



The Construction of Income Data from the Survey of Consumer Finances: 1983 to 2019

John F. Cogan and Daniel L. Heil*

Economics Working Paper 22126

HOOVER INSTITUTION
434 GALVEZ MALL
STANFORD UNIVERSITY
STANFORD, CA 94305-6010

November 2022

This paper provides a technical overview of the methods used to construct a consistent set of income variables across multiple years of the Survey of Consumer Finances. These variables are used in *Changes in Income Among Senior Citizens: 1982-2018* (Cogan and Heil, 2022), which analyzes trends in senior incomes both in absolute terms and relative to younger households. We provide detailed descriptions of the methods used to address shortcomings in the data and offer additional comparisons between measurements of senior income in the Survey of Consumer Finances and the Current Population Survey.

Keywords: Senior incomes, elderly employment, retirement income, defined contribution retirement plans, defined benefit retirement plans, Survey of Consumer Finances, Current Population Survey

JEL Codes: D31, E21, E24, J14, J32, H55

The Hoover Institution Economics Working Paper Series allows authors to distribute research for discussion and comment among other researchers. Working papers reflect the views of the authors and not the views of the Hoover Institution.

* Senior Fellow and Policy Fellow, the Hoover Institution, Stanford University.

Introduction

This paper provides a technical overview of the data and methods used to construct income variables for the analysis in Cogan and Heil (2022). That paper uses the Survey of Consumer Finances (SCF) to analyze the trend in incomes of households headed by persons age 65 and older (termed senior households) in both absolute terms and relative to younger households (termed non-senior households) from 1982 to 2018. It supplements the findings from the SCF with data from the Current Population Survey (CPS). This paper provides more detailed descriptions of the methods used to address shortcomings in SCF data and offers additional comparisons between measurements of senior income in the SCF and CPS.

The SCF is conducted on a moderately sized sample of U.S. households at three-year intervals. It is designed to capture information on asset holdings and income from assets, as well as other sources. Since senior households rely heavily on income from retirement plans and assets held outside of retirement plans, the SCF is particularly well-suited to studying the behavior of their incomes. The survey also provides information about Social Security benefits that allows analyses of various issues relating to Social Security policy. Indeed, the survey has been referred to as the “gold standard” for the measurement of senior household income and wealth (Chen et al., 2018).¹ Earlier work by Czajka and Denmead (2008, 2011, 2012) also finds that the SCF paints a far more accurate measure of retirement income than the CPS. Similarly, Dettling, et. al (2015) reports that aggregate incomes from the SCF closely match aggregate income from National Income and Product Accounts.

The SCF, however, suffers from shortcomings that limit its usefulness in analyzing senior income trends over time. First, a significant survey redesign in 1989 altered the way certain income information was collected compared to earlier years. Second, beginning in 1989, deficiencies in the way income sources are reported limit researchers’ ability to disaggregate income growth by source. Third, prior to 2004, the survey did not include specific questions about withdrawals from defined contribution retirement plans. This paper describes our methodology for addressing these issues and the results derived from it. An outcome of this work is a dataset composed of SCF survey data from 1983 to 2019—corresponding to income data from 1982 to 2018—that takes advantage of the

¹ Chen, et al (2018) finds that the 2016 SCF captures 98 percent of aggregate income among seniors, 99 percent of retirement plan income, 106 percent of interest and dividends, and 95 percent of Social Security income.

survey’s detailed income measures and contains data imputations that address the aforementioned shortcomings.²

We compare our SCF estimates of senior income levels and trends to those obtained from the CPS. The CPS is the federal government’s primary source of income, poverty, and employment data and is widely used for U.S. income analysis. The survey, however, suffers from well-documented shortcomings that limit its effectiveness in estimating senior incomes. The most important of these shortcomings is that it is designed primarily to capture “regular” income sources. As a result, the survey fails to capture a significant portion of income from defined contribution retirement plans and income from assets held outside of retirement plans. While a major survey redesign in 2014 reduced the degree of underreporting, the problem remains (Copeland 2016). Despite their differences and shortcomings, the SCF and CPS tell similar stories about senior income trends. This finding bolsters our confidence in both surveys—especially the CPS—as sources for estimating long-term trends in senior incomes.

The paper is organized as follows. Section 1 provides an overview of the SCF sample and highlights changes to the sample and unit of analysis over time. The SCF household definition differs from the more conventional definition used by the CPS and other surveys. The household definition within the SCF also differs between the 1983 and 1989–2019 surveys. This section assesses the importance of these differences in determining the number and composition of households headed by persons age 65 and older and those headed by persons under age 65. Section 2 discusses the various asset and income data available in the SCF and the potential issues with these data sources. Section 3 describes the methods we used to correct for two SCF shortcomings in measuring income from Social Security and private defined benefit plans. The first is a significant underreporting of Social Security benefit reciprocity in the 1983 and 1994–2001 surveys. The second is that the SCF combines prior-year Social Security income and private pension income into a single variable in the 1989–2019 surveys. Section 4 discusses the SCF’s underreporting of income from defined contribution retirement plans in the 1983–2001 surveys, and the method used to correct for it. The section also reports the results from using this method. Section 5 reports the impact of our income adjustments on household income and shows the sensitivity of the SCF household

² See <https://www.federalreserve.gov/econres/aboutscf.htm> for a description of the survey and links to the public use files and related codebooks. Our modified dataset is available at <https://www.dropbox.com/s/mxp2c4w6p0bz323/scf.work.withimp.ind.replicate.dta?dl=0>

definition to the more conventional definition. Section 6 offers some concluding thoughts on the value of the SCF as a tool for measuring changes in senior household income over time.

1 Overview of the SCF Sample

The SCF provides detailed information on prior-year income and current-year assets, as well as household demographic characteristics.³ We use surveys from 1983 to 2019, except the 1986 survey. The 1986 survey was excluded because it was designed as a supplement to the 1983 survey and re-interviewed respondents to that earlier survey.⁴ Here we provide an overview of the sample, discuss how it compares to the CPS, and examine how changes to the survey’s unit of analysis may affect reported income trends. Throughout the paper, when we discuss issues involved with the surveys themselves, we will refer to survey years (1983, 1989, 1992, ..., 2019). When discussing income data, we will refer to the year for which the data are reported, i.e., the year prior to the survey (1982, 1988, 1991, ..., 2018).

1.1 Senior and Non-Senior Households

We divide households into senior and non-senior households according to the age of the household head. Senior households are those in which the head is age 65 or older.⁵ Correspondingly, non-senior households are those in which the head’s age is less than 65. Table 1 reports the sample sizes of senior households and non-senior households by year. The number of senior households is not large, especially in the survey’s early years. The number of senior households surveyed ranges from a low of 678 in the 1983 survey to a high of 1,449 in the 2016 and 2019 surveys.⁶ The number of non-senior households, on the other hand, is considerably larger, ranging from 2,451 in 1988 to 5,226 in 2009.

³ Financial information is collected from a single respondent within the dwelling unit who SCF interviewers determined was “most knowledgeable” about household financial matters. The respondent is not necessarily the head of household. For a discussion of this issue, see Lindamood, Hanna, and Bi (2007).

⁴ The 1989 survey also reinterviewed 1983 survey participants that had not moved but the survey added new participants for cross-sectional analysis.

⁵ The survey’s age question is as of the time of the survey. Since our primary focus is on income received in the year prior to the survey, we construct a new age variable using the head’s reported birth year to set the head’s age as of December 31 in the year prior to the survey.

⁶ The 1983 survey also interviewed a special sample selected among high-income tax filers. Our estimates do not include this sample. We also exclude households who are missing particular income sources.

Table 1. SCF sample size

	Senior heads	Non-senior heads
1983	678	2981
1989	690	2451
1992	855	3049
1995	889	3408
1998	863	3439
2001	872	3568
2004	841	3669
2007	905	3487
2010	1200	5226
2013	1276	4711
2016	1449	4748
2019	1449	4293

Notes: Sample sizes exclude observations with missing income values.

The wide variance in the distribution of senior income is important for public policy analysis (Poterba 2014). Unfortunately, the SCF’s relatively small number of observations on senior households limits its value in examining distributional changes. The sample sizes are sufficiently large to allow estimation of income and asset distributions across quartiles of the senior household income distribution, but a finer breakdown of the senior income distribution or of income sources within quartiles produces statistically unreliable results. As a consequence, Cogan and Heil (2022) use quartiles to examine changes over time in the distribution of senior household income. When examining the importance of individual sources of income and demographic factors, however, the survey’s small sample size limits our distributional analysis to senior households with incomes above and below the median senior household income.

Table 2 compares population estimates of the number of all U.S. households, and the number of senior and non-senior households between the SCF and the CPS. The total number of households in the two surveys match each other closely. The SCF has fewer senior households, although this difference is minor except for the 1983 survey and, perhaps, the 2019 survey. The SCF correspondingly has more non-senior households by a similarly small margin. These differences may be due to how each survey assigns the head. In the CPS, the head (called “householder” in the CPS) is the person in “whose name the housing unit is owned or rented or, if there is no such person, any

adult member, excluding roomers, boarders, or paid employees.”⁷ In the SCF, the head (now called reference person in the SCF) is the “economically dominant” person in the household.⁸ Despite these differences, the SCF is generally in accord with the CPS, except for the unexplained 2.1 million fewer SCF senior households in 1983.

Table 2. Household counts in SCF and CPS (millions)

	All households			Senior households			Non-senior households		
	SCF	CPS	% Diff	SCF	CPS	% Diff	SCF	CPS	% Diff
1983	83.8	83.9	-0.1%	15.7	17.7	-11.6%	68.2	66.2	3.0%
1989	93.0	92.8	0.2%	19.0	19.8	-4.1%	74.0	73.0	1.4%
1992	95.9	95.7	0.3%	20.2	21.0	-4.2%	75.8	74.6	1.5%
1995	99.0	99.1	-0.1%	20.6	21.6	-5.0%	78.4	77.4	1.3%
1998	102.5	102.6	0.0%	21.3	22.0	-3.1%	81.2	80.6	0.8%
2001	106.5	106.5	0.0%	21.7	22.5	-3.6%	84.8	84.0	0.9%
2004	112.0	112.1	-0.1%	23.0	23.8	-3.1%	89.0	88.4	0.7%
2007	115.9	116.1	-0.2%	23.5	24.5	-4.3%	92.4	91.6	0.9%
2010	117.3	117.6	-0.3%	25.1	26.1	-4.0%	92.2	91.5	0.8%
2013	122.4	122.5	-0.1%	28.0	28.9	-3.0%	94.3	93.6	0.7%
2016	125.7	126.1	-0.3%	30.7	31.9	-3.8%	95.1	94.2	1.0%
2019	128.5	128.7	-0.2%	32.7	35.1	-6.9%	95.8	93.6	2.3%

Table 3 compares the distribution of senior households by household type between the two surveys. Both surveys tell a similar story. About half of all senior households are married couples (including unmarried partners) and this fraction shows no trend over time. The SCF share of households that are married couples is slightly higher in the SCF than in the CPS in most years. The rather large differences in 1983 and 2001 are noteworthy and are unexplained exceptions. The proportion of senior households headed by single females and males in the two surveys track each other well in most years. Both surveys show a marked decline over time in the share of senior households that are headed by single females and a corresponding rise in the share that are headed by single males.

