January 2022

Conference Draft

Simulating Income Tax Liabilities in the Survey of Consumer Finances

William G. Gale

William Gale is the Arjay and Frances Fearing Miller Chair in Federal Economic Policy and Senior Fellow at Brookings and Co-Director of the Urban-Brookings Tax Policy Center

Swati Joshi

Swati Joshi is a senior research assistant at Brookings

Christopher Pulliam

Christopher Pulliam is a research analyst at Brookings

John Sabelhaus

John Sabelhaus is Nonresident Senior Fellow at Brookings

This report is available online at: https://www.brookings.edu/research/simulating-income-tax-liabilities-in-the-survey-of-consumer-finances/

The Brookings Economic Studies program analyzes current and emerging economic issues facing the United States and the world, focusing on ideas to achieve broad-based economic growth, a strong labor market, sound fiscal and monetary policy, and economic opportunity and social mobility. The research aims to increase understanding of how the economy works and what can be done to make it work better.

B | Economic Studies at BROOKINGS

Abstract

The Survey of Consumer Finances (SCF) is a triennial household survey with extensive demographic, income, and balance sheet information. The SCF is unique among public-use household surveys because it oversamples wealthy households and is thus suitable for studying wealth and income inequality across the entire population, but the survey does not collect detailed information about tax filing or tax liabilities. We develop a method for separating SCF households into tax units and creating the necessary inputs for estimating taxes using the NBER TAXSIM model. We validate our methods by benchmarking against published tax filings. Our methodology provides measures of after-tax incomes and effective tax rates for SCF households under current law and allows analysis of tax changes in ways that are not possible using administrative tax data alone.

Acknowledgements

We thank the Peter G. Peterson Foundation and Arnold Ventures for generous financial support. The analysis and conclusions set forth are those of the authors and do not necessarily reflect those of any other individual or organization. Email contact: jsabelhaus@gmail.com.

I. Introduction

This paper describes methods for estimating income tax liabilities in public-use Survey of Consumer Finances (SCF) micro data files. Most recently conducted in 2019, the SCF is a triennial household survey with extensive demographic, income, and balance sheet information, for the designated survey respondent, and if present, the respondent's spouse/partner. The survey also collects basic demographic information, financial dependency indicators, and summary income measures for up to ten additional household members. The SCF is unique among public-use household surveys because it oversamples wealthy households and is thus suitable for studying trends in top wealth and income shares (Bhutta, et al., 2020, and Bricker, et al., 2016). Like most household surveys, however, the SCF does not ask detailed questions about household tax filing or tax liabilities.

In this paper, we develop a methodology to divide SCF households into tax units, reconcile survey and taxable incomes measures, and create the other necessary inputs for estimating income tax liabilities. We then estimate income tax liabilities for SCF tax unit micro files in conjunction with the most recent version of NBER's on-line tax calculator TAXSIM. TAXSIM replicates U.S. federal income tax rules over time, including the 1995 to 2019 period (tax years 1994 to 2018) spanned by the SCF micro data files that we use.[1]

We proceed in several steps. Section II describes our methods for creating tax units within SCF households. For most SCF households—such as a single person or married couple living alone or with dependent children—this process is simple. These households also account for the vast majority of income. Some households, however, contain multiple potential filing units – because they consist of either different generations or unrelated individuals. In these cases, we use data on demographic relationships, financial dependence measures, marital histories, and incomes to simulate tax filing units. We also benchmark our simulated outcomes against published tax filings in the Statistics of Income (SOI).

Section III describes our approach to mapping SCF incomes into taxable (SOI) concepts. SCF incomes are largely intended to be consistent with their taxable counterparts, but even after resolving conceptual differences, we show that the survey values are systematically higher than the published tax values. Although we do not exhaustively explore aggregate and distributional differences across income categories here, the key observation that emerges is that the gap in business incomes (mathematically) accounts for most of the overall income gap.

Section IV describes how we model itemized deductions. Taxpayers can choose between itemized deductions and a standard deduction that varies with filing status. The SCF captures about half of itemizable expenses, and we impute the other half using published SOI deductions. Our two key benchmarks are how well we track the number of tax filers who choose to itemize and the total value of itemized deductions.

In section V, we present baseline tax liability estimates, before and after credits, using the NBER TAXSIM model and benchmark those against published SOI values. Because incomes are systematically higher in the SCF relative to SOI, our estimated tax liabilities are also higher. Because the gap between SCF and SOI incomes is concentrated at the top of the income distribution, and the tax system is progressive, the gap in tax liabilities is not surprisingly larger than the income gap.

In section VI, we conclude by noting that the results contained in this methodology paper, especially the differences in business income across data sources, have important implications for recent controversies regarding the distribution of income and wealth. We explore these topics in a companion paper (Gale, et al., 2021) that builds on the methodology developed here.

1 Our program also processes SCF data from 1989 and 1992, but differences in key demographic and economic input variables make those years less comparable to the data from 1995 and later.

II. Identifying Tax Units Within SCF Households

In 2019 there were approximately 120 million households in the U.S., and those households filed roughly 150 million income tax returns (for tax year 2018) with the IRS. The SCF covers most of the tax filing population, missing only those living in institutional settings and dependent filers living in the households of their parents. The first step in our tax unit version of the public-use SCF is allocating SCF household members across tax units. Many of those tax units are straight-forward—a married couple or a single person at the head of a household, with or without dependent children. Although the relatively simple cases are the vast majority, it is important to capture the more complicated situations so we can benchmark the SCF tax unit micro file against SOI tax filings.[2]

Relationships and Financial Independence

The starting point for splitting SCF households into tax units is the difference between the Primary Economic Unit (PEU) and all other household members, collectively referred to as the Non-Primary Economic Unit (NPEU). The PEU subset of the SCF household automatically contains the survey respondent (R) and the respondent's spouse or partner (SP) if that person is present.[3] But there are up to ten other household members on the SCF household roster, and they can be in either the PEU or the NPEU. Those other household members are associated with the PEU if they are financially dependent on the R and SP. Thus, the PEU, which includes (by rule) children under 18 years of age, also includes other family members who are financially dependent on the R and/or SP. Detailed demographic and income information is collected for the R and SP, and a handful of demographic variables and aggregated income information is collected for members of the NPEU.[4]

As in many household surveys, the key variable for understanding household composition—and the starting point for building out tax units—is each household member's relationship to the R. The SCF relationship codes are standard, including options like "child" or "parent" and other codes that indicate family relationships. The SCF relationship codes also include values for non-relatives such as roommates and boarders, and that distinction is important for allocating dependents across tax units. In addition to the relationship to R, the other key demographic variables collected for every household member are age and marital status.

Building on the familial relationships, financial dependency is the second key input for constructing tax units within SCF households. During the survey, the R is asked a relationship-tailored financial dependence question for everyone on the household roster (again, except children under 18, for whom the answer is assumed). For example, the financial dependence question about the sister of a married (different sex) female R is worded,

> *"Does your sister depend on you and your husband for most of her support, or is she financially independent for the most part?"*

If one or more household member is reported to be financially independent, they are assigned to the NPEU, and there is a series of follow-up questions about the NPEU generally. These questions used to identify financially independent household members are closely related to the criteria used for dependent

2 The methods used here build directly on previous work by Kevin B. Moore at the Federal Reserve Board. His SAS program (taxsimpub.sas) can be downloaded at http://users.nber.org/~taxsim/to-taxsim/scf27-32/. That earlier version of the SCF tax unit program was used, for example, to estimate trends in effective tax rates by wealth class in a series of papers. See Moore, Pack, and Sabelhaus, 2016, and Bricker, et. al., 2020. Other projects have similarly adapted tax calculators to survey data sets such as the Survey of Income and Program Participation (SIPP), and Current Population Survey (CPS-ASEC) for various research purposes. See, for example, Jones and O'Hara, 2016, Wheaton and Stevens, 2016, Warren, Fox, and Edwards, 2020, Hertz, et al., 2020, and Meyer, et al., 2021.

3 The SCF always classifies the male in a different sex couple as the respondent, regardless of which member of the couple answered the survey questions. There is additional information on who answered the questions for interested researchers, but the fact that the same person-level information exists both for R and SP means (for our purposes) that the distinction is not important, because we can treat the R and SP symmetrically.

4 Incorporating SCF NPEUs in the analysis here is quantitatively significant. In the 2019 SCF, total NPEU wages were 5.6 percent of PEU total wages, and NPEU total income was 4.0 percent of PEU total income.

claiming on income tax returns. The follow up questions about levels and sources of income within the NPEU are inputs to the tax filing for those NPEU household members deemed to file their own returns.

One type of tax unit slips generally through the cracks of the SCF relationship and dependency framework. The NPEU income questions are only asked for financially independent household members. Thus, a financially *dependent* household member with earnings, such as an 18-year-old child with a part-time job living in their parent's home and still financially dependent on their parents, is not included in the NPEU. The SCF income questions for R and SP are very detailed, and (as described in more detail below) there is information about wages for R and SP in the SCF labor module. There is also information about PEU *total* wages in the SCF income module. In principle, the wording of the PEU total wages question in the family income module could include the part-time earnings of the 18-year-old student, but in practice the family earnings value is generally just the sum of the R and SP earnings reported in the labor modules. In any event, it is somewhere between difficult and impossible to allocate earnings to dependent filers within the PEU because of the reference periods for which incomes are measured.

In addition—and unlike other household data sets such as the Current Population Survey (CPS)—there is no direct information in the SCF about the labor force activities of financially dependent household members such as the 18-year-old child of the R. (At the same time, there are no questions about financial dependence in data sets like the CPS, and thus a tax unit version of the CPS relies on reported earnings to determine whether someone is financially dependent on the respondent.) The bottom line is that financially dependent filers such as the example 18-year-old child living at home will not be captured in the SCF tax unit disaggregation implemented here.[5] We expect (and confirm) a shortfall in single-filer tax units when we compare the SCF tax unit counts to published SOI tallies, and that shortfall is consistent with the number of dependent tax filers reported by the SOI.

Respondent and Spouse/Partner: One Tax Unit or Two?

Every SCF household has a survey respondent (R), and most households have a spouse/partner (SP) as well. In most situations where an SP is present the R and SP are deemed to file taxes jointly, but there are also cases where the R and SP are deemed to file separate returns. Our decision rule for joint versus separate filing makes use of the marital status variables for R and SP, a survey variable that captures how long the couple has been married, and the answers to direct (but sometimes confusing) survey questions about "who" in the couple filed a tax return in the previous year. In general, unmarried partners, newly married couples, and those couples who clearly state they filed separate tax returns are assigned to separate tax units. We do allow demographics to overrule some answers to the "who filed" tax questions. For example, sometimes a long-married couple reports that "only R" or "only SP" filed a tax return, but those answers seem to be conflating who earned the income or did the actual tax preparation and filing with whether the couple filed a joint (specifically, a Married Filing Jointly) return.

