.059	0.078		-0.214	0.127	
Education: High school omitted						
Less than high school	0.001	0.022		0.003	0.030	
Some college	-0.036	0.016	*	-0.030	0.022	
Bachelor's degree or more	0.013	0.016		-0.014	0.022	

Type of household: Married couple omitted						
Male-headed family, no wife present	0.116	0.080		0.117	0.128	
Female-headed family, no husband present	0.128	0.068		0.049	0.114	
Male living alone	0.109	0.074		0.111	0.119	
Female living alone	0.099	0.074		0.093	0.119	
Veteran status: Non-veteran omitted						
Veteran	0.012	0.018		0.033	0.023	
Disability status: Non-disabled omitted						
Disability	0.025	0.014		0.010	0.019	
Region: Northeast omitted						
Midwest	0.017	0.019		0.065	0.025	*
South	-0.002	0.017		0.040	0.024	
West	0.008	0.020		0.040	0.027	
Residence location: Principal city of metro area omitted						
Metro area, not principal city	0.021	0.015		-0.020	0.020	
Nonmetro	0.038	0.016	*	-0.011	0.022	
Asset records: Did not use asset records for interview omitted						
Used asset records	0.114	0.013	*	0.053	0.017	*
Retirement income: self-reported omitted						
Imputed retirement income	0.225	0.017	*	-0.338	0.021	*
1099-R retirement income source: Pension only omitted						
IRA	-0.439	0.018	*	-0.248	0.042	*
Both pension and IRA	-0.043	0.015	*	-0.370	0.018	*
1099-R retirement income quartile: Q1 omitted						
Q2	0.123	0.019	*	0.073	0.031	*
Q3	0.259	0.020	*	0.107	0.031	*
Q4	0.250	0.020	*	0.085	0.031	*
Q5	0.295	0.022	*	-0.024	0.033	
Intercept	0.461	0.051	*	0.599	0.076	*
Unweighted N	5000			2900		
R^2	0.273			0.253		

Notes:

* P<.05

The unweighted N is rounded for disclosure avoidance.

Source: Survey of Income and Program Participation, 2014 Panel, Wave 1; SSA Detailed Earnings Record; SSA OASDI payment records; SSI payment records; IRS 1040; IRS 1099-R.

Estimates are person-year level for respondents age 65 or older among the civilian, non-institutionalized population of the United States. The SIPP weights are adjusted for the differential probability of being assigned a PIK. Federal surveys, including the 2014 SIPP Panel, give respondents the option of reporting more than one race. These data can be shown in two ways: (1) as mutually exclusive from other race groups, which may be denoted by "alone" or (2) not mutually exclusive with other race groups, denoted by "alone or in combination with other race groups". The first method is used in this report.

Appendix Table 1: SIPP variables and administrative data sources

Content	SIPP variable	Administrative benchmark
Earnings	RPEARN	SSA DER wages and self employment
Interest Dividends	TINC_BANK TINC_BOND TINC_STMF	1040
Social security	ESSSAMT	SSA PHUS OASDI benefits
SSI	RSSI_AMT	SSA SSR federal and state benefits
Defined benefit pensions, 401(k), 403(b), traditional IRA, Roth IRA, Keogh, SEP, and other retirement income	ERET1AMT--ERET8AMT EDIS1AMT--EDIS10AMT ESUR1AMT--ESUR13AMT ELIFEAMT ELMPAMT (excluding EROLLOVER and non-retirement lump sum payments)	1099R
Other income	RPTOTINC minus variables above	No administrative benchmark

Appendix Table 2: Income deciles in SIPP and administrative data for white and black respondents, reference year 2013

	White respondents			Black respondents				
Measure	SIPP PIK sample	Replace retirement	Percentage difference	SIPP PIK sample	Replace retirement	Percentage difference	Black-white ratio in SIPP	Black-white ratio in admin data
P10	12,790	12,740	-0.4%	9,589	8,758	-8.7%	0.75	0.69
P20	17,720	19,200	8.4%	12,250	12,230	-0.2%	0.69	0.64
P30	22,650	24,740	9.2%	15,320	16,110	5.2%	0.68	0.65
P40	27,130	30,880	13.8%	19,050	20,900	9.7%	0.70	0.68
P50	32,260	37,650	16.7%	22,750	25,090	10.3%	0.71	0.67
P60	38,930	45,120	15.9%	28,090	32,280	14.9%	0.72	0.72
P70	47,250	56,140	18.8%	33,160	39,960	20.5%	0.70	0.71
P80	59,120	71,430	20.8%	42,010	52,450	24.9%	0.71	0.73
P90	82,610	101,700	23.1%	55,220	70,610	27.9%	0.67	0.69
P90/P10 ratio	6.46	7.98		5.76	8.06			
P80/P20 ratio	3.34	3.72		3.43	4.29			

Source: Survey of Income and Program Participation, 2014 Panel, Wave 1; IRS 1099-R.

Notes: Estimates are equivalence-adjusted, household-year level for households where the household head is age 65 or older among the civilian, non-institutionalized population of the United States. SIPP PIK sample estimates are survey-reported and imputed estimates for those respondents assigned a Protected Identification Key (PIK). For the PIK sample, the SIPP weights are adjusted for the differential probability of being assigned a PIK. The Admin sample replaces survey-reported income estimates with administrative records of income from SSA and IRS for PIK sample respondents. See text for definitions of income from earnings, interest and dividends, Social Security, SSI, and retirement. Federal surveys, including the 2014 SIPP Panel, give respondents the option of reporting more than one race. These data can be shown in two ways: (1) as mutually exclusive from other race groups, which may be denoted by "alone" or (2) not mutually exclusive with other race groups, denoted by "alone or in combination with other race groups". The first method is used in this report.