⁷ See <https://www.census.gov/programs-surveys/cps/technical-documentation/subject-definitions.html#householder> for a full definition.

⁸ In married households, the SCF assigns the male partner as the reference person (in same sex households, the older partner is chosen). In the CPS, the householder can be either the male or female partner if the house is owned or rented in both of their names. To ensure consistency across the surveys, however, we reassign householder status in the CPS to the male partner in all married households.

Table 3. Share of senior households by household type (SCF and CPS)

	Married couples			Single female heads			Single male heads		
	SCF	CPS	% Diff	SCF	CPS	% Diff	SCF	CPS	% Diff
1983	52%	45%	7%	39%	43%	-4%	9%	12%	-3%
1989	46%	45%	1%	44%	43%	1%	11%	12%	-2%
1992	44%	44%	-1%	44%	43%	0%	13%	12%	0%
1995	48%	44%	4%	41%	43%	-1%	11%	13%	-2%
1998	47%	44%	3%	43%	43%	0%	10%	13%	-3%
2001	53%	45%	8%	34%	42%	-7%	13%	14%	-1%
2004	49%	44%	4%	38%	41%	-4%	14%	14%	-1%
2007	47%	44%	3%	41%	41%	0%	12%	15%	-3%
2010	50%	45%	5%	37%	40%	-3%	13%	15%	-2%
2013	49%	46%	3%	38%	37%	1%	13%	16%	-4%
2016	51%	45%	5%	33%	38%	-5%	16%	17%	-1%
2019	49%	46%	3%	35%	36%	-1%	16%	18%	-2%

Table 4 shows the same distributions among non-senior households. Both surveys show a decline in married couples as a percentage of all non-senior households and an offsetting rise in households headed by single individuals of both sexes. The magnitude of these trends is greater in the CPS. There are at least two reasons for the differences in trends between surveys. First, the SCF treats couples living as partners as married regardless of their legal marital status. Second, as discussed above, the surveys have different methods for selecting the head of household.

Table 4. Share of non-senior households by household type (SCF and CPS)

	Married couples			Single female heads			Single male heads		
	SCF	CPS	% Diff	SCF	CPS	% Diff	SCF	CPS	% Diff
1983	67%	63%	4%	20%	22%	-2%	12%	14%	-2%
1989	61%	59%	2%	24%	24%	0%	14%	17%	-2%
1992	61%	58%	3%	23%	25%	-2%	15%	17%	-2%
1995	61%	57%	4%	25%	25%	0%	14%	18%	-4%
1998	62%	55%	6%	23%	26%	-3%	15%	19%	-4%
2001	62%	54%	8%	24%	26%	-2%	14%	20%	-6%
2004	60%	53%	7%	25%	27%	-2%	15%	20%	-5%
2007	62%	52%	9%	24%	27%	-3%	15%	21%	-6%
2010	60%	51%	9%	24%	27%	-4%	16%	22%	-6%
2013	60%	49%	11%	24%	28%	-4%	16%	23%	-7%
2016	59%	49%	10%	25%	28%	-4%	17%	23%	-6%
2019	58%	49%	9%	23%	28%	-5%	18%	23%	-5%

The similar household composition between the surveys is not surprising, as the SCF survey weights are in part based on population estimates from the CPS. Nevertheless, given the difference in the definition of households between the two surveys and CPS's use as the standard survey for measuring the demographic characteristics of the U.S. population, the similarity is reassuring for the SCF.

1.2 Unit of Analysis

The SCF's "unit of analysis" is used to define the household for the purposes of our study. The SCF's unit of analysis differs slightly between the 1983 and subsequent surveys. The 1983 unit of analysis, termed a "primary family," consists of all persons living together who are related by blood, marriage, or adoption, i.e., the household head, spouse, and their relatives.⁹ In the 1989–2019 surveys the unit of analysis, termed a primary economic unit (PEU), consists of all household members living together who are financially interdependent with the household head.¹⁰ Thus, in the 1983 survey, the unit of analysis includes some relatives who are financially independent of the head, while the 1989–2019 surveys exclude such persons. Conversely, in the 1989–2019 surveys, the unit of analysis may include non-relatives who are financially dependent on the household head, while the 1983 survey would exclude such persons.

Upon inspection of the 1989–2019 survey data, we find that very few PEU members are not relatives of the PEU head. The share of PEU members that are not relatives of other PEU members averages only 0.3 percent and never exceeds 0.5 percent. Unfortunately, the SCF does not report the income of these non-relatives separately. If their income had been excluded from PEU income in the 1989–2019 surveys, as it was in the 1983 survey, it would presumably have had only a minor impact on household income.¹¹

In the 1989–2019 surveys, relatively few relatives of PEU members were excluded from the PEU. If these financially independent relatives were included, their inclusion would increase the

⁹ The primary family could be a one-person unit (if the respondent did not live with relatives) as well as a unit of two or more individuals. Only one family in a dwelling unit was interviewed. Sub-families are excluded.

¹⁰ The 1989–2019 surveys define the unit of analysis as an "economically dominant single individual or couple (married or living as partners) in a household and all other individuals in the household who are financially interdependent with that individual or couple."

¹¹ Officially, the PEU annual income estimates should include all PEU members, but Gale, et al. (2022) generally find that reported income from the PEU consists only of the household head and the spouse if present. Thus, the actual effect on household income from including some non-relatives should be minimal.

number of PEU members only by 5 percent to 7 percent in any given year.¹² If their income were included in the PEU's income, as it was in the 1983 survey, it would have increased household income, which would cause our analysis to understate the growth in senior household income. As we show in section 5, the increase is minor for both senior and non-senior households. Thus, as a practical matter, the different SCF definitions of the unit of analysis between the 1983 survey and the 1989–2019 surveys matter little for analysis of income changes over time.

Our definition of senior households excludes some seniors. First, it excludes seniors who are not part of the primary family in the 1983 survey or the PEU in the 1989–2019 surveys.

Unfortunately, the 1983 survey provides financial and demographic information only on persons who are included in the primary family. The 1989–2019 surveys, on the other hand, provide age, sex, and limited financial information on all individuals not in the PEU. Table 5 uses this information to determine how many seniors are included in our analysis. Column 1 shows the total number of seniors in each survey. Column 2 shows the share of seniors that remain in the SCF sample of households after eliminating those who are excluded by the survey's household definition. From 1989 to 2019, the SCF definition captures between 92 percent and 96 percent of all seniors.

Table 5. Number and share of seniors in PEUs (1989 to 2019)

	Total seniors	Seniors in PEU	Share in PEU
1989	29,208,765	26,939,923	92%
1992	29,770,659	27,499,995	92%
1995	30,269,654	28,185,572	93%
1998	32,651,206	30,483,551	93%
2001	33,591,645	31,704,553	94%
2004	33,707,656	31,984,180	95%
2007	34,527,258	32,520,212	94%
2010	37,450,780	35,474,682	95%
2013	41,915,706	39,755,545	95%
2016	47,299,275	45,274,640	96%
2019	50,306,438	47,451,822	94%

Notes: The 1983 survey did not include any information on household members who were not part of the primary family.

¹² The 1983 SCF codebook suggests that the income of financially independent family members may not be included in its measure of 1982 income. The code book states: "This income measure, which was constructed for all survey households using reported 1982 income data, is roughly comparable to the IRS measure of adjusted gross income plus excluded realized capital gains." Although "financial independence" is in the mind of the beholder, it is no stretch to think that financially independent family members would file their own IRS returns. As noted above, Gale, et al (2022) reach a similar conclusion for more recent survey years.

Second, senior households do not include seniors who live in households headed by persons under age 65. This could include spouses married to heads under age 65 and financially dependent parents of the head or spouse. Table 6 provides a breakdown of seniors according to whether they are included in a senior-headed PEU, a non-senior PEU, or are not in a PEU. As shown in the table, around 90 percent of all seniors are included in senior households in each year.¹³

Table 6. Number and share of seniors by PEU status (1989 to 2019)

	Senior-headed households		Non-senior households		Not in PEU	
	# (millions)	Share	# (millions)	Share	# (millions)	Share
1983	21.5	89%	2.7	11%	n/a	n/a
1989	26.0	89%	1.0	4%	2.3	8%
1992	26.4	89%	1.1	4%	2.3	8%
1995	27.3	90%	0.9	3%	2.1	7%
1998	29.4	90%	1.1	4%	2.2	7%
2001	30.5	91%	1.2	4%	1.9	6%
2004	30.7	91%	1.3	4%	1.7	5%
2007	30.8	89%	1.7	5%	2.0	6%
2010	33.7	90%	1.7	5%	2.0	6%
2013	37.5	90%	2.2	6%	2.2	5%
2016	42.8	90%	2.5	5%	2.0	4%
2019	45.1	90%	2.4	5%	2.9	6%

Notes: The 1983 survey did not include any information on household members who were not part of the primary family.

2 SCF Household Income and Asset Data

The SCF reports household income received during the calendar year prior to the survey by the primary family in 1983 and the PEU in the 1989–2019 surveys. Income sources that are reported include labor and self-employment income, business income, Social Security, income from employer sponsored defined benefit and defined contribution pension plans, Individual Retirement Account withdrawals, government transfer benefits, alimony payments, and income from various non-retirement plan assets, including interest, dividends, rents, royalties, and realized capital gains. A major problem in the 1989–2019 surveys is that prior-year annual income from Social Security and private defined benefit pension plans are combined into a single variable. This issue is addressed

¹³ While about 90 percent of seniors are in senior-headed households, a small share of this group are neither the head nor the spouse in the PEU. The SCF does not capture detailed income and asset information for these PEU members.

below. Each survey also includes detailed information on financial assets and liabilities, including home values and mortgages. The value of these assets is recorded as of the time of the survey.

The surveys also report separate income information for heads of households and, in the case of married couple households, spouses. These income variables are generally reported as of the time of the survey. This information includes whether each individual is receiving Social Security and the monthly amount received, each individual's employment status, the number of hours usually worked per week, and the number of weeks worked during the prior year. These "current" data permit us to address several shortcomings in the SCF. The "current" Social Security information allows us to separate the combined prior-year Social Security income from defined benefit pension income. Similarly, "current" weeks worked and the wage information allows us to separately account for the individual contributions of husbands and wives to changes in household earnings. Importantly, these data are generally not collected for other members of the PEU, thus most of analysis is focused exclusively on heads and spouses.

The 1989–2019 surveys report only a limited amount of information on financially independent household members (i.e., household members not included in the PEU). The only demographic information reported is the age and sex of these members. Income received by these individuals is reported, but all income except earnings is pooled into a single variable.¹⁴ This information allows us to evaluate how using the PEU definition of the household rather than the broader definition, which includes the income of all persons in a dwelling unit, affects our conclusions. This evaluation is taken up in section 5. Unfortunately, the 1983 survey does not provide any household income information for members who are not part of the primary family.

3 Imputing Missing Social Security Benefits

As noted above, the 1989–2019 SCF surveys combine prior-year Social Security benefits with prior-year income from private defined benefit pensions into a larger "pension income" category. The 1983 survey does not have this problem. Like the 1989–2019 surveys, the 1983 survey includes a combined Social Security and pension variable, but it also includes prior-year Social Security income received for the head and spouse as a separate variable. In contrast, the 1989–2019 surveys only report monthly Social Security and employer provided pension benefits received by the household head and, if present, the spouse at the time of the survey. The survey also reports the

¹⁴ The data also include indicator variables that identify the sources of the pooled income variable but not the amount from each source.

number of years each person has been receiving both benefits. We use these monthly benefits to obtain separate estimates of the household's annual Social Security and defined benefit pension income during the prior year. Since the person-level monthly Social Security benefit is the starting point for our Social Security income estimates, some assessment of the accuracy of these monthly benefits is warranted before describing the method we employed.