NPEU Tax Units

In addition to the respondent (R) and spouse or partner (SP), there are ten other slots on the SCF household person list (or household "roster" in SCF parlance) and thus ten potential tax filing units. The first criteria for a potential NPEU tax filing unit is financial independence, which rules out (again, by SCF rule) children under 18. Every financially independent household member is a potential tax unit, but the rules we use for allocating total NPEU income across the financially independent NPEU members will eliminate some of those potential tax filers. Also, although most NPEU tax units are single filers, there are some cases where a married (both financially independent of the PEU) couple within the NPEU is deemed to file together.

There are two SCF questions about NPEU income amounts, covering (1) wages and (2) all other forms of income combined. Conditional on reporting any non-wage income, the respondent is then also asked to

5 One possibility for future work is imputing the earnings of dependent children living at home, using household and person demographics. Data sets like the Current Population Survey (CPS) have the information about earnings for dependent children and the necessary imputation variables.

report (in a "check all that apply" format) the types of income included in the "other" income measure, with categories including Social Security, business income, interest, dividends, rental income, pension income, transfers, capital gains, and unemployment insurance. In practice, the "other" income measure usually only includes one or perhaps two types of income (these are, after all, individuals living in someone else's household). Indeed, a typical case of positive "other" income is an older family member, such as a parent, reporting only Social Security.

The relative importance of incomes like wages and Social Security among financially independent household members often makes it straight forward to allocate incomes across the NPEU members and thus identify tax units. Wages are allocated equally to NPEU individuals ages 69 and younger (unless there are none), and Social Security is allocated to NPEU individuals 62 and older (again, unless there are none). In all other cases and for all other incomes the NPEU aggregates are allocated equally across all financially independent adults in the household. Again, the typical SCF NPEU case is one financially independent adult (a roommate, working and financially independent older child, or a parent on Social Security) so the allocation of NPEU income is straightforward.

The marital status variables for all SCF household members also makes it possible to identify married couples within the NPEU, and those couples are deemed to file jointly (meaning their assigned incomes are added together, and their filing status is set to Married Filing Jointly). Good examples of such cases are both parents living in a child's home, or a sister and brother-in-law couple living (financially independent from the PEU) in their sibling's home. We also flag all cases where there is any related (family) filing unit in the NPEU and identify which of those family units has the highest income, because that designation becomes relevant when allocating dependents across tax units within the household.

Identifying and Removing Non-Filers

Demographic variables and financial dependence are the starting point for identifying tax units within SCF households, but if a potential tax unit has no income—or has income below a level at which they would be expected to file a tax return—their potential tax return status is irrelevant. That is, they become a "non-filer" in our tax unit version of the SCF. Indeed, if a potential tax filer does not file because they have little or no income, it is important to establish that fact so that some other (filing) tax unit within the SCF household can claim that individual as a dependent (if they are eligible to be claimed). Unfortunately, official SOI filing thresholds are not particularly useful in this regard, because many individuals file returns even though their income is below the official filing thresholds. Those individuals are filing to recover taxes they or their employers paid during the year, or to take advantage of refundable government benefits such as the Earned Income Tax Credit (EITC). Nonetheless, the published SOI filing statistics we are using to benchmark our SCF tax unit file includes those individuals, so we want to include them in our micro file.

Our solution to the non-filer issue is to use a very low dollar threshold ($1,000) for most types of income for which withholding is likely to occur. Other types of income where withholding does not occur—especially Social Security—do not trigger tax filing in the absence of other incomes, because, for example, an elderly couple with only Social Security would have no reason to file a tax return. Social Security is only taxable (and then only partially taxable) if the couple has substantial other forms of income. Ultimately, the choice of low filing thresholds is justified by the results. By using the low filing thresholds, we see overall tax unit counts consistent with published SOI tax return tallies over time.

Dependency Assignments

Having established tax units within the SCF household, the next step is to assign dependents to those tax units. The dependency assignment rules are key to determining filing status because single individuals (every tax unit not deemed to be Married Filing Jointly or Separately as defined above) can be deemed to file as Single, Head of Household, or Qualified Widow(er). The distinction is important because tax parameters vary substantially over the three filing statuses. That is, for a given income, an individual filing Head of Household pays less in tax than an individual filing Single, and a Qualified Widow(er) pays less

than either. In addition, the assignment of dependents in various age categories matters for various tax credits and deductions.

As with identifying tax units within households, most dependency assignments are straightforward. For example, a child of the R or SP will be deemed to be a dependent on the main household return if the R and SP file jointly. A financially independent adult member of the NPEU will not be claimed as a dependent on someone else's return. However, dependency assignments can become complicated quickly, including situations where a member of the NPEU is deemed to file a return and may be the one to claim some other household member, situations where the R and SP file separate returns, and situations where the R or SP reports having children living elsewhere.[6]

The simplest set of rules apply to children of the R and SP. If R and SP file jointly, all children are dependents on their returns. We assume that children between ages 18 and 24 who are reported to be financially dependent on the R and SP are students, or otherwise eligible to be claimed as dependents on the main household tax return.[7] Indeed, if there is no identified NPEU tax filing unit, all financially dependent family members are deemed eligible to be claimed by the R and SP. In that sense, SCF financial dependence is assumed to be equivalent to the income tax criteria for providing support to a dependent.

If there is an identified NPEU tax filing unit in the household, the dependency allocation rules are adjusted accordingly. The NPEU family tax unit (and the unit with the highest income if there are multiple NPEU family tax units) claims other non-filing family members (except, of course, the children of R and SP). We make no attempt beyond that simple rule to allocate dependents in a way that minimizes overall household tax liability.[8] In no circumstances does any tax unit claim a non-family member as a dependent, though non-family NPEU members can of course comprise a single tax unit.

The situations where family dependents are present in the household and R and SP are deemed to file separately are the most uncertain. The primary limitation is that the relationship codes don't have the detail we would need—we don't know whether a "child" is related to the R or SP, for example. Even if we did have that detail, it might not matter—the higher income partner may be the one to claim that child (or any other dependent family member for that matter). Our solution is to allocate the first dependent to the higher income of R or SP based on who has higher income. If there are two or more dependents in the household, we assign one dependent to the lower income of R and SP (one dependent is enough to improve their filing status situation) and all other dependents to the filer with higher income.

There is also another group of dependents to consider—children of the R and SP living elsewhere, usually the result of divorce or separation. The SCF asks about the number of such children, in age ranges less than 18, 18 to 24, and 25 or older. Many of those children are claimed as dependents for tax purposes in the household where they are living, but some can be claimed (often by a legal settlement) by the household who reports them living elsewhere. We distinguish the two cases using reported child support payments. If the SCF household reports children living elsewhere under age 25, and the respondent reports having paid child support, we count those children as dependents on their return.

---

6 Technically, we specify rules to allocate dependents and set filing status across tax units in nine distinct types of households: (1) R and SP file jointly, no NPEU returns, (2) R and SP file jointly, one or more NPEU returns, (3) R and SP file separately, no NPEU returns, (4) R and SP file separately, one or more NPEU returns, (5) R files but SP does not, no NPEU returns, (6) R files but SP does not, one or more NPEU returns, (7) R does not file but SP does, no NPEU returns, (8) R does not file but SP does, one or more NPEU returns, and (9) only NPEU returns. In addition, some households have no deemed filing tax units.

7 The public-use SCF micro files from 2004 forward round the ages of household members except R and SP for non-disclosure purposes, but the effect on our dependency allocation is not first-order. For example, we know with certainty if a household member is less than 18 years old. If a respondent is between 18 and 27, their age is rounded to 25, and we assign an age using actual distributions within the range 18 to 27 for the financially dependent and financially non-dependent before 2004. This approach captures the fact that, conditional on being between 18 and 27, a financially dependent child is much more likely to be in the younger end of the range. We similarly assign ages to household members between 13 – 17, whose ages are rounded to 17 after 2004.

8 There is some evidence that households do allocate dependents in ways that maximize tax benefits. See, for example, Jones and O'Hara, 2016. Much of that seems to involve shifting dependents between R, SP, and NPEU tax units. For example, in a three-generation household, the highest earner (parent or grandparent) claims the grandchild.

Filing Status

All tax units file as one of five types: Married Filing Jointly (MFJ), Married Filing Separately (MFS), Single (S), Qualified Widow(er) (QW), or Head of Household (HH). The federal income tax system is designed such that there is no advantage to a married couple in switching from MFJ to MFS. There are, however, distinct advantages for a single filer in moving from S to QW or HH. The only simple filing status assignment from our perspective is MFJ, which goes hand in hand with tax units where R and SP are deemed to file to file jointly (as described above). All other filing status designations involve reported marital status and dependency assignments, and the order in which the possible filing status is evaluated matters.

The SCF collects marital status for R, SP, and all other adult household members. The marital status options are standard in household surveys, including married, divorced, separated, widowed, and never married. For all tax units other than R and SP filing together as MFJ, the first decision is whether they are a QW. The QW designation is only available for single Rs and Rs in households where they are deemed to file separately (that is, not MFJ). The first criterion is that the R reports marital status as widowed, the second criterion is that the widowing occurred in a recent year (based on tax law), and the third criterion is that they have dependents. The advantage of QW is substantial, because the QW gets the benefits of filing MFJ.

If the tax unit is not a QW, the second option is MFS, and the first criteria is that they report their marital status as separated. However, there are many situations where (for example) a single R living alone reports their marital status as separated, but the questions about when they were separated indicate the separation occurred many years ago, and the individual is very unlikely to still be filing as MFS with that former partner. Thus, we limit the MFS designation to recent separations (within the past five years). The "when separated" question is only asked of the R, but there is a complementary question that applies to the SP, asking how long they have been with their current partner, so (correspondingly) we allow MFS only if the SP has been with the R for a few years. It is not possible to put a time limit on the NPEU tax filers who report being separated (but also not quantitatively significant).

The final option for filing status is HH. A tax unit headed by a single individual can claim HH if they have a dependent family member who meets the criteria for support, which (again) we base on the SCF "financially dependent" question and relationship codes. Note that many single tax units could be either MFS or HH—they report being separated but also have (based on our dependency assignment algorithm) dependents. This is where the order of precedence for assigning filing status becomes important. Allowing MFS status to be assigned before (and thus preclude) HH gives us a distribution by filings tatus that better matches the SOI published distribution of tax units by filing status.