3.1 Individual-Level Social Security Reciprocity and Benefits

Our assessment involves comparisons of SCF data to administrative data from the Social Security Administration. Social Security Administration benefit data are only available for individuals, irrespective of the composition of their household. Thus, our comparison is at the individual, as opposed to the household level. The administrative data include all individuals age 65 and older.¹⁵ The SCF comparison group is slightly different. Social Security reciprocity and benefit data are only reliably collected for the head and, if present, their spouse. Thus, our senior SCF group consists of heads and their spouses age 65 and older in a PEU. As shown in table 7, this group accounts for 87 to 92 percent of all seniors included in each survey.

Table 7. Share of seniors by relationship to head

	Seniors in sample	Share of seniors		
		Head	Spouses	All Others
1983	24,201,434	65%	23%	13%
1989	29,208,765	65%	25%	10%
1992	29,770,659	68%	22%	11%
1995	30,269,654	68%	23%	9%
1998	32,651,206	65%	25%	10%
2001	33,591,645	65%	27%	8%
2004	33,707,656	68%	23%	9%
2007	34,527,258	68%	23%	9%
2010	37,450,780	67%	24%	9%
2013	41,915,706	67%	24%	9%
2016	47,299,275	65%	26%	9%
2019	50,306,438	65%	26%	9%

Notes: The 1983 survey did not include any information on household members who were not part of the primary family.

¹⁵ The administrative data on the percentage of seniors who receive Social Security and on average benefit levels are derived from the Annual Statistical Supplements of the Social Security Bulletin. We use the supplements from the year of the data, which report enrollment and benefit data from December of the prior year.

Table 8 compares the share of seniors who receive Social Security and the average benefit among recipients in the SCF to Social Security Administration (SSA) data in each survey year from 1983 to 2019. The participation rate in administrative data exhibits no trend and hovers around 90 percent. The SCF participation rates closely approximate and slightly exceed the administrative rates in all years except 1983 and 1995–2001. In these years, the SCF substantially undercounts the administrative participation rates. Average benefit levels in the SCF, on the other hand, closely approximate the administrative benefit levels in all years.¹⁶

Table 8. Social Security reciprocity and benefits: SCF and Social Security Administration

	Participation rate			Mean benefit		
	SCF	SSA	Difference	SCF	SSA	% Difference
1983	81%	90%	-9%	\$11,100	\$11,250	-1%
1989	93%	91%	3%	\$11,450	\$11,850	-3%
1992	92%	91%	1%	\$12,300	\$12,350	0%
1995	84%	91%	-7%	\$12,650	\$12,750	-1%
1998	85%	91%	-6%	\$12,700	\$13,250	-4%
2001	87%	92%	-5%	\$13,250	\$13,950	-5%
2004	95%	92%	3%	\$14,200	\$14,450	-2%
2007	93%	91%	3%	\$14,550	\$15,150	-4%
2010	94%	90%	4%	\$15,350	\$16,150	-5%
2013	92%	90%	2%	\$16,150	\$16,550	-2%
2016	93%	89%	4%	\$17,200	\$17,100	1%
2019	93%	89%	4%	\$17,550	\$17,750	-1%

Notes: SSA data are for the month of December of the prior year. The administrative participation rates use population estimates from the Social Security Trustees annual report (see table V.A3 in the Trustees supplemental tables.). We use the mean of the survey year and the previous year to estimate the over-65 population as of December of the year prior to the survey. Income data are adjusted for inflation using the PCE price index. Data from the 1983 survey corresponds to the level of benefits received in 1982.

Except for the 1983 and 1995–2001 surveys, the SCF appears to provide a good measure of Social Security income and participation rates for the 65 and older population. The substantial undercount in participation rates during the excepted years, however, is concerning and warrants further examination. Table 9 sheds some light on the reasons for the undercount. The table shows

¹⁶ The automatic withdrawal of Medicare Part B premiums from Social Security benefits may affect the reported benefit levels in some years. Prior to the 2007 survey, the Social Security benefit questions did not specify whether the respondent should provide their gross Social Security benefit amount or their benefits net of Medicare premiums. Beginning in 2007, the question explicitly asked for the gross amount. As a sensitivity check, we estimate the change in Social Security benefits in 1982 if all respondents age 65 and older paid Medicare premiums and reported the net amount. The mean Social Security benefit for seniors would rise by \$300 or about a 3 percent increase of reported Social Security benefits.

the proportion of seniors (limited to heads and their spouses) who report receiving Social Security by household type and, among married couple households, for husbands and wives separately. The percentage of single male heads of household who receive Social Security dips slightly in 1989 and 1992, but it then returns to the mid-90s percentage range, so it is not part of the explanation for the mid-to-late 1990s undercount. The corresponding percentage of single senior females shows a somewhat more pronounced drop in 1992, 1995, and 1998 from prior and subsequent years, so it plays only a minor role in the overall undercount.

Table 9. Social Security reciprocity by sex and marital status (65 and older)

	Single		Married	
	Males	Females	Males	Females
1983	95%	91%	91%	53%
1989	91%	95%	94%	92%
1992	90%	91%	94%	93%
1995	95%	92%	78%	78%
1998	95%	92%	79%	82%
2001	96%	96%	83%	83%
2004	96%	98%	95%	90%
2007	94%	95%	94%	91%
2010	95%	97%	92%	95%
2013	95%	94%	92%	88%
2016	93%	95%	92%	93%
2019	95%	96%	91%	91%

Notes: Data are from SCF and only include respondent and, if present, spouse.

The major reason for the undercounts lies in the underreporting of benefit reciprocity among married couples. In 1983, the percent of married males who reported having received Social Security is in line with the share in later years in which there is no apparent undercount. In contrast, the share of females who are reported as recipients is substantially lower than in later years. Thus, the 1983 undercount appears to be the result of too few married females who are reported as receiving Social Security benefits. The survey question for that year asks about Social Security benefits based on “his/her own past contributions.”¹⁷ This wording may have led respondents to omit the receipt of spousal benefits. The cause of the undercount from 1995–2001 appears to be different. During

¹⁷ The 1983 questionnaire is available at https://www.federalreserve.gov/econres/files/1983_quex83.pdf. The head’s Social Security questions begin on page 67; the spouse’s Social Security questions begin on page 85.

these years, the share of both married males and married females who reported receiving Social Security benefits dropped precipitously from their levels in the preceding and subsequent years.

Table 10 provides a more detailed picture of the 1995–2001 underreporting. As the data show, the share of households in which both spouses report receiving Social Security declines sharply from 1995–2001 while the percent of households in which one or the other spouse reports receiving benefits shows no corresponding decline.¹⁸ This suggests that the SCF inadvertently records the receipt of Social Security benefits for only one of the two spouses when it is likely that both are receiving benefits. Moreover, as shown in table 9, it appears that occasionally the husband’s benefit fails to be recorded and at other times, the wife’s benefit is not recorded.

Table 10. Social Security reciprocity among married senior households

	At least one spouse receiving	Both spouses receiving
1983	93%	38%
1989	95%	80%
1992	96%	77%
1995	91%	51%
1998	94%	55%
2001	95%	60%
2004	97%	71%
2007	97%	70%
2010	95%	74%
2013	95%	68%
2016	96%	74%
2019	94%	73%

Notes: Married senior households include all households where the household head is 65 or older.

To explain why the underreporting occurs in the 1995–2001 surveys but not in either the 1989 survey or the post-2001 surveys, we examined the survey questionnaires for these years to see if there are differences in the wording of Social Security related questions. The wording of the Social Security questions changes slightly from year to year, but the differences are minimal.¹⁹ A potential

¹⁸ The sample used for this table differs slightly from the sample used in the previous tables. Table 10 is based on households in which the head is age 65 or older. Thus, it includes some married males who have spouses that are less than age 65 and it excludes households where the spouse is age 65 or older and the head is under age 65.

¹⁹ Married heads were first asked whether either spouse received Social Security. If they answered yes, then they were asked a follow-up question about which spouse receive the benefit. In 1992, married heads were asked “Are you both receiving Social Security payments or is only one of you? (Which one?)” while in 1995, married heads were asked “Are you and your (spouse/partner) currently receiving Social Security payments or is only one of you? (Which one?).”

explanation is that the SCF transitioned from paper questionnaires in 1992 to a “computer-assisted personal interviewing program” in 1995; this change may have affected the survey design in ways not evident in the published codebooks and questionnaires.

Although the underlying reasons for the 1983 and 1995–2001 undercounts appear to be different, in both cases the undercounts are a consequence of failing to capture the receipt of Social Security by one of the spouses in married couple households. This finding has noteworthy implications for our subsequent analysis of the role that Social Security plays in senior household income. Since the underreporting affects only whether both spouses receive Social Security benefits, as opposed to only one, the underreporting does not affect the number of households that receive benefits. The finding has influenced our approach to correcting the underreporting problem. In both periods, we impute a Social Security benefit to a subset of spouses. The imputation is based on Social Security’s spousal benefit formula. Specifically, the imputed benefit is equal to 50 percent of the reporting spouses benefit to all married spouses age 62 and older who do not report receiving Social Security benefits and who are not in the labor force. This method will likely understate a spouse’s benefit level in cases where they qualify based on their own earnings. It could overstate benefit levels for spouses who receive a spousal benefit if the non-reporting spouse retired at a younger age than the reporting spouse.²⁰

To test the appropriateness of this imputation, we compare Social Security income data in the CPS and the SCF among married couples age 62 and older who both reported receiving Social Security. Table 11 compares the medians of the ratio of reported Social Security benefits of the lesser-earning spouse to the higher-earning spouse. In the CPS, the lesser earning spouse reported Social Security income at about 50 percent of their spouses until the early 2000s when the ratios began to rise. The SCF reveals a similar trend, albeit slightly delayed. The one exception is 1983 where the SCF ratio is much higher, at 57 percent. This is consistent with our earlier finding that the 1983 SCF omitted Social Security income from wives earning a spousal benefit, which is typically less than or equal to half of the husband’s benefit. Similarly, the fact that there is no discernable change in the SCF ratio from 1992 to 2004 is consistent with the finding that the 1995–2001 undercount is not due to omitted spousal benefits. In any case, the fact that the median ratio is close

²⁰ Officially, the spousal benefit formula is half of the spouse’s primary insurance amount (PIA) if the spouse begins collecting benefits at the full retirement age. In 2022, a spouse that begins collecting benefits at 62 would only receive 32.5 percent of their spouse’s PIA. Our calculation uses the reporting spouse’s current monthly benefit, which may be higher or lower than their PIA depending on what age the reporting spouse began collecting benefits.

to 50 percent throughout the 1980s and 1990s suggests that our imputation strategy for non-reporting spouses is reasonable.²¹

Table 11. Ratio of SS income of lesser-earning spouse to higher-earning spouse

	CPS	SCF
1983	50.0%	57.2%
1989	50.0%	48.5%
1992	50.0%	48.3%
1995	49.8%	48.2%
1998	50.0%	50.0%
2001	50.8%	50.0%
2004	50.9%	50.0%
2007	51.4%	48.7%
2010	53.0%	50.0%
2013	53.3%	54.0%
2016	55.0%	53.8%
2019	56.1%	55.0%

Notes: Samples are limited to married couple households where both spouses were age 62 or older and reported Social Security income. Data are based on reported levels before our imputations.

Table 12 shows the imputation's effect on the share of seniors who receive Social Security benefits and on average benefits received by recipients. The revised amounts are also compared to administrative data. The imputation produces a substantial increase in the percentage of recipients in all four years. It also brings the percentage in line with administrative data. The imputation lowers the mean Social Security benefit, but the estimates are still within 10 percent of the administrative data. While the 1982 and 1995–2001 surveys appear to inadvertently count only one of the two spouses in married couple households as a survey year Social Security recipient, this error does not affect the household's prior year total income. Total household income is obtained from a set of separate questions that pertain solely to the prior year and exclude any consideration of survey-year Social Security benefits. Thus, our imputations leave total household income unchanged in all years. The imputations only affect how we divide the combined Social Security and defined pension income variable between these two income sources and among each spouse. The allocation procedure is discussed next as part of our description of the method used to obtain household Social Security and pension income.

²¹ It is possible that in the 1995, 1998, and 2001 surveys, we are missing Social Security income of the higher-earning spouse, meaning our imputation strategy will understate Social Security benefits in these households.

Table 12. Social Security reciprocity and benefits after spousal imputation: SCF and SSA

	Participation rate			Mean benefit		
	SCF before imputations	SCF with imputations	SSA	SCF before imputations	SCF with imputations	SSA
1983	81%	91%	90%	\$11,100	\$10,650	\$11,250
1995	84%	90%	91%	\$12,650	\$12,150	\$12,750
1998	85%	92%	91%	\$12,700	\$12,150	\$13,250
2001	87%	94%	92%	\$13,250	\$12,700	\$13,950

Notes: SSA data are for the month of December of the prior year. The administrative participation rates use population estimates from the Social Security Trustees annual report (see table V.A3 in the Trustees supplemental tables.). We use the mean of the survey year and the previous year to estimate the over-65 population as of December of the year prior to the survey. Income data are adjusted for inflation using the PCE price index.

3.2 Estimating Prior-Year Social Security and Private Defined Pension Plan Income

As noted above, all survey years report combined prior-year income from Social Security and private pensions. The 1983 survey also included a separate question for prior-year Social Security income earned by the head and spouse, if present. Thus, pension and Social Security income can be easily decomposed in the 1983 survey. Beginning in 1989, however, separate Social Security and pension data are only available as of the time of the survey. Thus, to analyze the sources of senior household income changes over time using prior-year income data, we must separate prior-year Social Security and pension income into its component parts.

The method we used to separate the components preserves the combined value of the two prior-year income sources. First, after imputing missing spouses' benefits in the 1995–2001 surveys, we reduce the survey year monthly Social Security benefits for household heads and spouses to account for Social Security's annual cost-of-living adjustments.²² Second, we annualize these reduced benefits. Persons who report having received Social Security benefits for a year or more are assumed to have received benefits for the entire prior year. Those who report having received Social Security benefits for less than a year are assumed to have received benefits for six months during the previous year. Spouses for whom Social Security benefits are imputed are assumed to begin collecting benefits at age 62 or at the same time the reporting spouse began collecting benefits, whichever was later. Third, we use the annualized Social Security income as our measure of the

²² We also cap benefits to the maximum one could receive if one always had earnings at or above the taxable maximum. Each respondent is assigned the minimum of their reported benefit amount or the annual scheduled benefit amount for someone with "steady maximum earnings" in the year the person reaches full retirement age (adjusted for early or delayed retirement). We use the estimated annual scheduled benefit amounts from the 2021 Social Security Trustees Report to determine the maximum benefit (see table V.C7 in <https://www.ssa.gov/OACT/TR/2021/lrIndex.html>).

household's prior-year Social Security income if it did not exceed the combined household Social Security and private pension income reported for the prior year. Private pension income, therefore, is treated as a residual. If, on the other hand, our computed household Social Security income measure exceeds the reported prior-year combined value, we set the household's Social Security income equal to the total income from Social Security and private pensions. In this case private pension income is set to zero.²³

Table 13. Social Security participation rates and mean benefits among senior households

	Participation rate			Mean benefit among recipient households		
	SCF	CPS	Difference	SCF	CPS	% Difference
1982	92%	93%	-1%	\$14,600	\$14,050	4%
1988	93%	94%	-1%	\$14,500	\$14,550	0%
1991	89%	93%	-4%	\$14,650	\$15,050	-3%
1994	88%	93%	-5%	\$15,100	\$16,700	-10%
1997	87%	93%	-5%	\$15,900	\$17,650	-10%
2000	95%	93%	3%	\$17,750	\$18,350	-3%
2003	97%	92%	5%	\$18,700	\$18,950	-1%
2006	94%	90%	3%	\$18,850	\$19,450	-3%
2009	95%	89%	6%	\$21,050	\$21,750	-3%
2012	93%	88%	6%	\$21,600	\$21,350	1%
2015	93%	86%	8%	\$23,800	\$22,050	8%
2018	92%	85%	7%	\$24,300	\$22,250	9%

Notes: Income data are adjusted for inflation using the PCE price index.

Our method for allocating the two income sources is based on our assumption that Social Security benefits are more likely to be accurately reported. Social Security benefits are received monthly from a single source while employer provided pensions are often received less regularly and often from more than one source. Table 13 reports the results obtained from this method for the share of senior households that receive Social Security income and the mean amount of Social Security income among recipient households. The table compares these results to those computed from CPS data.

An alternative method of separating the combined Social Security and private defined benefit income into its two components is to allocate each income source according to its respective

²³ Obviously, this procedure can misestimate the allocation of income between Social Security and other pensions for persons who, at the time of the survey, began collecting either Social Security or private pension benefits at some point during the year prior to the survey. In the absence of knowing the number of months such individuals have been recipients, there is no way to determine what portion of prior-year benefits were received.

proportion of the combined total. This alternative assumes that both measures are equally likely to be reported accurately. Tables 14 and 15 show the sensitivity of our results among senior households to the choice of method.²⁴ Table 14 compares the share of senior households that receive Social Security benefits and the average Social Security income between the two methods. Social Security participation rates are largely unaffected by the choice of allocation method while income levels are 4 to 9 percent higher under our chosen method.

Table 14. Participation and benefit levels for Social Security under varying assumptions

	Participation rates			Benefit levels		
	SS-first method	Proportional method	Difference	SS-first method	Proportional method	% Difference
1982	92%	92%	0.0%	\$14,600	\$14,600	0%
1988	93%	93%	-0.1%	\$14,500	\$13,700	6%
1991	89%	89%	-0.1%	\$14,650	\$13,650	7%
1994	88%	88%	-0.2%	\$15,100	\$13,850	9%
1997	87%	88%	-0.2%	\$15,900	\$14,750	8%
2000	95%	95%	0.0%	\$17,750	\$16,650	7%
2003	97%	97%	0.0%	\$18,700	\$17,800	5%
2006	94%	94%	-0.1%	\$18,850	\$17,900	5%
2009	95%	95%	0.0%	\$21,050	\$20,100	5%
2012	93%	94%	-0.2%	\$21,600	\$20,250	6%
2015	93%	94%	-0.2%	\$23,800	\$22,800	4%
2018	92%	92%	-0.1%	\$24,350	\$23,100	5%

Notes: The SS-first method treats defined benefit pension income as a residual after calculating Social Security benefit levels. The proportional method calculates each income source first and then allocates them proportionally such that their sum matches the combined annual income variable. Income data are adjusted for inflation using the PCE price index.

Table 15 compares defined pension participation rates and average defined benefit pension income levels under the two alternative methods. Scaling pension benefits produces higher average benefits, but its effect on the average benefit and the growth in average benefits over time is minimal.

²⁴ Social Security and pension income constitute only a small fraction of non-senior household income. Fewer than 10 percent of these households report receiving Social Security benefits or pension benefits. Hence, differences in Social Security income and pension income that are due to the choice of allocation method are inconsequential.

Table 15. Participation and benefit levels for private pensions under varying assumptions

	Participation rates			Benefit levels		
	SS-first method	Proportional method	Difference	SS-first method	Proportional method	% Difference
1988	57%	54%	3.4%	\$14,300	\$16,050	-11%
1991	52%	49%	2.7%	\$15,100	\$16,750	-10%
1994	48%	46%	1.8%	\$16,750	\$19,300	-13%
1997	50%	48%	1.5%	\$18,000	\$20,200	-11%
2000	52%	50%	2.5%	\$18,650	\$21,450	-13%
2003	61%	59%	2.3%	\$21,900	\$23,650	-7%
2006	52%	50%	1.3%	\$19,900	\$21,400	-7%
2009	57%	55%	2.2%	\$20,950	\$22,800	-8%
2012	56%	54%	2.2%	\$22,350	\$24,800	-10%
2015	56%	54%	2.2%	\$27,850	\$29,850	-7%
2018	52%	51%	1.0%	\$26,550	\$28,500	-7%

Notes: The SS-first method treats defined benefit pension income as a residual after calculating Social Security benefit levels. The proportional method calculates each income source first and then allocates them proportionally such that their sum matches the combined annual income variable. Income data are adjusted for inflation using the PCE price index.

4 Correcting for Defined Contribution Plan Withdrawals

The SCF contains detailed information on participation in defined contribution retirement plans of the respondent and their spouse, if present. This includes whether the individual is enrolled in a plan, the type of plan (i.e., an employer sponsored defined contribution plan such as a 401k, Keogh, 403b, or 457), or an Individual Retirement Account, the amount of any plan contributions, plan assets, and withdrawals from the individual's plan. The type of plan and plan assets are recorded as of the time of the survey, while withdrawals are recorded as part of prior-year income.

Income from defined contribution plans in surveys prior to the 2004 survey appears to be significantly underreported. In the 1989–2001 surveys, withdrawals from IRAs and employer sponsored defined contribution plans were obtained from a catch-all question in which respondents were asked about prior-year income from “other sources.” The list of possible other sources contained over 20 items, including gambling winnings, income tax refunds, and honoraria. It appears that many survey respondents failed to identify defined contribution plan withdrawals. A small number of those who did so appear to have included rollovers of defined contribution plan assets into IRAs in their response. In the 1983 survey no information about withdrawals is reported. In 2004, the SCF significantly changed the questions related to defined contribution plan withdrawals. Starting with the 2004 survey, respondents were asked a separate question about IRA and employer sponsored defined contribution plan withdrawals. Also starting in 2004, respondents were asked to

distinguish between retirement plan withdrawals and rollovers of balances in employer-sponsored plans to IRA or other retirement savings vehicles.

The underreporting can distort comparisons between incomes in the 1990s and years after the survey change. It should not appreciably affect income comparisons between the 1983 survey and surveys after 2001 because IRAs and 401k plans—now the dominant types of defined contribution plans—were in their infancy in 1983. With little time available to build significant balances, IRAs were not yet an important source of retirement income in the early 1980s. Internal Revenue Service regulations for newly authorized 401k plans were not issued until 1982. Although Keogh, 457 plans, and 403b plans had been in operation for a considerable amount of time by 1982, they were only available to certain government employees, employees of non-profit organizations, and small business owners. Thus in 1983, only 11 percent of senior households were enrolled in defined contribution plans and the contribution of these plans to total senior household income was likely to be quite small.