Both dependency and filing status assignments require assumptions that are likely to be wrong in some situations. That is, if we linked SCF household members to their actual tax records, we would certainly find many cases where the data or assumptions gave us the wrong answer. The SCF can never be linked to tax filings, however, so the evidence that we are getting the assignments right on average comes from looking at SCF versus SOI distributions by filing status over time.

Comparisons to Published SOI Data Part 1: Counts of Tax Units by Filing Status

Barring a direct linkage of SCF households with tax filings (prohibited by the IRS guidelines under which the SCF is conducted) there is no direct way to evaluate our approach to splitting SCF households into tax units. The alternative is to compare the number of simulated SCF tax returns by filing status to corresponding IRS published counts over time.[9] In this section we show that our simulated SCF tax filings by filing status track actual filings quite well, after adjusting for the missing dependent filers in the SCF.

We begin with the total number of simulated tax filings by year (Table 1). The number of actual tax returns filed with IRS increased from 115.9 million in 1994 to 153.8 million in 2018. The corresponding

9 The SOI counts are from the Table 1.3 series on the SOI Tax Stats page, https://www.irs.gov/statistics/soi-tax-stats-individual-income-tax-returns-complete-report-publication-1304.

simulated tax filings in the SCF increased from 105.0 million to 147.6 million, and the ratio of SCF to SOI hovers around 90 percent with an upward trend. These observations are all consistent with expectations, because the number of dependent filers we are missing in the SCF is likely in the single digit millions in recent years, and likely declining over time.[10] The actual number of dependent filers in the SOI is higher, but the SCF will capture (for example) dependent filers living outside their parents' home but still financially dependent.

The more important challenge for our methodology is tracking counts of tax returns by filing status (Figures 1A-1D). In general, the fit by filing status is quite good, and the growth of tax filings across the various sub-types is clearly captured in the SCF. In addition, the gap between the total count of simulated and actual tax filings is (as expected) reflected in a persistent gap in the number of single returns, which is where we would expect to find the missing dependent filers. The other noticeable gap is in the Head of Household filing status, especially in the earlier part of our sample period.

When looking at Figure 1 it is important to remember that the SCF is a small survey—between 4,500 and 6,500 observations over this period. In general, the fit gets better as we move from tax year 1994 to tax year 2018, and that is in part because the SCF sample sizes increase over time and have been consistently above 6,000 for the past few survey waves. To put this in perspective, the number of SCF observations underlying the Married Filing Single group is in the low hundreds, even in recent years. Sampling variability is having a substantial impact on simulated filings.

The small gap in Head of Household returns, especially in the early part of our sample period, is more noticeable, systematic, and likely attributable to factors that go beyond simple sampling variability. One possibility is taxpayer non-compliance. Head of Household status requires an eligible dependent. In some situations, it is not obvious (for example) which partner in a divorced or separated couple can (or should) claim a child for tax purposes. It is easy to imagine situations where both parents—living separately and therefore (statistically) in multiple SCF households—claim the same child. Indeed, we could in principle (with some randomization) build in such non-compliance and create more Head of Household filers (they would come out of the Single filer category). Support for the non-compliance argument is evident in the time series of the gap—as dependent tracking through SSN requirements have improved over time, the gap has closed.

III. Taxable Incomes

Given the assignment of SCF household members to tax units, the next step is identifying, classifying, and assigning taxable values of survey income values to those tax units. Our approach to allocating incomes across tax units within the household is focused on creating the necessary inputs for TAXSIM, and we also want to benchmark the distribution of incomes against published SOI values to better understand how the survey incomes are diverging from incomes reported on tax returns.

Published SOI tables with summary information from U.S. income tax returns generally show numbers of returns and tax variable dollar values organized by Adjusted Gross Income (AGI). The concept of AGI is closely related to the sum of (mostly taxable) incomes collected in household surveys, but some incomes (especially Social Security) are not fully taxable, and a handful of statutory adjustments to taxable income (deductions for self-employment taxes, student loan interest, alimony paid, and contributions to IRAs outside of employer-sponsored plans) are important for various subsets of taxpayers. Some of those AGI adjustments correspond directly to SCF questions while others can be inferred from related SCF variables.

SCF Incomes

As with the allocation of household members into tax units, the starting point for allocating incomes is the distinction between the Primary Economic Unit (PEU) and Non-Primary Economic Unit (NPEU). The

10 A deeper dive on this would involve using a data set like the Current Population Survey, computing the fraction of young people living in their parent's home who are working for pay, and applying those fractions to the SCF.

income module for the PEU comes towards the end of the survey, but there is also relevant information collected along the way in the business and labor modules.

The SCF income module collects total PEU incomes in several categories: wages and salaries, proprietor's income reported on Schedule C, non-taxable interest income, taxable interest income, dividend income, capital gains income, various other types of business incomes such as S-Corps, rent, partnerships reported on Schedule E, unemployment insurance income, child support and alimony received, government transfers, retirement incomes, and a catch-all "other category" which is matched to an "other sources" code frame. For example, two tax-relevant components of SCF "other" income are gambling earnings and net operating loss carry forwards.

The SCF income module is the key input into the tax simulation, but other SCF modules are also important. For example, the labor module collects information about labor earnings for the respondent (R) and spouse-partner (SP) separately, and those are used to split the total PEU income in situations where the household is deemed to have separate tax units for the R and SP. The labor module also collects information about Social Security benefits, pension benefits, and other retirement account withdrawals. Those variables are used to both split the total retirement income category into Social Security versus other retirement incomes and to further divide each of those across the R and SP in multiple tax return households.

The SCF business module is also relevant for processing income inputs to the TAXSIM tax calculator. The business module collects information about legal form of organization, active versus passive, and broad industry for multiple businesses owned and/or operated by the PEU. The information complements information about "shares" of profits collected as incomes in the labor module, and the respondent-reported primary industry. The business type and industry variables are most important for the last year covered by our sample period, because they determine whether the income is eligible for special treatment (QBI and SSTB) after 2017.

In general, the SCF incomes correspond to the taxable income categories reported in regular SOI tables. Indeed, SCF interviewers have access to corresponding tax form and line numbers for each income category, but relatively few respondents follow the interviewers' suggestion to retrieve and refer to their previous year tax return when answering the income module questions. As we will see, one of the biggest income module omissions is failure to report negative incomes and/or overstate positive incomes, particularly among business owners and other high wealth individuals.

The public-use SCF also introduces an additional hurdle. For disclosure avoidance purposes, some very high incomes are blurred, and some losses are reported only with indicator variables. For example, the Schedule C, E, and capital gains income fields are filled in with "-9" if the value is an undisclosed loss and overall income is also an undisclosed negative. Although we don't know exactly what the non-disclosed income losses are, we can infer that some income must be sufficiently negative to offset any observed positives and adjust the taxable income accordingly.

NPEU Incomes

Researchers using the SCF often discount the importance of the Non-primary Economic Unit (NPEU), but those household members account for about six percent of total household income. The NPEU income questions come at the very end of the survey, and the details are limited to minimize the incremental respondent burden. In addition, the total incomes are collected for the entire NPEU, rather than for individual NPEU members, so we assign the incomes across household members using the (limited) demographic information.

There are two broad income questions. The first collects total wages earned by the NPEU. The second collects "all other" income using a "check all that apply" approach to distinguishing what types of income were included in that total. The categories of other income include Social Security, pensions, disability, public assistance, dividends, interest, business, real estate, alimony, child support, and capital gains. The "check all" approach means there will be cases where we can't tell exactly how much of a given type of

income was received. For example, a respondent may report the NPEU earned $20,000, and indicate that included both Social Security and dividends—we assume an equal division across the two types of income in that situation.

The two limitations of the NPEU income module—aggregating across NPEU members and aggregating "other" incomes—are not a serious problem in most cases, because most NPEU situations are relatively simple and well-captured. For example, one typical situation is a single roommate who tends to be younger and have only wage earnings, and the respondent usually has a pretty good sense of what the roommate earns. Another typical situation is multi-generational families. If the parents are the PEU, the children or other relatives living with them tend to have only wages—remember, they are living in their relative's home. If the younger generation is the PEU, the parents or grandparents might have Social Security income, but usually little else.

Solving for Adjusted Gross Income (AGI)

The front page of Form 1040 is well known to most U.S. taxpayers. The top third of the front page collects basic taxpayer and dependent information, and assigns filing status—basically the tax filing steps described in the previous section. The middle third of the front page of Form 1040 collects gross and taxable incomes (wages, interest, dividends, capital gains, retirement incomes, and incomes from business) that together add up to what Form 1040 refers to as "total income." The bottom third of the front page of Form 1040 is where taxpayers enter the statutory adjustments needed to solve for AGI.[11]

In general, the incomes collected in the SCF are already taxable and thus correspond directly to TAXSIM inputs. One exception is Social Security. The SCF value is clearly gross benefits, but the share that is taxable varies between 0 and 85 percent, depending on the taxpayer's other incomes (technically, all other income plus half of Social Security benefits and non-taxable interest). Single individuals with other incomes less than $25,000 include no Social Security benefits in their AGI, those between $25,000 and $34,000 other income include half, and those above $34,000 in other income pay tax on 85 percent of their benefits. Imposing the condition brings gross and taxable SCF Social Security benefits in concordance with gross and taxable benefits in published IRS tables.

Other SCF variables are used to estimate the major statutory adjustments needed to solve for AGI. The deduction for one-half of self-employment taxes is estimated using Schedule C (proprietor's) income, tax rates, and other tax parameters. We estimate Social Security taxes (up to the taxable maximum) and Medicare (unlimited in our sample period) separately.

Alimony paid is solved for by starting with the SCF question on total alimony and child support paid but set to zero in cases where the respondent reports having own children under age 18 living outside the household, because that suggests the payment is more likely child support than alimony (and brings the estimated deduction in line with published IRS reports).

Student loan interest became deductible for some taxpayers in the last part of our sample, and we use information from the SCF education loan module to populate that TAXSIM inputs. Our estimate is based on outstanding loan balances, interest rates, and whether the household reports making payments. Finally, contributions to IRAs outside employer sponsored plans are collected in the financial assets module after respondents are asked about balances, but unfortunately the questions are only available in recent survey waves. Although not shown here, our estimated adjustments to income align well with published SOI data.[12]

11 The "bottom third" of the front page of Form 1040 was moved to a separate "Schedule 1" after the 2017 tax law changes went into effect. Although the content of the AGI adjustments section was unchanged, moving AGI adjustments from Form 1040 to Schedule 1 made Form 1040 appear shorter (and thus "simpler") for the taxpayers who would have otherwise entered zeroes in the AGI adjustment lines.