Table 16. Defined-contribution plans statistics among senior households

	Share with plan	Mean assets	Share of participants with withdrawals	Mean withdrawal among participants
1983	11%	\$63,400	n/a	n/a
1989	19%	\$95,200	5%	\$607
1992	22%	\$91,150	1%	\$567
1995	26%	\$120,450	2%	\$130
1998	32%	\$140,800	2%	\$116
2001	36%	\$218,900	10%	\$1,700
2004	36%	\$229,250	60%	\$8,996
2007	40%	\$244,650	60%	\$10,517
2010	41%	\$293,800	46%	\$8,801
2013	39%	\$413,700	55%	\$13,435
2016	45%	\$376,900	66%	\$15,627
2019	43%	\$406,600	62%	\$15,366

Notes: DC plan data include IRAs. Withdrawal data are for the preceding year. Mean withdrawals are among all participants including those who reported making no withdrawal. Data are adjusted for inflation using the PCE price index.

Table 16 provides evidence of the underreporting in the 1983–2001 surveys. The table reports the percent of senior households that were enrolled in defined contribution plans, the mean assets of enrollees, the percent of enrollees who withdrew income from their plan, and the mean

amount withdrawn.²⁵ The first two columns reveal a steady increase in the share of senior households who are enrolled in defined contribution plans and the average asset balance of these plans from the early 1980s to the early 2000s followed by a jump in recent years. Both the enrollment and asset levels are recorded as of the survey date and appear to be accurately reported in the 1983–2001 surveys. The differences in reported defined contribution withdrawals between the pre- and post-2004 surveys, shown in the next two columns, are large. Prior to 2004, both the reported defined contribution plan participation rate and mean withdrawals by plan participants are negligible and markedly lower than the corresponding numbers in 2004 and later years. Although the 10 percent plan participation rate and mean withdrawal amount in 2000 show some improvement over prior years, both statistics are considerably lower than in subsequent years.

Table 17. Effect of reported DC withdrawals on senior household total income

	Median			Mean		
	With DC withdrawals	Without DC withdrawals	% Difference	With DC withdrawals	Without DC withdrawals	% Difference
1982	\$25,450	\$25,450	0.0%	\$43,150	\$43,150	0.0%
1988	\$26,850	\$26,850	0.0%	\$60,500	\$60,350	0.2%
1991	\$24,700	\$24,550	0.6%	\$49,350	\$49,000	0.7%
1994	\$25,350	\$25,350	0.0%	\$52,550	\$52,500	0.1%
1997	\$26,900	\$26,900	0.0%	\$53,500	\$53,450	0.1%
2000	\$32,750	\$32,750	0.0%	\$64,800	\$64,200	1.0%
2003	\$35,100	\$33,750	4.0%	\$62,800	\$59,550	5.4%
2006	\$35,150	\$32,400	8.5%	\$78,550	\$74,500	5.4%
2009	\$39,000	\$37,850	3.1%	\$69,600	\$66,000	5.5%
2012	\$38,950	\$36,700	6.1%	\$80,600	\$75,350	7.0%
2015	\$46,800	\$43,150	8.5%	\$95,600	\$88,500	8.0%
2018	\$47,000	\$44,150	6.4%	\$91,950	\$85,250	7.9%

Notes: DC plan data include IRAs. Data are adjusted for inflation using the PCE price index.

Table 17 provides an indication of the likely magnitude of the underreporting on median and mean senior household income. The left-side columns compare median senior income excluding defined contribution withdrawals and median income including withdrawals, and the percentage by which reported defined contribution income increases household income. The right-side columns show the same data for mean income. Looking first at the years from 2004 to 2018, including

²⁵ Median withdrawals are zero in all years before 2004, and hence, not reported.

defined contribution income raises total household income by only a small amount. The median increase ranges from a low of 3 percent to a high of 9 percent. Similarly, the mean increase ranges from 5 percent to 8 percent. Including defined contribution income raises the average growth in median and mean incomes by 3 percentage points from 2003 to 2018. While the impact of defined contributions on mean income rises over time, its impact on median income exhibits no pronounced trend. The percent increase in income from 2003 to 2018 likely provides an upper-bound estimate of the true impact of defined contribution plan income on total household income prior to 2004. As shown in table 16, the percent of households that were enrolled in defined contribution plans in the 1980s was only one-quarter to one-half compared to the percent enrolled from 2004–2019. Plan asset levels relative to income were only one-fifth to one-third as large. In the 1990s, the defined contribution plan participation rate was only two-thirds the average rate from 2004–2019 surveys and plan asset levels were only about half as large.

To correct for the underreporting in survey years prior to 2004, we impute defined contribution plan withdrawals for individual senior households who are enrolled in such plans from regressions applied to SCF data for the survey years from 2004 to 2019. Specifically, two separate regression equations estimated using pooled data are used to predict withdrawals among plan participants in 2003–2018.²⁶ The sample of senior households used for these regressions is extended to include households headed by persons age 60–64. A non-trivial percentage of these younger households have income from defined contribution plans, which is likely to also be underreported in the 1983–2000 surveys.²⁷ The first equation uses the sample of defined contribution plan participants to estimate the probability that households made withdrawals from their defined contribution plans. It is estimated with probit. The second equation uses the sample of plan participants who withdrew income from their defined contribution plans to estimate expected withdrawals conditional upon withdrawing any income. The dependent variable for this equation is the natural log of withdrawals. This equation is estimated with ordinary least squares. The product of the two equations produces an estimate of the expected withdrawal amount conditional on being

²⁶ This model is broadly similar to the multi-part model described in Duan, Manning, Morris, and Newhouse (1984).

²⁷ From 2003 to 2018, 17 percent of participating households with heads age 60 to 64 withdrew funds from their defined contribution plan. In comparison, withdrawal rates were 4 percent for participating households with heads age 50 to 59 and under 3 percent for households with heads under age 50. The large increase at age 60 is likely due to differential tax treatments; withdrawals taken before a plan participant reaches age 59 ½ are generally subject to a 10 percent tax penalty.

enrolled in a defined contribution plan. The parameters of both equations are applied to plan participants in years 1983 to 2001 to estimate their withdrawal amounts in the preceding year.²⁸

Separate regression equations are estimated for married couple and single headed households and, within each of these household types, those in which the head is age 70 or older and those in which the head is age 60–69. The distinction between households at age 70 is made because during the years covered by our analysis, except 2009, federal law required individuals to begin making required minimum withdrawals (RMDs) from defined contribution plans when they reach age 70½. Previous research has shown that many individuals do not start making withdrawals until they reach this age (Brown, Poterba, and Richardson 2017).²⁹ Thus, we have four sets of estimates: married and single senior households headed by persons age 70 and older and married and single senior households headed by persons age 60 to 69.

The same basic regression specification was used in all equations, except spousal information is included in the married couple equations. The demographic variables in all equations include dummy variables for age and education. In the age 60–69 unmarried household regressions, the regression includes a dummy variable that represents households whose heads are age 65–69. The omitted group, therefore, is households with heads age 60–64. In the older household regressions, we include a dummy variable for households headed by individuals age 70. This is to account for the fact that some age 70 individuals may not have reached the age required to make an RMD (70 ½) in the prior year.³⁰ The omitted group is households with heads age 71 or older. The married couple equations contain dummy variables that allow full interaction between head of household and spouse ages using the same age groupings: 60–64, 65–69, 70, and 71+. The omitted group in the younger household regressions is households in which the head is age 65–69 and the spouse is age 71 or older. The omitted group in the older household regressions is households in which the head and spouse are both age 71 or older.

²⁸ To minimize the effect of outliers, the regressions exclude observations with defined contribution asset balances greater than \$4 million.

²⁹ In summarizing the research on the effect of RMDs, Brown, Poterba, and Richardson (2017) find “substantial numbers of qualified plan participants do not take any distributions prior to [the RMD] age.” A joint test of equality between parameter estimates of the younger and older senior household regressions rejected the hypothesis that the regressions were the same.

³⁰ Our age variable is defined as age at the end of the year prior to the survey. Because the SCF does not provide the month survey respondents were born, we are unable to determine whether an individual reached 70 ½ at the end of the prior year.

The education variables are dummy variables for whether the head attended but did not complete college and whether the head attained a college or higher-level degree. The omitted group is households in which the household head completed a high school education or less. The regressions on single household heads include a dummy variable that represents male household heads. Married couple households in the SCF are nearly always headed by a male, so no dummy variable for sex of household head is included.

Both sets of equations also include economic variables. Many of these variables are jointly determined with both the probability of withdrawing any income from a defined contribution plan and the amount of income withdrawn. While the parameter estimates of the impact of these variables aid in obtaining predictions used for our imputations, the parameter estimates should not be considered as structural estimates from an underlying behavioral model.

The economic variables include two variables that measure the household's financial assets: one for the defined contribution asset balance and the other for the household's financial assets that are held outside retirement accounts.³¹ To capture the impact of continued employment, the regressions include a dummy variable for the head of household's (and spouse in married couple households) employment status and a continuous variable measuring the household's labor earnings.³² Separate dummy variables are included for whether either the head or spouse was enrolled in a defined benefit plan and whether either person was receiving Social Security benefits. The amount of household income received from each of these sources are also included. All continuous variables are in natural log form. Finally, dummy variables for each year are included.

Since 1989, the SCF has used a multiple imputation method to impute missing data. This method yields five "implicates" for each respondent that allow researchers to account for imputation error when estimating standard errors. Similarly, since 1989 the SCF has included replicate weights to estimate bootstrapped standard errors that reflect the survey's complex survey design. The Federal Reserve provides a Stata package, *SCFCOMBO*, that accounts for the multiple imputations and produces bootstrapped standard errors.³³ Accordingly, we estimate two versions of each set of

³¹ The timing mismatch between the value of assets, both those in defined contribution plans and those held outside of retirement plans, and retirement plan withdrawals introduces errors in our imputation process. Asset values are measured months after the end of the year in which withdrawals occur and, hence, measure the value of assets after the withdrawals were made. Also, some households that were not, in fact, plan participants during the prior year are mistakenly treated as plan participants who have not made withdrawals.

³² Employment status is defined according to whether the household head or spouse worked at all during the year.

³³ For details, see https://www.federalreserve.gov/econres/files/standard_error_documentation.pdf.

regression equations. The first set is estimated using the SCF-provided sampling weight and a single implicate.³⁴ The second set uses *SCFCOMBO* to correctly estimate standard errors using SCF data. The latter set are shown here; the appendix provides the estimates from the single implicate. The predicted values, the estimated coefficients, and standard errors are not materially different between the two estimation techniques.

Tables 18 and 19 provide the summary statistics and the individual parameter estimates for the probit and OLS regressions. The primary objective of the regressions is to obtain predicted values so the regression summary statistics are particularly important in providing an indication of how well the regressions fit the data.³⁵ Both the equations estimating the probability of making a defined contribution withdrawal and the amount of the withdrawal conditional upon making one are statistically significant for all four demographic groups. The percent of the observed variation for whether withdrawals are made that is explained by the probit model is on the low side of tolerability; the percent of explained variation from the amount withdrawn is higher and is reasonable for micro data.³⁶

Table 18 shows the probit parameter estimates of the probability of making a withdrawal among plan participants. In three of the four groups, employment of the head is statistically associated with a lower probability of the household withdrawing funds from its defined contribution plan. Larger defined plan asset balances are also statistically related to higher withdrawal probabilities among all households. The coefficients on the age dummy variables among older households capture to some degree the effects of the required minimum distribution. Recall in our data, we cannot determine whether individuals who are age 70 at the end of the previous year were at least 70 ½ at the end of the year. The inclusion of an age 70 dummy variable distinguishes

³⁴ Specifically, the observations for these regressions were selected using the first of the SCF's five implicates.

³⁵ These summary statistics are taken from the regressions estimated using a single implicate. The Federal Reserve's statistical package used to calculate standard errors, *SCFCOMBO*, does not include these summary statistics.

³⁶ The number of observations on both married couple and single older households which make withdrawals from defined contribution plans is about 25 percent less than the corresponding number of households which are enrolled in such plans. It appears that some SCF households whom we would expect to be making mandatory defined contribution withdrawals are not doing so. One reason is that some individuals who are required to take the minimum withdrawal may simply not do so. IRS Statistics of Income data indicate that this isn't likely the case. Fewer than 7,000 persons paid a penalty to IRS for not taking their RMD in 2019. It is possible that filers are simply ignoring this tax provision, but there are several other possible data related reasons. First, some enrollees who are age 70 during the survey year haven't yet reached the age at which withdrawals are required. Second, some enrollees are in Roth IRAs. Third, some enrollees are in 401k plans and are still working for the sponsoring employer. Fourth, in married couple households the spouse may be the enrollee and is less than the age at which distributions are required. None of these reasons, however, can fully account for the phenomenon.

such individuals from those who were at least 71 during the prior year. Older single households headed by persons who are age 70 are less likely to make withdrawals than those headed by persons who are age 71 or older.

Beyond these results, there is little that can be concluded about relationships between withdrawal probabilities and the regressors. For example, younger single households where the head has at least some college education are statistically less likely to make withdrawals than those with a high school or less, but this relationship is not observed among any of the other household types.

Table 19 shows OLS parameter estimates of the amount withdrawn among households that make withdrawals. There are only a few estimated relationships that are statistically significant across the four groups. Among all household groups, the larger the defined contribution asset balance, the larger the withdrawal. Among older households, a 10 percent increase in asset balances is associated with a 6.4 to 6.9 percent increase in the annual amount withdrawn. Among younger households, a 10 percent increase in asset balances is associated with a 2.1 to 4.8 percent increase. There is some evidence that higher education levels are statistically related to higher levels of withdrawals. Among older senior households, those headed by college graduates have higher withdrawals than those headed by persons with a high school or less education. Those who attended college but did not receive a degree make statistically higher withdrawals than heads of similar households with a high school education or less.

Tables 20 and 21 compare the predicted and actual values for each year and group obtained from the bootstrap technique. Table 20 shows the fitted and predicted values for the probability of withdrawal. The percent differences in the probabilities of making withdrawals for each group across all years are small. They never exceed 2 percent in any year. The percent differences between the predicted and actual amounts withdrawn by defined contribution plan enrollees (shown in table 21) show larger year-to-year fluctuations.

Table 18. Probit parameter estimates of probability of withdrawal

	Married/Under70	Married/70 and over	Single/Under70	Single/70 and over
Head has some college (no degree)	-0.00433	-0.01489	-0.39956***	0.06528
Head has college degree	0.01938	-0.10598	-0.34683***	-0.0366
DC plan assets	0.18953***	0.18305***	0.08163***	0.15173***
All other assets	-0.00273	-0.00189	-0.01406	-0.015182
HH has DB plan income	0.45135	0.58735	0.85562	0.542
HH receives SS	-0.40494	-0.06163	-0.25735	0.75517
Head works	-0.2527***	-0.27669**	-0.38191*	-0.27553
Earnings	-0.07133***	-0.02582**	-0.04483***	-0.01114
Social Security income	0.07449	-0.00485	0.05126	-0.07316
DB plan income	-0.04304	-0.06828	-0.06647	-0.0665
Non-retirement plan investment income	-0.0269***	0.03545***	-0.03362***	0.02555*
2006 dummy	-0.01617	-0.19702	-0.18406	0.4894**
2009 dummy	0.03073	-0.86205***	-0.1979	-0.33009
2012 dummy	-0.09313	-0.08593	-0.23495	0.30629
2015 dummy	0.08245	0.11034	0.1245	0.53963***
2018 dummy	0.24922**	0.01622	-0.30123	0.36673*
Spouse works	0.05907	-0.4009***		
Age: Head: 60 to 64 Spouse: Under 60	-0.51711***			
Age: Head: 60 to 64 Spouse: 60 to 64	-0.4061**			
Age: Head: 60 to 64 Spouse: 65 to 69	-0.4655**			
Age: Head: 60 to 64 Spouse: 70	-1.51644***			
Age: Head: 60 to 64 Spouse: Over 70	-0.12544			
Age: Head: 65 to 69 Spouse: Under 60	-0.32122*			
Age: Head: 65 to 69 Spouse: 60 to 64	-0.53385***			
Age: Head: 65 to 69 Spouse: 65 to 69	-0.53485***			
Age: Head: 65 to 69 Spouse: 70	0.36226			
Age: Head: 70 Spouse: Under 60		-0.24841		
Age: Head: 70 Spouse: 60 to 64		-0.72713		
Age: Head: 70 Spouse: 65 to 69		-0.67247***		
Age: Head: 70 Spouse: 70		0.48413		
Age: Head: 70 Spouse: over 70		-1.00142***		
Age: Head: Over 70 Spouse: Under 60		-1.11345***		
Age: Head: Over 70 Spouse: 60 to 64		-0.62029***		
Age: Head: Over 70 Spouse: 65 to 69		-0.66411***		
Age: Head: Over 70 Spouse: 70		-0.31277		
Age: Head: 65 to 69 No spouse			0.10873	
Age: Head: 70 No spouse				-0.68537***
Single male head			0.13878	0.4795***
Constant	-2.024***	-0.47789	-0.87749**	-1.12853
Observations	2692	1397	745	627
Likelihood ratio	237.742	187.367	90.882	52.266
Pseudo R-squared	0.150	0.238	0.150	0.131

Notes: ***=p<.01, **=p<.05, *=p<.1. Income and asset variables are adjusted for inflation using the PCE price index. We use the natural log of the income and asset variables (plus one). Sample is limited to households that reported having a DC or IRA plan and had head age 60 or older. Standard errors were calculated using bootstrapped method. Summary regression statistics are based on single implicate regressions using the survey sample weight.

Table 19. OLS parameter estimates of the natural log of the withdrawal amount

	Married/Under70	Married/70 and over	Single/Under70	Single/70 and over
Head has some college (no degree)	0.32543**	0.20437**	0.42246**	0.25505**
Head has college degree	0.1614	0.36027***	0.30847*	0.09187
DC plan assets	0.47692***	0.69125***	0.21237***	0.63944***
All other assets	-0.01278	-0.00381	-0.00782	0.02094*
HH has DB plan income	0.27918	0.16584	-1.34219	0.5081
HH receives SS	-2.99641**	0.13734	2.9056	-1.55678
Head works	0.20075	-0.03932	0.34268**	-0.3463**
Earnings	-0.03761***	-0.0025	-0.00362	0.00606
Social Security income	0.30704**	0.01982	-0.26177	0.23892*
DB plan income	-0.04526	-0.01845	0.12891	-0.04313
Non-retirement plan investment income	-0.01503	0.00348	0.0274*	-0.0061
2006 dummy	-0.55016*	-0.06052	-0.23033	0.04326
2009 dummy	-0.21649	0.1336	0.05033	-0.02515
2012 dummy	-0.37375	-0.18667*	0.12156	-0.09621
2015 dummy	-0.19043	-0.02098	-0.0818	0.01312
2018 dummy	-0.437*	0.05359	-0.55782**	0.12302
Spouse works	-0.22563	-0.17936*		
Age: Head: 60 to 64 Spouse: Under 60	0.72847***			
Age: Head: 60 to 64 Spouse: 60 to 64	0.70809***			
Age: Head: 60 to 64 Spouse: 65 to 69	0.24979			
Age: Head: 60 to 64 Spouse: 70	-3.5342**			
Age: Head: 60 to 64 Spouse: Over 70	0.51611			
Age: Head: 65 to 69 Spouse: Under 60	0.63867**			
Age: Head: 65 to 69 Spouse: 60 to 64	0.83868***			
Age: Head: 65 to 69 Spouse: 65 to 69	0.1055			
Age: Head: 65 to 69 Spouse: 70	-1.01758***			
Age: Head: 70 Spouse: Under 60		0.46589		
Age: Head: 70 Spouse: 60 to 64		0.10534		
Age: Head: 70 Spouse: 65 to 69		-0.18291		
Age: Head: 70 Spouse: 70		-0.73521**		
Age: Head: 70 Spouse: over 70		-0.19276		
Age: Head: Over 70 Spouse: Under 60		-0.15808		
Age: Head: Over 70 Spouse: 60 to 64		-0.37768**		
Age: Head: Over 70 Spouse: 65 to 69		-0.28481***		
Age: Head: Over 70 Spouse: 70		-0.04077		
Age: Head: 65 to 69 No spouse			-0.12249	
Age: Head: 70 No spouse				0.14242
Single male head			-0.31094**	-0.22657**
Constant	3.68333***	0.52762	6.40024***	0.52531
Observations	440	1194	168	489
F-Stat	10.810	28.245	2.886	37.591
Adjusted R-squared	0.368	0.635	0.202	0.576

Notes: ***=p<.01, **=p<.05, *=p<.1. Income and asset variables are adjusted for inflation using the PCE price index. We use the natural log of the income and asset variables (plus one). Sample is limited to households that reported a positive withdrawal from a DC or IRA plan and had a head age 60 or older. Standard errors were calculated using bootstrapped method. Summary regression statistics are based on single implicate regressions using the survey sample weight.

Table 20. Probability of withdrawal: Actual vs predicted

	Married/Under 70			Married/70 and older		
	Actual	Predicted	% Difference	Actual	Predicted	% Difference
2003	18%	18%	1.8%	77%	77%	-0.6%
2006	20%	20%	-1.3%	72%	72%	0.6%
2009	21%	21%	0.0%	54%	54%	-0.4%
2012	18%	19%	0.7%	79%	80%	0.6%
2015	22%	22%	0.4%	83%	83%	-0.2%
2018	27%	27%	0.1%	80%	81%	0.4%
	Single/Under 70			Single/70 and older		
	Actual	Predicted	% Difference	Actual	Predicted	% Difference
2003	30%	30%	-1.5%	67%	67%	0.4%
2006	25%	25%	1.0%	85%	85%	0.4%
2009	20%	20%	-0.5%	60%	60%	0.3%
2012	23%	23%	-0.5%	79%	79%	0.2%
2015	32%	31%	-1.3%	86%	85%	-0.8%
2018	18%	19%	1.7%	79%	80%	0.8%

Notes: Sample is limited to households that reported having a DC or IRA plan and had head age 60 or older.