12 Interested users can inspect the benchmarking files that are part of the download package described in Appendix A.

Comparisons to Published SOI Data Part 2: Taxable Incomes

Most household surveys have total income below corresponding published aggregates, and for a very simple reason: rich people are less likely to take surveys, and rich people have a lot of income. The SCF is unique among public-use household surveys because it oversamples wealthy households and is thus much better at capturing the top of the income distribution. Indeed, estimated top income shares in the SCF align closely with widely accepted values based on synthesized administrative and survey data.[13] However, that does not necessarily mean the SCF income measures will align with SOI taxable incomes, either in total or for any given component.

Taxable incomes in our tax unit version of the SCF are above the corresponding published SOI values (Table 2). The gap is some 10 to 20 percent across the nine tax years in our sample, and there is little evidence of trend. Most years the ratio is between 100 and 114 percent—so very systematic—and there is no clear trend in the ratio over time. Tax years 2003 and 2015 stand out in terms of the gap between SCF and SOI incomes, with a gap closer to 20 percent. In tax year 2018, the SCF has 11 percent more income than in published SOI tables, which amounted to just over $1.3 trillion in additional taxable income.

Although a comprehensive income by income analysis is a topic for future research, there is one source of income that stands out in terms of aggregate divergence. Our category of business income combines sole proprietor, S-corporation, partnership, rent, royalty, and other incomes generally reported on Schedules C and E, along with net operating losses. The gap between SCF and SOI business incomes is proportionally much larger than the overall income gap, and thus accounts for a disproportionate share of the overall gap (Table 3). In tax year 2018, SCF respondents reported nearly $1 trillion more in business incomes, and that in turn accounted for about 75 percent of the overall income gap.[14]

The systematic differences in taxable income aggregates map directly into consistent differences between SCF and SOI income distributions (Table 4). Considering both number of returns and total taxable income across six broad AGI classes reveals a stable pattern over time. The SCF has many fewer filers in the "None" AGI category, which includes business owners whose losses exceed their positive incomes (Appendix C has additional detail and for every year of the sample). The SCF has more filers and more taxable income in the two highest AGI categories shown, which is consistent with business owners *not* reporting losses in the same way they do for tax purposes. Finally, tying back to the point about returns by filing status above, most of the gap in returns is in the $1 to $25,000 AGI class, which is where one would expect to find the SCF dependent filers we are not currently simulating.

IV. Itemized Deductions

The first step in moving from AGI to tax liability involves subtracting deductions. Taxpayers are entitled to take either a fixed standard deduction associated with their filing status, or they can claim itemized deductions reported on Schedule A. As with the adjustments to income needed to solve for AGI described in the previous section, some itemizable expenses are directly observed or can be inferred using SCF data, while other expenses must be imputed using published SOI tables. Although we identify "likely" itemizers based on a comparison of measured itemizable expenses and the available standard deduction for the tax unit, the itemizing decision and estimated itemized deductions (including AGI limitations for the tax years before 2018) are ultimately determined by TAXSIM.

Itemizable Expenses

In most of the tax years spanned by our SCF sample, nearly one third of taxpayers chose to claim itemized deductions, meaning the sum of their itemizable expenses exceeded the standard deduction. That changed

13 See the discussion of SCF versus Congressional Budget Office incomes in Bricker, et. al. 2016.

14 One possibility is that SCF business owners are misreporting the wages they paid to themselves as a component of business income. Given access to IRS tax data, one could address that directly by comparing the distributions of wages reported by SCF business owners with wages of Schedule C and E filers in the tax data.

dramatically after 2017—when increases in the standard deduction and limits on deductions for state and local taxes implemented as part of the 2017 tax reforms made the standard deduction more beneficial for nearly half of the taxpayers who would have itemized under prior year tax laws.

The shift in itemizing behavior after 2017 underscores the need to compare the benefits of itemizing relative to claiming the standard deduction for every tax unit in our sample. Modeling itemizing behavior involves adding up potentially itemizable expenses for every SCF tax unit and then comparing the value of such expenses to the available standard deduction for that tax unit. As with AGI adjustments, the SCF does not directly collect information about itemized deductions, but it is possible to model itemizing behavior using information available in the survey along with published SOI tabulations.

The list of itemizable expenses allowed under the income tax is detailed in Schedule A of Form 1040. We model deductions across ten broad categories: interest expense, real estate taxes, charitable contributions, state and local income and sales taxes, state and local personal property taxes, state and local other personal taxes, theft/casualty losses, limited miscellaneous expenses, and unlimited miscellaneous expenses. The first three components of itemizable expenses are directly observed or easily inferred in the SCF. The remaining seven components are imputed using published SOI data.

The first observed itemizable expense is interest paid. SCF respondents are not asked directly about the interest they paid during the tax year, but they are asked details about several different types of loans. [15] The details include outstanding balance and annual interest rate, and—to a first approximation—the interest paid and deductible for tax purposes is just the interest rate multiplied by the outstanding balance. The SCF questions allow us to distinguish and include only interest paid on real estate holdings not owned by a business. We also apply the statutory limits on loan balances ($1 million on loans taken out between 1988 and 2017, $750,000 after 2017) and thus limit the estimated interest paid expense.

When we add up the potentially itemizable interest paid in the survey, the total is—as expected—greater than the interest paid deduction as reported by SOI in published tables.[16] This makes sense because many of the survey respondents for whom we observe interest paid will ultimately choose the standard deduction, and the SOI tables only capture itemizable expenses for those who ultimately choose to itemize. However, the value of deductible interest paid in the survey is closer to published SOI values, because the SCF respondents with higher interest paid are more likely to ultimately choose to itemize.

The SCF questions on real estate taxes and charitable contributions are more directly related to the information needed to compute itemizable expenses. Respondents are asked to report real estate taxes paid during the year prior to the survey. Charitable contributions are also collected for the year prior to the survey—the tax year we are modeling—but respondents are only asked about "substantial" contributions, where substantial is a total amount above $500. Thus, we are in principle missing charitable contributions below $500 made by taxpayers who choose to itemize. As with interest paid, the sum of potentially itemizable real estate taxes and charitable contributions in the survey are above the totals in published SOI tables, because some the survey respondents who report real estate taxes and charitable contributions will ultimately choose the standard deduction.

Interest paid, real estate taxes, and charitable contributions account for about half of itemized deductions in published SOI tables. Our strategy for imputing the other half of itemized deductions—state and local income and sales taxes, state and local personal property taxes, state and local other personal taxes, theft/casualty losses, limited miscellaneous expenses, and unlimited miscellaneous expenses—is determined by the data available on tax returns and published in SOI tables. As with the three categories of itemizable expenses in the SCF, the published SOI data only captures spending for those who ultimately itemize. Thus, the strategy involves using survey collected interest paid, real estate taxes, and charitable

15 Technically, we should only be counting interest deducted on loans that were outstanding in the year prior to the survey. However, imposing that condition in the SCF is problematic. If a respondent has a new loan because they sold an existing mortgaged home and bought a new house, or because they refinanced an existing loan, they paid interest in the year prior to the survey. The loans in force during that prior year are not captured in the survey, and thus our best estimate of the interest they paid is based on the new loan information.

16 These comparisons are based on the SOI Table 2.1 series.

contributions to determine if a respondent is *likely* to itemize, then imputing the missing expenses for just those likely itemizers. If the sum of the observed itemizable expenses is greater than 50 percent of the standard deduction, that respondent is deemed a likely itemizer.

Given that a tax unit is identified as a likely itemizer, we use SOI data on itemization by income to assign the missing half of itemizable expenses. SOI Table 2.1 has counts and dollars of each of the seven expenses by income class for itemizing tax returns. We assign the incidence of a given expense using probability of having that expense—the number claiming the expense divided by the total number of itemizers in the given income category. We then assign a conditional dollar amount using the average expense within the income category and adjust that for where in the income range the SCF tax unit is observed.

Comparisons to Published SOI Data Part 3: Itemizers and Itemized Deductions

Our approach to estimating itemizing behavior generates reasonable results over most of the sample period but slightly understates the drop in itemizing between 2015 and 2018 (Table 5). In most tax years, the SCF tax unit file (after running through TAXSIM) has an itemizer count within a few percentage points of published SOI values, though we are underestimating itemizing behavior in 2012 and 2015 by more substantial margins. Even though the estimated number of itemizers in the SCF tax unit file fell by more than half between 2015 and 2018 as higher standard deductions kicked in, the SOI reported itemizer count fell even more, and the estimated itemizer count in the SCF is above the published SOI in that year.

The ratio of SCF to SOI itemizable expenses is generally higher and more variable. Itemizable expenses for most taxpayers are modest relative to their incomes, including things like mortgage interest, state and local taxes paid, real estate taxes, and regular modest charitable contributions. However, some categories of deductible spending such as large charitable contributions and "miscellaneous" expenses are more the domain of the wealthy, and thus harder to pin down. Even so, the estimated itemized deductions in our SCF tax unit file are generally 10 to 20 percent above published SOI values, and though volatile, there is little trend in the ratio over time.

V. Estimating Tax Liability Using TAXSIM

Much of the work needed to estimate income tax liability in the SCF micro data files is described in the preceding sections. We have described creating tax units within households, assigning filing status and dependents, mapping incomes into taxable counterparts, subtracting statutory income adjustments, and estimating itemizable expenses. Estimating tax liability requires the additional step of processing those inputs through complex income tax formulas that vary over time. Thankfully, survey data users can make use of the NBER TAXSIM program to complete that final step.[17]

We use TAXSIM version 32. In addition to generating tax liabilities using survey income values under the tax law applicable for that survey year, TAXSIM makes it possible to run simulations with popular alternative tax rules, or even the tax law that existed in years other than the year in which the survey was conducted.

TAXSIM begins with 32 user inputs for incomes, demographics, itemizable expenses, and other key variables, and returns estimated tax liability.[18] The correspondence between SCF tax unit file variables and TAXSIM inputs is shown item-by-item in Appendix B. Much of the discussion about constructing variables in the preceding sections is focused on creating the necessary inputs for TAXSIM, and thus the table in Appendix B provides a correspondence between the underlying SCF information and what we are feeding into TAXSIM.

17 The latest version of TAXSIM is available at https://users.nber.org/~taxsim/.

18 TAXSIM calculates all limits, phaseouts, etc. For example, we provide itemizable interest for each tax unit to TAXSIM and allow it to add in statutory limits and arrive at the deductible interest amount. Ultimately, TAXSIM will determine whether that tax unit itemizes.

TAXSIM generates a long list of output variables. The two outputs we focus on here are federal income tax before credits and federal income tax after credits. One of the reasons to carefully track ages of dependents, for example, is so that TAXSIM can compute credits which depend on age, such as the EITC and Child Tax Credit. However, there are key inputs to credits (such as child-care expenses) that we have not yet attempted to model, and other credits (such as for education) that are not directly modeled in TAXSIM.

Comparisons to Published SOI Data Part 4: Estimated Tax Liability

The SCF tax unit file is close to published SOI in terms of tax filing but has noticeably more income. In addition, we know that much of the incremental income is in the top tail, which (because we have a progressive tax system) suggests those incremental dollars are subject to an above average tax rate. Thus, the expectation is that running the SCF tax unit file through TAXSIM will generate aggregate taxes that exceed SOI reported values, and by more than the gap in taxable incomes.

The expectations about SCF TAXSIM tax liability are borne out (Tables 7, 8). In general, estimated tax liability varies between 15 and 30 percent above published SOI tables. The ratio varies somewhat between before and after credits, which suggests further research on inputs for credits (or how TAXSIM models credits) is warranted. However, the ratios of estimated to published tax liabilities in Tables 7 and 8 vary over time with the ratio of SCF to SOI published taxable incomes shown in Table 2.

Much remains to be done in terms of reconciling SCF and SOI taxable incomes, as well as other inputs to income tax calculations. But if there is indeed incrementally useful information about (especially business) incomes in the SCF, these TAXSIM output tables provide a glimpse of what that information would suggest for overall income tax revenues. The magnitude and location (in the income distribution) of that income suggests substantial foregone revenue.

VI. Conclusion

Having measures of tax liabilities for SCF households will undoubtedly prove useful in a wide variety of contexts. A specific motivation for this paper is developing a better understanding of the taxation of business income. Income aggregates—particularly related to business income—vary substantially across sources and are lowest in the income tax data published by the SOI, followed by the SCF, and highest in the National Income and Product Accounts (NIPAs). Alternative ways of resolving these differences lead to first order differences in estimated overall income distributions. The debate over distributing "missing" business income plays an important role in the differences between estimates of top income shares (Piketty, Saez, and Zucman 2018; Auten and Splinter 2019). Kopczuk and Zwick (2020) discuss the issues with measuring business income. See also Bricker, et al. (2016) and Sabelhaus and Park (2020). Likewise, the lack of wealth data in administrative SOI micro tax return files is also the basis for the debate over appropriate "capitalization" of taxable incomes to estimate wealth, as reflected in the debate over top wealth shares (Bricker, et al., 2016, Saez and Zucman, 2016, and Smith, Zidar, and Zwick, 2021).

Although some researchers (e.g., Bhandari, et al., 2020) argue that SCF business income data are therefore unreliable, we show there is important information in the differences. Unlike the SOI, the SCF contains data on wealth and both taxable and non-taxable incomes for the same households. Because we can use reported business ownership and business wealth to help assign missing business incomes to the appropriate business owners, replicating tax filings in the SCF may help resolve outstanding questions about income distribution. We explore these issues further in a companion paper (Gale, et al., 2021) that uses the methodology developed here.

Figure 1.

Figure 1A. Married Filing Joint and Widow(er) Returns

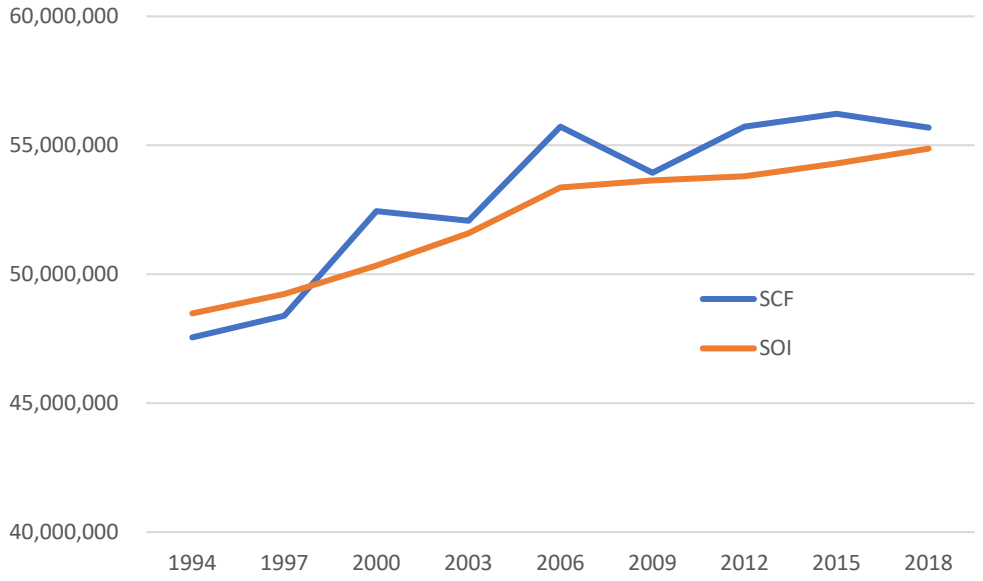

Figure 1B. Married Filing Single Returns

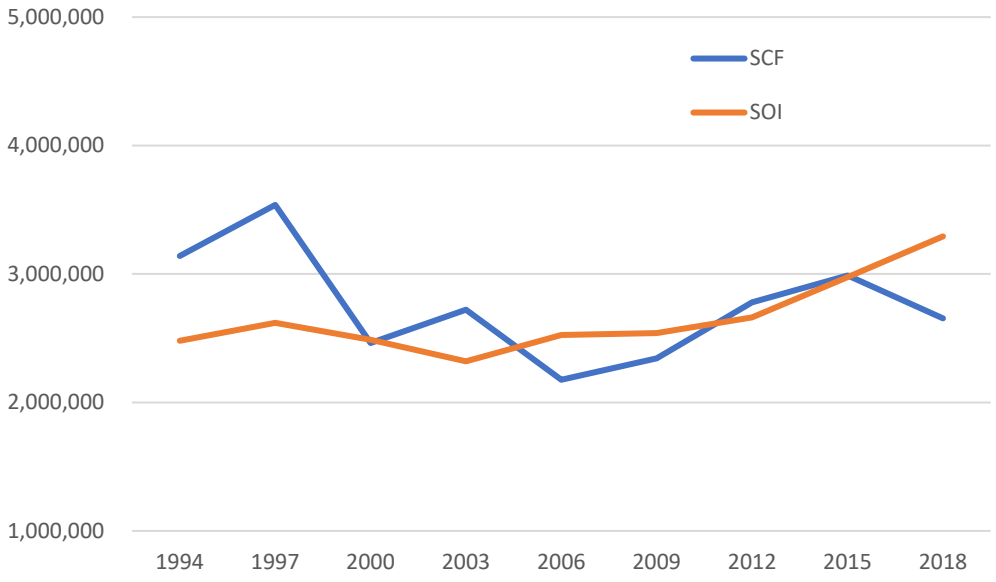

Figure 1C. Head of Household Returns

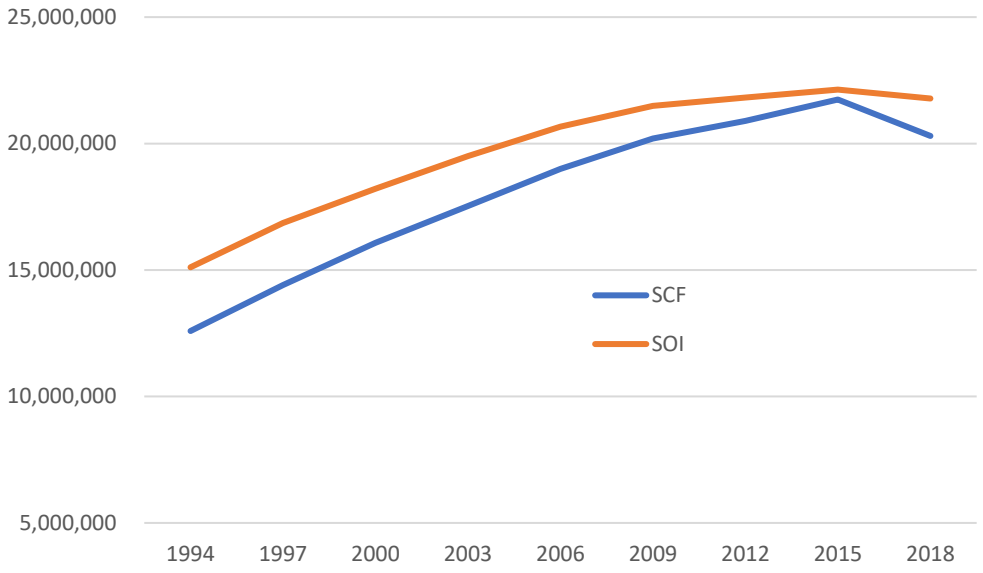

Figure 1D. Single Returns

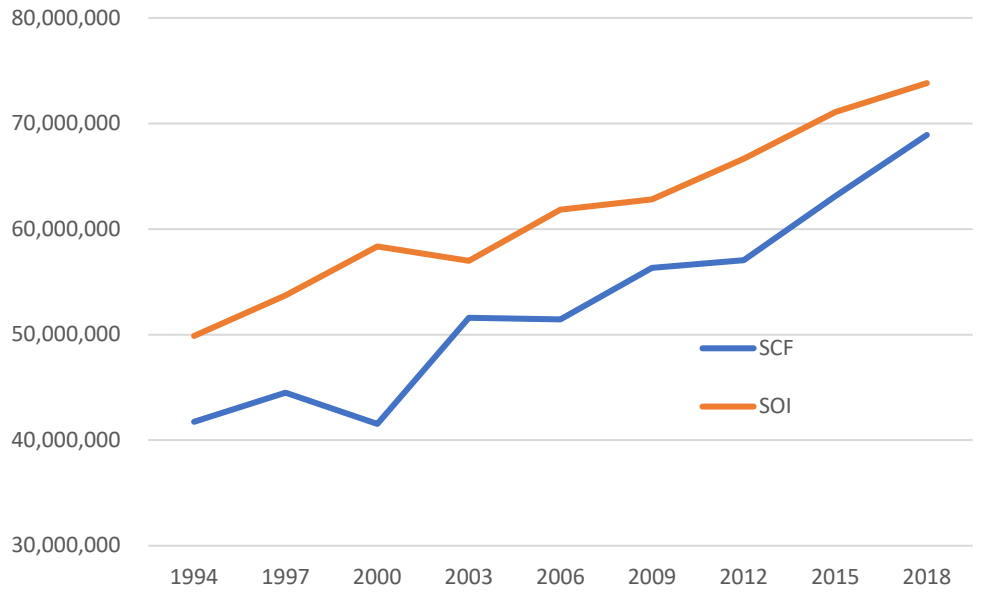

Table 1. Simulated (SCF) and Actual (SOI) Tax Returns (Millions)