Table 21. Amount of withdrawal: Actual vs predicted

	Married/Under 70			Married/70 and older		
	Actual	Predicted	% Difference	Actual	Predicted	% Difference
2003	\$27,100	\$32,500	20.0%	\$16,750	\$18,150	8.3%
2006	\$23,150	\$25,700	10.9%	\$15,100	\$16,100	6.7%
2009	\$32,200	\$32,150	-0.2%	\$16,900	\$18,800	11.5%
2012	\$32,750	\$28,200	-13.9%	\$22,950	\$24,600	7.2%
2015	\$27,200	\$33,250	22.3%	\$26,350	\$29,400	11.6%
2018	\$25,950	\$30,550	17.7%	\$25,600	\$29,000	13.2%
	Single/Under 70			Single/70 and older		
	Actual	Predicted	% Difference	Actual	Predicted	% Difference
2003	\$17,100	\$19,100	11.6%	\$5,750	\$8,050	39.9%
2006	\$12,500	\$11,400	-8.7%	\$15,800	\$11,400	-27.9%
2009	\$17,450	\$17,200	-1.2%	\$12,850	\$12,650	-1.5%
2012	\$19,500	\$20,750	6.3%	\$13,550	\$13,300	-1.8%
2015	\$13,450	\$17,250	28.0%	\$15,700	\$15,700	-0.3%
2018	\$7,950	\$10,000	25.6%	\$19,500	\$20,650	5.7%

Notes: Sample is limited to households that reported a positive withdrawal from a DC or IRA plan and had a head age 60 or older. Income data are adjusted for inflation using the PCE price index.

Table 22 combines estimates from the equations that separately estimate the probability of making withdrawals and the amount withdrawn to show predictions of the average amount of income withdrawn by households which are enrolled in defined contribution plans. The table is divided between senior households and households where the head is age 60 to 64. The predicted average withdrawal among senior households is within 13 percent of their actual average withdrawal in each year.

Table 22. Actual and Predicted Withdrawals among DC plan enrollees

	Senior Households			Households headed by 60- to 64-year-olds		
	Actual	Predicted	% Difference	Actual	Predicted	% Difference
2003	\$8,500	\$9,600	12.8%	\$5,100	\$4,950	-3.7%
2006	\$9,250	\$8,150	-12.1%	\$3,450	\$3,250	-6.3%
2009	\$8,500	\$9,150	7.7%	\$4,250	\$4,700	11.5%
2012	\$11,750	\$10,750	-8.7%	\$3,950	\$3,350	-14.9%
2015	\$14,550	\$15,550	6.7%	\$4,500	\$5,400	19.7%
2018	\$14,050	\$14,800	5.3%	\$4,450	\$5,100	14.6%

Notes: Sample is limited to households that reported having a DC or IRA plan. Income data are adjusted for inflation using the PCE price index.

Table 23. Effect of fitted DC withdrawal rates among all senior households

	Among plan participants			Share with imputed withdrawals	Mean change in income
	Mean reported withdrawal	Mean imputed withdrawal	Difference		
1982	n/a	\$2,200	n/a	11%	\$250
1988	\$600	\$4,300	\$3,700	19%	\$700
1991	\$550	\$5,050	\$4,450	22%	\$800
1994	\$150	\$6,000	\$5,850	26%	\$1,550
1997	\$100	\$6,350	\$6,250	32%	\$2,000
2000	\$1,700	\$9,250	\$7,550	36%	\$2,750

Notes: Income data are adjusted for inflation using the PCE price index. The 1983 survey did not report withdrawals.

Table 23 shows the regression predictions for the years in which withdrawals are underreported. The first column shows the average reported withdrawal among senior enrollees. Column 2 shows the average imputed withdrawal amount by the same enrolled households. Column 3 shows the net impact of correcting for the underreporting on average defined contribution income among senior household enrollees. Column 4 shows the percentage of senior households which are enrolled in a defined contribution plan. Column 5 reports the net impact of our imputation on

senior household defined contribution income. As expected, the amount of imputed income from DC plans rises over time, but the change is small relative to other income sources.

5 Household Income Estimates

This section assesses the effect of our imputations on household income and the impact of the SCF household definition on household income.

5.1 Effect of income imputations on income estimates

In the previous sections, we showed the impact of adjustments for missing Social Security benefits and underreported defined contribution income on each income source. As we noted above, the Social Security imputations have no effect on total incomes; they only affect the relative share attributed to Social Security versus private pensions. The defined contribution imputations, however, affect estimated income from 1982 to 2000. Table 24 shows the effect the imputation has on median and mean senior household income and their income at the 25th and 75th percentiles of the senior income distribution.

Table 24. Senior incomes before and after imputations

	Median			Mean		
	Before	After	% Change	Before	After	% Change
1982	\$25,450	\$25,450	0.0%	\$43,150	\$43,400	0.5%
1988	\$26,700	\$27,200	2.0%	\$60,500	\$61,200	1.2%
1991	\$24,850	\$25,450	2.4%	\$49,350	\$50,100	1.6%
1994	\$25,200	\$25,900	2.7%	\$52,550	\$54,100	2.9%
1997	\$27,350	\$28,750	5.3%	\$53,500	\$55,450	3.7%
2000	\$32,950	\$34,050	3.4%	\$64,800	\$67,550	4.3%
Growth rate (1982 to 2018)	85%	85%		113%	112%	
	25th Percentile			75th Percentile		
	Before	After	% Change	Before	After	% Change
1982	\$13,300	\$13,300	0.0%	\$46,100	\$46,200	0.3%
1988	\$15,000	\$15,250	1.7%	\$45,850	\$46,050	0.4%
1991	\$14,050	\$14,300	2.1%	\$47,550	\$49,600	4.3%
1994	\$12,950	\$13,700	5.5%	\$47,700	\$49,900	4.6%
1997	\$14,950	\$14,950	0.0%	\$51,750	\$54,600	5.6%
2000	\$17,300	\$18,500	6.9%	\$65,200	\$69,350	6.3%
Growth rate (1982 to 2018)	93%	93%		94%	93%	

Notes: Income data are inflation adjusted using the PCE price index.

The impact on both the median and mean incomes reported in the table is small. The imputation raises the overall medians and means by less than 6 percent in each year. Data for the

25th and 75th percentiles show a similarly small effect of the imputation. Since our imputations apply only to the early survey years, they slightly reduce the estimated growth in mean senior incomes from 1982 to 2019. In 1982, the imputations do not affect the median or the 25th percentile, but slightly change the 75th percentile. These changes do not materially affect the total growth rates.

5.2 Effect of household definition on income estimates

As discussed in section 1, the SCF unit of analysis changed between the 1983 and later surveys. The 1983 survey's unit of analysis included all members of the primary family, and the survey did not capture any income or asset information for household members that were not related to the primary family. Beginning with the 1989 survey, the SCF's unit of analysis, the primary economic unit (PEU), included non-family members that were financially dependent on the head or the spouse of the head, if present. Conversely, the PEU excluded family members who were financially independent of the PEU.

Table 25. Change in incomes from including financially independent family members

	Senior PEUs		Non-Senior PEUs	
	Median	Mean	Median	Mean
1988	5.0%	3.5%	1.1%	2.8%
1991	5.1%	2.8%	4.9%	2.2%
1994	7.2%	3.0%	3.7%	2.9%
1997	5.3%	3.1%	0.8%	2.0%
2000	1.6%	1.0%	1.3%	1.5%
2003	7.7%	2.7%	2.1%	2.3%
2006	8.3%	3.6%	4.6%	2.0%
2009	8.2%	3.7%	3.0%	2.0%
2012	2.9%	2.6%	4.0%	2.1%
2015	4.9%	2.5%	4.3%	2.3%
2018	5.0%	3.4%	6.5%	2.4%

The change from primary family to PEU may affect our observed income trends. The reported income amounts from of primary families in the 1983 survey may include income from financially independent family members that would have been excluded from the later PEU definition. If this income represents a significant share of total income, failure to account for the subsequent change in household definition would understate the growth in incomes since 1983. To test this possibility, we examine the incomes of non-PEU family members who are related to the household head or spouse in the 1989–2019 surveys. Table 25 shows how median and mean income

of senior and non-senior households would change if the income of financially independent family members were included. As we noted in section 1, the effect on income is relatively small. The effect is slightly larger for senior households relative to non-seniors, particularly for the median values.³⁷

Table 26. Effect on median income from financially independent household members

	Senior households			Non-senior households		
	PEU	+NPEU	% Change	PEU	+NPEU	% Change
1988	\$27,200	\$28,600	5.0%	\$57,350	\$59,600	3.9%
1991	\$25,450	\$26,650	4.7%	\$53,100	\$56,450	6.3%
1994	\$25,900	\$27,850	7.5%	\$55,200	\$57,750	4.6%
1997	\$28,750	\$30,750	7.0%	\$59,300	\$61,900	4.4%
2000	\$34,050	\$34,700	1.9%	\$63,250	\$65,650	3.8%
2003	\$35,200	\$37,600	6.9%	\$63,000	\$67,300	6.8%
2006	\$35,150	\$37,850	7.7%	\$63,100	\$67,250	6.6%
2009	\$39,300	\$42,150	7.3%	\$59,050	\$61,800	4.7%
2012	\$38,950	\$40,050	2.9%	\$55,600	\$58,950	6.0%
2015	\$46,650	\$49,450	6.0%	\$59,800	\$64,700	8.2%
2018	\$47,000	\$49,700	5.7%	\$63,650	\$69,800	9.7%
Growth rate (1988 to 2018)	73%	74%		11%	17%	

Notes: Data are inflation adjusted using the PCE price index. PEU median income includes DC imputations.

The 1989–2019 surveys collected information on income received by all household members, even those who financially independent and thus are not included in the PEU. When these non-PEU household members are counted, the household includes all persons living in the same dwelling unit and resembles the more conventional household as defined in the CPS and other surveys. If there are pronounced trends in incomes among these financially independent household members, it could affect our conclusions about the relative changes in household incomes between seniors and non-seniors. Table 26 compares median income by year using the PEU household definition and the expanded definition for both senior and non-senior households. Among seniors, the yearly income differences between the two household definitions are modest. The differences range from 2 to 8 percent and exhibit no trend. The yearly differences among non-seniors are in a

³⁷ As noted above, there are reasons to be skeptical that the reported income from the PEU includes income from PEU members that are not the head or spouse. While officially the income questions should include all PEU members, it appears respondents often report income data from the primary tax unit of the PEU. Gale, et al. (2022), for example, has noted that the PEU aggregates earnings value is most often equal to the reported earnings of the head and the spouse only.

similarly modest range. These differences are somewhat larger from 2003 onward than in earlier years and are especially larger during the last two survey years. Most important, the substantial growth in senior incomes in both absolute terms and relative to non-seniors is not materially affected by the household definition.³⁸

6 Summary and Conclusions

This paper serves as a supplement to Cogan and Heil (2022). That paper uses the Survey of Consumer Finances to examine the growth in the incomes of senior citizens from 1982 to 2018. This paper documents the imputations and adjustments that were made to overcome certain data shortcomings when using the SCF for time series income comparisons. Our documentation focuses on three major data issues: adjusting for underreported Social Security reciprocity in the 1983 and 1995–2001 surveys, separating the combined Social Security and pension income amounts into their component parts in the 1989–2019 surveys, and adjusting for underreported withdrawals among defined contribution plan participants in the 1983–2001 surveys. We also check the sensitivity of income levels to changes in the SCF’s household definition between 1983 and the 1989–2019 surveys.