Year	SCF	SOI	SCF/SOI	SOI - SCF
1994	105.0	115.9	90.6%	10.9
1997	110.8	122.4	90.5%	11.6
2000	112.5	129.4	87.0%	16.9
2003	123.9	130.4	95.0%	6.5
2006	128.3	138.4	92.7%	10.1
2009	132.8	140.5	94.5%	7.7
2012	136.5	144.9	94.2%	8.5
2015	144.1	150.5	95.7%	6.4
2018	147.6	153.8	96.0%	6.2

Table 2. Estimated (SCF) and Actual (SOI) Taxable Income (Billions)

Year	SCF	SOI	SCF/SOI	*SCF - SOI*
1994	4,537.6	3,958.6	114.6%	579.0
1997	5,542.7	5,016.9	110.5%	525.8
2000	7,134.7	6,261.8	113.9%	872.9
2003	7,543.5	6,295.2	119.8%	1,248.3
2006	9,265.6	8,144.7	113.8%	1,121.0
2009	8,825.9	7,739.4	114.0%	1,086.5
2012	10,054.2	9,122.4	110.2%	931.8
2015	12,173.2	10,360.4	117.5%	1,812.8
2018	13,116.6	11,785.3	111.3%	1,331.3

Table 3. Estimated (SCF) and Actual (SOI) Taxable Business Income (Billions)

Year	SCF	SOI	SCF/SOI	SCF - SOI
1994	531.0	230.4	230.5%	300.6
1997	723.9	305.2	237.2%	418.8
2000	836.3	381.4	219.2%	454.8
2003	850.6	421.6	201.8%	429.0
2006	1,300.7	614.2	211.8%	686.5
2009	1,162.3	446.8	260.1%	715.5
2012	1,337.7	665.3	201.1%	672.4
2015	1,701.4	771.8	220.5%	929.6
2018	1,825.6	830.1	219.9%	995.5

Table 4. Estimated (SCF) and Actual (SOI) Returns and Taxable Income By AGI

Tax Year 1994

AGI Class	Number of Returns			Total Income (Millions)		
	SCF	SOI	SCF/SOI	SCF	SOI	SCF/SOI
None	317,247	953,210	33.3%	-8,887	-55,827	15.9%
$1 to Under $25,000	49,581,333	63,406,306	78.2%	567,690	706,911	80.3%
$25,000 to Under $50,000	29,348,107	29,158,867	100.6%	1,051,576	1,037,238	101.4%
$50,000 to Under $100,000	19,024,047	17,910,518	106.2%	1,291,278	1,200,710	107.5%
$100,000 to Under $1,000,000	6,625,159	4,444,294	149.1%	1,301,720	848,846	153.4%
$1,000,000 or More	128,463	69,935	183.7%	334,218	220,687	151.4%
Total	***105,024,356***	***115,943,130***	***90.6%***	***4,537,595***	***3,958,565***	***114.6%***

Tax Year 2006

AGI Class	Number of Returns			Total Income (Millions)		
	SCF	SOI	SCF/SOI	SCF	SOI	SCF/SOI
None	841,335	2,675,594	31.4%	-11,135	-89,369	12.5%
$1 to Under $25,000	42,545,343	55,981,752	76.0%	519,433	692,416	75.0%
$25,000 to Under $50,000	33,061,532	33,588,778	98.4%	1,189,421	1,227,183	96.9%
$50,000 to Under $100,000	30,422,151	29,995,325	101.4%	2,133,311	2,144,784	99.5%
$100,000 to Under $1,000,000	20,721,814	15,799,214	131.2%	3,952,853	2,950,766	134.0%
$1,000,000 or More	598,629	354,093	169.1%	1,481,764	1,218,909	121.6%
Total	***128,190,803***	***138,394,756***	***92.6%***	***9,265,647***	***8,144,688***	***113.8%***

Tax Year 2018

AGI Class	Number of Returns			Total Income (Millions)		
	SCF	SOI	SCF/SOI	SCF	SOI	SCF/SOI
None	507,115	1,962,253	25.8%	-11,002	-200,109	5.5%
$1 to Under $25,000	38,914,043	50,453,810	77.1%	476,879	647,707	73.6%
$25,000 to Under $50,000	37,928,376	36,512,304	103.9%	1,378,423	1,340,764	102.8%
$50,000 to Under $100,000	37,577,727	35,146,085	106.9%	2,698,081	2,534,215	106.5%
$100,000 to Under $1,000,000	31,680,727	29,160,637	108.6%	6,466,956	5,670,128	114.1%
$1,000,000 or More	906,347	539,207	168.1%	2,107,287	1,792,574	117.6%
Total	***147,514,334***	***153,774,296***	***95.9%***	***13,116,623***	***11,785,278***	***111.3%***

Table 5. Number of Itemizers (Thousands)

Year	SCF TAXSIM	SOI	SCF TAXSIM/SOI	SCF TAXSIM-SOI
1994	33,084	33,018	100%	66
1997	38,237	36,625	104%	1,613
2000	41,738	42,534	98%	-796
2003	44,565	43,950	101%	616
2006	48,251	49,124	98%	-872
2009	43,422	45,696	95%	-2,274
2012	41,030	45,582	90%	-4,552
2015	41,081	44,567	92%	-3,486
2018	19,557	17,533	112%	2,024

Table 6. Value of Itemized Deductions (Millions)

Year	SCF TAXSIM	SOI	SCF TAXSIM/SOI	SCF TAXSIM-SOI
1994	575	480	120%	95
1997	696	598	117%	99
2000	908	784	116%	124
2003	1,056	873	121%	183
2006	1,390	1,194	116%	195
2009	1,306	1,191	110%	114
2012	1,282	1,239	103%	43
2015	1,459	1,210	121%	249
2018	801	649	124%	153

Table 7. Tax Liabilities Before Credits (Billions)

Year	SCF TAXSIM	SOI	SCF TAXSIM/SOI	SCF TAXSIM-SOI
1994	722.6	541.6	133.4%	181.1
1997	894.0	739.5	120.9%	154.5
2000	1,238.5	1,018.2	121.6%	220.3
2003	1,057.0	790.0	133.8%	267.0
2006	1,323.7	1,082.9	122.2%	240.8
2009	1,206.5	976.0	123.6%	230.5
2012	1,417.3	1,261.0	112.4%	156.3
2015	1,935.9	1,516.2	127.7%	419.7
2018	1,903.6	1,651.8	115.2%	251.8

Table 8. Tax Liabilities After Credits (Billions)

Year	SCF TAXSIM	SOI	SCF TAXSIM/SOI	SCF TAXSIM-SOI
1994	708.2	532.6	133.0%	175.5
1997	873.6	727.3	120.1%	146.2
2000	1,200.0	980.5	122.4%	219.5
2003	996.8	748.0	133.3%	248.8
2006	1,265.1	1,023.9	123.6%	241.2
2009	1,092.5	865.9	126.2%	226.6
2012	1,357.6	1,188.0	114.3%	169.6
2015	1,933.3	1,435.8	134.6%	497.5
2018	1,765.5	1,509.9	116.9%	255.6

References

Auten, Gerald, and David Splinter, 2019. "Income Inequality in the United States: Using Tax Data to Measure Long-term Trends," manuscript.

Bhandari, Anmol, Serdar Birinci, Ellen R. McGrattan, and Kurt See. 2020. "What Do Survey Data Tell Us About US Businesses?" *AER Insights*, 2(4):443-58.

Bhutta, Neil, Jesse Bricker, Andrew C. Chang, Lisa J. Dettling, Sarena Goodman, Joanne W. Hsu, Kevin B. Moore, Sarah Reber, Alice Henriques Volz, and Richard A. Windle. 2020. "Changes in U.S. Family Finances from 2016 to 2019: Evidence from the Survey of Consumer Finances," *Federal Reserve Bulletin*, 106(5): 1-42. (September)

Bricker, Jesse, Alice Henriques, Jacob Krimmel, and John Sabelhaus. 2016. "Measuring Income and Wealth at the Top Using Administrative and Survey Data," *Brookings Papers on Economic Activity, 1:2016*, p. 261-321.

Bricker, Jesse, Kevin B. Moore, Sarah J. Reber, and Alice Henriques Volz. 2020. "Effective Tax Rates by Income and Wealth Class," *National Tax Journal*, 73(4), 987-1004.

Congressional Budget Office, 2021. *The Distribution of Household Income, 2018*. (August)

Feenberg, Daniel. and Elisabeth Coutts. 1993. "An Introduction to the TAXSIM Model," *Journal of Policy Analysis and Management,* 12(1). (Winter)

Gale, William G., Swati Joshi, Christopher Pulliam, and John Sabelhaus. 2021. "Taxing Business Incomes," The Brookings Institution. (December)

Hertz, Tom, Pat Langetieg, Mark Payne, Alan Plumley, and Margaret Jones. 2020. "The Extent of Individual Income Tax Nonfiling," Internal Revenue Service and U.S. Census Bureau. (October).

Jones, Maggie R., and Amy B. O'Hara. 2016. "Do Doubled-Up Families Minimize Tax Burdens?" *National Tax Journal*, 69 (3): 613–640. (September)

Kopczuk, Wojciech and Eric Zwick. 2020. "Business Incomes at the Top," *Journal of Economic Perspectives*, 34(4): 27-51. (Fall)

Meyer, Bruce D., Derek Wu, Grace Finley, Patrick Langetieg, Carla Medalia, Mark Payne, and Alan Plumley. 2021. "The Accuracy of Tax Imputations: Estimating Tax Liabilities and Credits Using Linked Survey and Administrative Data," Forthcoming in *Measuring and Understanding the Distribution and Intra/Inter-Generational Mobility of Income and Wealth*, National Bureau of Economic Research, Studies in Income and Wealth. Chicago: University of Chicago Press.

Moore, Kevin B., Sarah J. Pack, and John Sabelhaus. 2016. "Taxing Incomes of Top Wealth Holders," *National Tax Journal*, 69(4), 965-980.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman, 2018. "Distributional National Accounts: Methods and Estimates for the United States," *Quarterly Journal of Economics* 133, no. 2: 553-609.

Sabelhaus, John, and Somin Park. 2020. "U.S. income inequality is worse and rising faster than policymakers probably realize." Washington Center for Equitable Growth, , available at https://equitablegrowth.org/u-s-income-inequality-is-worse-and-rising-faster-than-policymakers-probably-realize/. (May)

Saez, Emmanuel and Gabriel Zucman. 2016. "Wealth Inequality in the United States Since 1913: Evidence from Capitalized Income Tax Data," *Quarterly Journal of Economics*, 131(2):519–578.

Smith, Matthew, Owen Zidar, and Eric Zwick. 2021. "Top Wealth in the United States: New Estimates and Implications for Taxing the Rich," U.S. Available at: https://scholar.princeton.edu/sites/default/files/wealth_main_resubmitted.pdf.

Warren, Lewis, Liana Fox, and Ashley Edwards. 2020. "The Supplemental Poverty Measure in the Survey of Income and Program Participation," U.S. Census Bureau, SEHSD Working Paper 2020-20. (November)

Wheaton, Laura, and Kathryn Stevens. 2016. *The Effect of Different Tax Calculators on the Supplemental Poverty Measure*. Washington, D.C.: Urban Institute.

Appendix A. Downloading and Running the SCF Tax Units and TAXSIM Programs

The programs, data sets, and auxiliary files associated with this project are available for download and use. Interested users should contact the authors to gain access to the download.

The main Stata .do file is scf_taxsim_main.do. That program calls three main subprograms: public_use_age_recode.do, compute_scf_itemizable_expenses.do, and impute_missing_itemizable_expenses.do. Together, these programs have nearly 4,000 lines of code, and the program run time is approximately 30 minutes.

The program set_file_paths.do is key to running the program in the user's specific environment. The Stata programs all refer to file paths using Stata globals such as "$datapath." Those globals are set relative to the main directory in which the project is being run. In the default configuration that directory is c:/scf/taxsim. Users can adapt the programs to their specific environment by changing the folder references in set_file_paths.do. Users will also need to make sure their computing environment is compatible with the TAXSIM call by running the following from a Stata command line:

```
net from "https://www.nber.org/stata"
net describe taxsim32
net install taxsim32
```

Full documentation for the TAXSIM 32 Stata ado file is available on the TAXSIM website, https://users.nber.org/~taxsim/taxsim32/stata-remote.html.

The SCF input files are all accessible from the Federal Reserve Board website, https://www.federalreserve.gov/econres/scfindex.htm.

There are specific versions of the SCF files in the download directory. Although SCF files are not generally updated, errors have been fixed at various times so newer files may be a little different. The summary variable "bulletin" files used in the project are all nominal, whereas the default files maintained on the FRB website have dollar variables in real terms. There is a program called public_stata_merge.do called near the top of the scf_taxsim_main.do program that combines extracts from SCF core and bulletin files in a particular way. Users can replace the result of this program with their own merged file.

The main program reads a few other Excel input files as well. The first is scf_taxsim_main_parameters, which has key income tax inputs by year. The second is public_use_age_recodes, which is used to assign specific ages given the (disclosure-avoidance) bracketed ages in the public-use SCF data. The third is itemizable_expenses_imputation, which uses published SOI data on itemized deductions by income to impute the missing expense categories given the subset of itemizable expenses that are observed in the SCF.

The program generates statistical outputs in the form of "benchmarking" files. The tables and graphs in this paper are taken directly from the Excel output files Filing_Status_Benchmarking, Incomes_Benchmarking, Itemized_Deductions_Benchmarking, and Tax_Liability_Benchmarking. Those files are automatically updated every time the program is run, and users can verify that their updated versions match the downloaded files.

Finally, the program produces a household level Stata output file scf_taxsim_out.dta, which aggregates tax liabilities for all tax units in the household. Those estimates can be merged with any public use SCF file using the variables y1 and year.

Appendix B. TAXSIM and SCF Variable Concordance

TAXSIM 32 has up to 32 distinct input variables for each tax return. The table below provides an overview of how TAXSIM inputs are constructed from SCF variables, though many details are left to the comments in the Stata code.

Table A1. SCF-TAXSIM Variable Crosswalk

TAXSIM Variable	**TAXSIM definition**	**SCF Concept and Variable Name**
taxsimid	Arbitrary Case ID	Arbitrary Case ID assigned
state	State of residence	No state in public use SCF, so set state=0, meaning TAXSIM does not calculate state taxes
mstat	Marital Status (unmarried, married, separate, or dependent taxpayer). TAXSIM determines whether individual taxpayer is filing singly or as head of household	Starting point is SCF households, code comments describe rules for splitting households into one or more tax units, and presence of spouse-partner and marital status/history determines mstat
page	Age of primary taxpayer	Age of designated tax unit primary
sage	Age of spouse	Age of designated tax unit secondary if present
depx	Number of dependents	Allocation of household members across tax units explained in code
dep13	Number of children under 13 with	Allocation of household members

	eligible childcare expenses for Dependent Care Credit.	across tax units explained in code, this group is ages<17
dep17	Number of children under 17 for the entire tax year for the Child Care Credit. This includes children under 13.	Allocation of household members across tax units explained in code, this group is ages <17
dep18	Number of qualifying children for EITC.	Allocation of household members across tax units explained in code, this group is ages <18 or ages18-24 and financially dependent (assumed student)
pwages	Wages and salary of primary taxpayer (includes self-employment but not QBI)	Starting point is SCF variable X5702, wages and salaries from income module Also includes X5704, Schedule C income, in all years before 2019 Schedule C income is allocated between QBI and SSTB categories in 2019 and later SCFs (see below)

		Allocation between primary and secondary in split R/SP households depends on labor income from work and earnings module For tax units that are constructed from non-primary economic units (NPEUs), we use salaries and wages and divide by the number of NPEU members that are under 70 years old
swages	Wages and salary of spouse (includes self-employment but not QBI)	See description above for pwages, but Schedule C income always assigned to primary
dividends	Dividend income (qualified dividends only for 2003 and beyond)	Starting point is SCF variable X5710, total dividends 50/50 allocation between primary and secondary respondents in split R/SP households All dividends are mapped into this

		category through the 2004 SCF survey, and 75 percent (the average ratio of qualified to total dividends) from that point forward; the remaining 25 percent goes to otherprop (see below)
intrec	Interest received	Starting point is SCF variable X5708, total taxable income 50/50 allocation between primary and secondary respondents in split R/SP households
stcg	Short term capital gains or losses (+/-)	Starting point is SCF variable X5712, total capital gains, which can be positive or negative 50/50 allocation between primary and secondary respondents in split R/SP households We use aggregate data on long-term and short-term capital

		gains/losses from the SOI Individual report (Table 1.4) to determine the share of gains/losses that are long and short term for three broad AGI classes (<50k, 50-100k, 100k+); the shares from this are then applied to the data by AGI class
ltcg	Long term capital gains or losses (+/-)	See stcg above
otherprop	Other property income subject to NIIT, including unearned or limited partnership and passive S-Corp profits; rent not eligible for QBI deduction; non-qualified dividends; and other income or loss.	Starting point is SCF variable X5714, income (+/-) reported on Schedule E, includes S-Corps, partnerships, rent, royalties Includes 25 percent of dividends (see above) in 2004 SCF and later 50/50 allocation between primary and secondary respondents in split R/SP households Schedule E income is allocated

		between active S Corp and QBI categories in 2019 and later SCFs (see below)
nonprop	Other non-property income not subject to Medicare NIIT such as alimony, nonwage fellowships, and state income tax refunds (itemizers only). Also includes adjustments and items such as alimony paid, Keogh/IRA contributions, foreign income exclusion, and NOLs, which can be entered as negative income.	Starting point is SCF variables X5718 (child support and alimony) plus parts of X5724 (other income) deemed "other" taxable or net operating loss carry forward (see codes in program) Subtractions for one-half self-employment tax, student loan interest, alimony paid, and deductible IRA contributions (see program for details) 50/50 allocation between primary and secondary respondents in split R/SP households
pensions	Taxable pensions and IRA distributions	Starting point is SCF variable X5722 (retirement income) but we also use Social Security incomes from labor modules to

		split total between pensions/Social Security and between R and SP
gssi	Gross Social Security Benefits	Starting point is SCF variable X5722 (retirement income) but we also use Social Security incomes from labor modules to divide total between pensions and Social Security We also directly observe social security income for NPEUs as a total and divide by the number of NPEU members that are over 62 to determine social security income
ui	Unemployment compensation received	Starting point is SCF variable X5716 If a household was split, the default is a 50/50 split of unemployment income. If we can directly observe if either the

		respondent or spouse was unemployment and the household was split, we allocate the full amount of unemployment income to the unemployed tax unit.
transfers	Non-taxable transfer incomes such as welfare, workers comp; veterans' benefits; and child support that would affect eligibility for state property tax rebates but would not be taxable at the federal level.	Starting point is SCF variable X5720, government transfers
rentpaid	Rent paid (only used for calculating state property tax rebates)	See code, uses SCF variables that capture rent for renter households
proptax	Real estate taxes paid – this is a preference for AMT and additionally used to calculate state property tax rebates	See code, directly observed in SCF
otheritem	Other itemized deductions that are a preference for AMT, including other state and local taxes (line 8, Schedule A) plus local income	See code, imputed using "likely" itemizer based on reported itemizable expenses, with incidence and average values

	tax; preference share of medical expenses (line 16); and miscellaneous deductions (line 27)	from published SOI tables
childcare	Childcare expenses	Currently set to zero
mortgage	Deductions that are not included in line 25 and are not a preference for AMT including: deductible medical expenses not on line 16; motor vehicle taxes paid (line 7); home mortgage interest (line 15); charitable contributions (line 19); and casualty and theft losses (line 20)	See code, combination of observed SCF components (interest paid and charitable contributions) and imputed items using "likely" itemizer based on reported itemizable expenses, with incidence and average values from published SOI tables
scorp	Active S-scorp income for a specialized service trade or business service (SSTB)	Zero through SCF year 2016, then part of Schedule E income (see above)
pbusinc	Primary taxpayer's qualified business income (QBI) subject to a preferential rate without phaseout. Assumes sufficient wages paid/capital to be eligible for the full deduction. Subject to	Zero through SCF year 2016, then part of Schedule E and Schedule C income (see above)

	SECA and Medicare Additional Earnings Tax.	
pprofinc	Primary Taxpayer's SSTB with preferential rate subject to claw-back. Subject to SECA and Medicare Additional Earnings Tax.	Zero through SCF year 2016, then part of Schedule C income (see above)
sbusinc	Spouse's QBI. Zero for non-joint returns.	No split between R and SP in current version
sprofinc	Spouse's SSTB. Zero for non-joint returns.	No split between R and SP in current version