The 1983 survey appears to underreport spousal benefits and the 1995–2001 surveys undercount the share of seniors that are Social Security recipients. Our adjustment for assigning a spousal benefit brings the share of seniors receiving Social Security and the average benefits more in line with Social Security Administration data and the CPS data during these survey years. The Social Security undercounts do not affect prior-year total household income, which is obtained from responses to questions in other parts of the survey. The 1989–2019 surveys do not provide separate prior-year Social Security income and income from defined benefit plans. Our method of separating the two income amounts into their component parts is designed to preserve the combined income amount. The separation method, combined with the aforementioned Social Security benefit imputations, produce Social Security benefits that reasonably match Social Security Administration and CPS data.

³⁸ Beginning in 1989, the SCF surveys used a multiple imputation strategy to impute missing values. This strategy entails imputing five separate responses for each missing response for each respondent. This yields five separate datasets (implicates), which allow researchers to account for imputation error. Our imputations are specific to each implicate, but we do not account for imputation error introduced by our imputations. This omission will not affect our point estimates but will yield lower, less conservative standard errors.

The SCF underreports withdrawals among defined contribution plan participants in the 1983–2001 surveys. Our imputation method uses information on withdrawals among plan participants in the 2004–2019 surveys to predict withdrawals in the earlier surveys. The imputation has a negligible effect on total household income, increasing median income by between 0 and 5 percent during 1982–2000. It does not affect the estimated growth in median incomes across our sample period.

As a separate matter, we also find that the unique household definition used in the SCF does not materially affect conclusions about the absolute growth in senior household income, nor its growth relative to non-senior households.

The results have an important implication for using the SCF to analyze senior household incomes. As noted earlier, the survey is regarded as the “gold standard” for analysis of senior incomes. Yet, the SCF has certain shortcomings when used to examine changes in senior incomes since the 1980s and 1990s. Our results suggest that while these shortcomings should be accounted for in analyses of income dynamics, they do not materially affect conclusions about the long-term growth in senior incomes.

Sources Cited

- Brown, Jeffrey R., James Poterba, and David P. Richardson (2017). “Do required minimum distribution rules matter? The effect of the 2009 holiday on retirement plan distributions.” *Journal of Public Economics*, 151: 96-109.
- Chen, Anqi, Alicia H. Munnell, and Geoffrey T. Sanzenbacher (2018). “How much income do retirees actually have? Evaluating the evidence from five national datasets.” *Center for Retirement Research at Boston College, CRR WP 2018-14*. Available at: https://crr.bc.edu/wp-content/uploads/2018/11/wp_2018-14_.pdf.
- Cogan, John F. and Daniel L. Heil (2022). “Changes in income among senior citizens: 1982 to 2018.” *Hoover Institution Economics Working Papers 22125*. Available at: <https://www.hoover.org/research/changes-income-among-senior-citizens-1982-2018>.
- Copeland, Craig (2016). “Another year after the Current Population Survey redesign and more questions about the survey's retirement plan participation estimates.” *EBRI Notes*, 37(12). Available at: <https://www.ebri.org/content/another-year-after-the-current-population-survey-redesign-and-more-questions-about-the-survey-s-retirement-plan-participation-estimates-3395>
- 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.
- Czajka, John L. and Gabrielle Denmead (2011). “What we’re missing: How to obtain better estimates of retirement income in our major household surveys.” *Paper presented at the Association for Public Policy Analysis and Management Fall Research Conference, November 3-5, 2011, Washington, DC*.
- Czajka John L. and Gabrielle Denmead (2012). “Income measurement for the 21st century: Updating the Current Population Survey.” *Final Report*. Washington, DC: Mathematica Policy Research.
- Dettling, Lisa J., Sebastian J. Devlin-Foltz, Jacob Krimmel, and Sarah J. Pack (2015). “Comparing micro and macro sources for household accounts in the United States: Evidence from the Survey of Consumer Finances.” *Finance and Economic Discussion Series* 86: 1-69.
- Duan, Naihua, Willard G. Manning Jr., Carl N. Morris, and Joseph P. Newhouse (1984). “Choosing between the sample-selection model and the multi-part model.” *Journal of Business & Economic Statistics*, 2(3): 283-289.
- Gale, William G., Swati Joshi, Christopher Pulliam, and John Sabelhaus (2022). *Simulating Income Tax Liabilities in the Survey of Consumer Finances*. Washington, DC: The Brookings Institution.
- Lindamood, Suzanne, Sherman D. Hanna, and Lan Bi (2007). “Using the Survey of Consumer Finances: Some methodological considerations and issues.” *Journal of Consumer Affairs*, 41(2): 195-222
- Poterba, James M. (2014). “Retirement security in an aging population.” *American Economic Review, Papers and Proceedings*: 1-30.

Appendix

Table A1. Probit parameter estimates of probability of withdrawal (using single implicate)

	Married/Under70	Married/70 and over	Single/Under70	Single/70 and over
Head has some college (no degree)	0.03845	0.02682	-0.40421**	0.0721
Head has college degree	0.04012	-0.07057	-0.34619**	-0.04726
DC plan assets	0.19097***	0.17705***	0.07081	0.1633***
All other assets	-0.002246	0.00087	-0.01465	-0.0132553
HH has DB plan income	0.56932	0.85524*	0.88879	0.32479
HH receives SS	-0.3342	0.13452	-0.58325	0.89672
Head works	-0.29102***	-0.27883*	-0.36212*	-0.2693
Earnings	-0.07156***	-0.02377	-0.04352**	-0.01243
Social Security income	0.06228	-0.02517	0.0864	-0.0802
DB plan income	-0.05402	-0.09043*	-0.07471	-0.04391
Non-retirement plan investment income	-0.03009***	0.03196**	-0.03231*	0.02609
2006 dummy	-0.03314	-0.15255	-0.17306	0.49167
2009 dummy	0.02026	-0.84291***	-0.18475	-0.28588
2012 dummy	-0.10386	-0.04807	-0.21962	0.31972
2015 dummy	0.06158	0.13988	0.1361	0.61366**
2018 dummy	0.23393	0.06584	-0.25728	0.37408
Spouse works	0.07274	-0.42239***		
Age: Head: 60 to 64 Spouse: Under 60	-0.52538*			
Age: Head: 60 to 64 Spouse: 60 to 64	-0.37924			
Age: Head: 60 to 64 Spouse: 65 to 69	-0.43957			
Age: Head: 60 to 64 Spouse: 70	-1.37271**			
Age: Head: 60 to 64 Spouse: Over 70	-0.25786			
Age: Head: 65 to 69 Spouse: Under 60	-0.34261			
Age: Head: 65 to 69 Spouse: 60 to 64	-0.51475*			
Age: Head: 65 to 69 Spouse: 65 to 69	-0.48631*			
Age: Head: 65 to 69 Spouse: 70	0.3881			
Age: Head: 70 Spouse: Under 60		-0.27448		
Age: Head: 70 Spouse: 60 to 64		-0.68484		
Age: Head: 70 Spouse: 65 to 69		-0.63777**		
Age: Head: 70 Spouse: 70		0.48997		
Age: Head: 70 Spouse: over 70		-1.00724***		
Age: Head: Over 70 Spouse: Under 60		-1.08852***		
Age: Head: Over 70 Spouse: 60 to 64		-0.60898***		
Age: Head: Over 70 Spouse: 65 to 69		-0.61799***		
Age: Head: Over 70 Spouse: 70		-0.26521		
Age: Head: 65 to 69 No spouse			0.11067	
Age: Head: 70 No spouse				-0.74405***
Single male head			0.15881	0.52008***
Constant	-2.01057***	-0.51712	-0.7921	-1.34845*
Obs	2999	1578	774	661
Likelihood ratio	179.172	192.372	82.117	45.967
Pseudo R-squared	0.119	0.208	0.139	0.102

Notes: ***=p<.01, **=p<.05, *=p<.1. Income and asset variables adjusted for inflation using the PCE price index. We use the natural log of the income and asset variables (plus one). Sample is limited to households that reported having a DC or IRA plan and had head age 60 or older. Regressions use the first implicate only and use the survey weight.

Table A2. OLS parameter estimates of the natural log of withdrawal amount (using single implicate)

	Married/Under70	Married/70 and over	Single/Under70	Single/70 and over
Head has some college (no degree)	0.21362	0.22432*	0.36034	0.34563**
Head has college degree	0.10972	0.42187***	0.2404	0.11036
DC plan assets	0.47862***	0.67721***	0.20567***	0.65605***
All other assets	-0.01289	-0.00469	-0.00511	0.02663**
HH has DB plan income	0.24143	0.11084	-1.21648	0.72819
HH receives SS	-2.93871**	0.12618	3.05008	-1.94613
Head works	0.23213	-0.07462	0.37783*	-0.31015
Earnings	-0.03991***	-0.00277	-0.0052	0.00048
Social Security income	0.3039**	0.02283	-0.27902	0.28679**
DB plan income	-0.04074	-0.01721	0.11447	-0.07167
Non-retirement plan investment income	-0.01018	0.00561	0.02407	-0.00798
2006 dummy	-0.47212	-0.09934	-0.33198	-0.01429
2009 dummy	-0.18126	0.09382	-0.04498	0.04917
2012 dummy	-0.2906	-0.20649	0.09535	-0.10322
2015 dummy	-0.11829	-0.02266	-0.1557	-0.0484
2018 dummy	-0.38179	0.05021	-0.66854**	0.06403
Spouse works	-0.20171	-0.19152		
Age: Head: 60 to 64 Spouse: Under 60	0.70561**			
Age: Head: 60 to 64 Spouse: 60 to 64	0.71445***			
Age: Head: 60 to 64 Spouse: 65 to 69	0.25278			
Age: Head: 60 to 64 Spouse: 70	-3.53992***			
Age: Head: 60 to 64 Spouse: Over 70	0.33074			
Age: Head: 65 to 69 Spouse: Under 60	0.62655*			
Age: Head: 65 to 69 Spouse: 60 to 64	0.87331***			
Age: Head: 65 to 69 Spouse: 65 to 69	0.09406			
Age: Head: 65 to 69 Spouse: 70	-1.11158**			
Age: Head: 70 Spouse: Under 60		0.49182		
Age: Head: 70 Spouse: 60 to 64		0.14405		
Age: Head: 70 Spouse: 65 to 69		-0.22551		
Age: Head: 70 Spouse: 70		-0.71052**		
Age: Head: 70 Spouse: over 70		-0.226		
Age: Head: Over 70 Spouse: Under 60		-0.16467		
Age: Head: Over 70 Spouse: 60 to 64		-0.42519**		
Age: Head: Over 70 Spouse: 65 to 69		-0.24777*		
Age: Head: Over 70 Spouse: 70		-0.0933		
Age: Head: 65 to 69 No spouse			-0.09627	
Age: Head: 70 No spouse				0.13062
Single male head			-0.30094*	-0.29145**
Constant	3.59561***	0.67826	6.59477***	0.30532
Observations	440	1194	168	489
F-Stat	10.810	28.245	2.886	37.591
Adjusted R-squared	0.368	0.635	0.202	0.576

Notes: ***=p<.01, **=p<.05, *=p<.1. Income and asset variables are adjusted for inflation using the PCE price index. We use the natural log of the income and asset variables (plus one). Sample is limited to households that reported a positive withdrawal from a DC or IRA plan and had a head age 60 or older. Regressions use the first implicate only and use the survey weight.