Appendix C. Simulated Versus Actual Taxable Income Distributions

Year = 1994								
AGI Class	**Number of Returns**		**Total Income (Millions)**		**Positive Income (Millions)**		**Negative Income (Millions)**	
	SCF	**SOI**	**SCF**	**SOI**	**SCF**	**SOI**	**SCF**	**SOI**
None	317,247	953,210	-8,887	-55,827	6,446	21,224	15,333	77,050
$1 to Under $25,000	49,581,333	63,406,306	567,690	706,911	570,066	734,709	2,376	27,799
$25,000 to Under $50,000	29,348,107	29,158,867	1,051,576	1,037,238	1,056,020	1,059,668	4,444	22,430
$50,000 to Under $100,000	19,024,047	17,910,518	1,291,278	1,200,710	1,296,566	1,224,229	5,287	23,519
$100,000 to Under $1,000,000	6,625,159	4,444,294	1,301,720	848,846	1,306,585	873,290	4,865	24,445
$1,000,000 or More	128,463	69,935	334,218	220,687	334,673	228,133	455	7,446
Total	***105,024,356***	***115,943,130***	***4,537,595***	***3,958,565***	***4,570,355***	***4,141,254***	***32,760***	***182,689***

Year = 1997								
AGI Class	**Number of Returns**		**Total Income (Millions)**		**Positive Income (Millions)**		**Negative Income (Millions)**	
	SCF	**SOI**	**SCF**	**SOI**	**SCF**	**SOI**	**SCF**	**SOI**
None	443,506	918,707	-5,042	-53,017	5,479	24,413	10,521	77,430
$1 to Under $25,000	47,470,704	61,862,801	552,120	730,470	553,775	756,831	1,654	26,362
$25,000 to Under $50,000	29,273,603	30,819,522	1,053,348	1,118,735	1,056,528	1,140,658	3,180	21,923
$50,000 to Under $100,000	24,274,003	21,635,161	1,661,452	1,490,789	1,663,004	1,517,747	1,552	26,958
$100,000 to Under $1,000,000	9,114,012	7,041,341	1,775,029	1,305,510	1,777,570	1,337,524	2,541	32,013
$1,000,000 or More	236,037	144,459	505,759	424,418	508,030	437,489	2,271	13,071
Total	***110,811,866***	***122,421,991***	***5,542,667***	***5,016,905***	***5,564,385***	***5,214,662***	***21,719***	***197,757***

Year = 2000								
AGI Class	**Number of Returns**		**Total Income (Millions)**		**Positive Income (Millions)**		**Negative Income (Millions)**	
	SCF	**SOI**	**SCF**	**SOI**	**SCF**	**SOI**	**SCF**	**SOI**
None	281,978	1,146,357	-3,014	-28,948	2,691	29,553	5,704	58,501
$1 to Under $25,000	41,182,130	59,370,058	520,482	683,771	522,678	702,912	2,196	19,141
$25,000 to Under $50,000	30,161,788	32,328,575	1,080,039	1,143,631	1,084,732	1,163,524	4,694	19,893
$50,000 to Under $100,000	26,823,245	25,673,487	1,870,923	1,746,286	1,875,549	1,772,845	4,626	26,558
$100,000 to Under $1,000,000	13,635,036	10,615,341	2,604,601	1,889,535	2,622,000	1,916,360	17,399	26,825
$1,000,000 or More	437,099	239,685	1,061,688	827,539	1,071,812	833,498	10,124	5,959
Total	***112,521,275***	***129,373,503***	***7,134,718***	***6,261,815***	***7,179,461***	***6,418,691***	***44,743***	***156,876***

Year = 2003								
AGI Class	**Number of Returns**		**Total Income (Millions)**		**Positive Income (Millions)**		**Negative Income (Millions)**	
	SCF	**SOI**	**SCF**	**SOI**	**SCF**	**SOI**	**SCF**	**SOI**
None	589,245	1,813,840	-13,969	-79,313	11,029	40,165	24,997	119,478
$1 to Under $25,000	45,420,463	57,328,158	562,040	703,506	565,782	737,146	3,742	33,640
$25,000 to Under $50,000	31,638,998	32,951,454	1,143,891	1,201,721	1,154,453	1,233,796	10,562	32,075
$50,000 to Under $100,000	29,244,271	26,915,091	2,040,571	1,910,662	2,051,185	1,951,794	10,614	41,132
$100,000 to Under $1,000,000	16,656,851	11,233,799	3,023,101	2,021,038	3,044,347	2,070,638	21,246	49,601
$1,000,000 or More	328,569	181,283	787,899	537,623	790,452	555,653	2,553	18,030
Total	***123,878,397***	***130,423,625***	***7,543,534***	***6,295,236***	***7,617,248***	***6,589,192***	***73,715***	***293,957***

Year = 2006								
AGI Class	**Number of Returns**		**Total Income (Millions)**		**Positive Income (Millions)**		**Negative Income (Millions)**	
	SCF	**SOI**	**SCF**	**SOI**	**SCF**	**SOI**	**SCF**	**SOI**
None	841,335	2,675,594	-11,135	-89,369	24,578	50,018	35,713	139,387
$1 to Under $25,000	42,545,343	55,981,752	519,433	692,416	522,152	728,403	2,718	35,987
$25,000 to Under $50,000	33,061,532	33,588,778	1,189,421	1,227,183	1,193,409	1,258,311	3,988	31,129
$50,000 to Under $100,000	30,422,151	29,995,325	2,133,311	2,144,784	2,142,349	2,191,809	9,038	47,025
$100,000 to Under $1,000,000	20,721,814	15,799,214	3,952,853	2,950,766	3,965,839	3,016,957	12,986	66,191
$1,000,000 or More	598,629	354,093	1,481,764	1,218,909	1,485,371	1,260,616	3,607	41,706
Total	***128,190,803***	***138,394,756***	***9,265,647***	***8,144,688***	***9,333,698***	***8,506,114***	***68,051***	***361,425***

Year = 2009								
AGI Class	**Number of Returns**		**Total Income (Millions)**		**Positive Income (Millions)**		**Negative Income (Millions)**	
	SCF	**SOI**	**SCF**	**SOI**	**SCF**	**SOI**	**SCF**	**SOI**
None	985,346	2,511,925	-26,210	-196,405	29,779	77,303	55,989	273,708
$1 to Under $25,000	47,187,128	56,546,597	581,223	713,389	589,526	757,692	8,303	44,302
$25,000 to Under $50,000	32,601,807	33,830,451	1,170,363	1,235,713	1,181,063	1,273,983	10,700	38,270
$50,000 to Under $100,000	32,054,727	30,158,618	2,247,373	2,163,160	2,262,125	2,219,208	14,752	56,049
$100,000 to Under $1,000,000	19,413,779	17,209,654	3,885,661	3,088,492	3,930,714	3,163,051	45,054	74,559
$1,000,000 or More	457,987	236,883	967,508	735,050	976,760	766,166	9,252	31,116
Total	***132,700,774***	***140,494,128***	***8,825,918***	***7,739,398***	***8,969,968***	***8,257,403***	***144,049***	***518,005***

Year = 2012								
AGI Class	**Number of Returns**		**Total Income (Millions)**		**Positive Income (Millions)**		**Negative Income (Millions)**	
	SCF	**SOI**	**SCF**	**SOI**	**SCF**	**SOI**	**SCF**	**SOI**
None	607,730	2,128,548	-3,070	-196,459	23,989	69,555	27,060	266,014
$1 to Under $25,000	47,799,257	56,752,718	566,389	713,462	570,101	754,084	3,712	40,621
$25,000 to Under $50,000	32,979,639	34,059,304	1,187,265	1,237,422	1,191,904	1,271,336	4,639	33,915
$50,000 to Under $100,000	31,801,978	31,089,262	2,257,863	2,225,862	2,266,010	2,276,091	8,147	50,230
$100,000 to Under $1,000,000	22,609,949	20,505,789	4,506,519	3,787,940	4,539,740	3,867,801	33,221	79,861
$1,000,000 or More	635,117	392,850	1,539,241	1,354,143	1,546,938	1,402,859	7,697	48,716
Total	***136,433,670***	***144,928,471***	***10,054,206***	***9,122,370***	***10,138,682***	***9,641,726***	***84,476***	***519,356***

Year = 2015								
AGI Class	**Number of Returns**		**Total Income (Millions)**		**Positive Income (Millions)**		**Negative Income (Millions)**	
	SCF	**SOI**	**SCF**	**SOI**	**SCF**	**SOI**	**SCF**	**SOI**
None	687,310	2,072,066	-1,275	-201,562	36,379	71,770	37,654	273,332
$1 to Under $25,000	46,342,707	54,962,676	558,509	698,677	561,644	739,901	3,135	41,224
$25,000 to Under $50,000	34,783,404	35,372,173	1,262,554	1,299,296	1,267,622	1,331,210	5,068	31,913
$50,000 to Under $100,000	34,948,213	32,801,908	2,505,104	2,367,670	2,514,072	2,414,437	8,968	46,767
$100,000 to Under $1,000,000	26,313,221	24,845,104	5,354,460	4,723,441	5,380,440	4,811,584	25,980	88,142
$1,000,000 or More	949,290	439,335	2,493,838	1,472,881	2,504,350	1,530,345	10,512	57,465
Total	***144,024,145***	***150,493,262***	***12,173,189***	***10,360,403***	***12,264,506***	***10,899,247***	***91,317***	***538,844***

Year = 2018								
AGI Class	**Number of Returns**		**Total Income (Millions)**		**Positive Income (Millions)**		**Negative Income (Millions)**	
	SCF	**SOI**	**SCF**	**SOI**	**SCF**	**SOI**	**SCF**	**SOI**
None	507,115	1,962,253	-11,002	-200,109	46,011	98,116	57,013	298,225
$1 to Under $25,000	38,914,043	50,453,810	476,879	647,707	479,536	690,574	2,658	42,867
$25,000 to Under $50,000	37,928,376	36,512,304	1,378,423	1,340,764	1,379,171	1,378,499	749	37,735
$50,000 to Under $100,000	37,577,727	35,146,085	2,698,081	2,534,215	2,711,783	2,586,907	13,702	52,693
$100,000 to Under $1,000,000	31,680,727	29,160,637	6,466,956	5,670,128	6,493,083	5,778,079	26,127	107,951
$1,000,000 or More	906,347	539,207	2,107,287	1,792,574	2,116,786	1,876,636	9,499	84,063
Total	***147,514,334***	***153,774,296***	***13,116,623***	***11,785,278***	***13,226,370***	***12,408,811***	***109,747***	***623,533***

The Brookings Economic Studies program analyzes current and emerging economic issues facing the United States and the world, focusing on ideas to achieve broad-based economic growth, a strong labor market, sound fiscal and monetary policy, and economic opportunity and social mobility. The research aims to increase understanding of how the economy works and what can be done to make it work better.

Questions about the research? Email communications@brookings.edu.
Be sure to include the title of this paper in your inquiry.