TAX POLICY CENTER
URBAN INSTITUTE & BROOKINGS INSTITUTION

SIMULATING INCOME TAX LIABILITIES IN THE SURVEY OF CONSUMER FINANCES

William Gale, Swati Joshi, Christopher Pulliam, and John Sabelhaus

February 2022

This policy brief summarizes a new paper (Gale et al. 2022a) in which we develop and refine 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; Bricker et al. 2016). Like most household surveys, however, the SCF does not ask detailed questions about household tax filing or tax liabilities.

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.

We proceed in several steps. The first step creates 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 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). Barring a direct linkage between SCF households and tax filings (prohibited by the IRS guidelines under which the SCF is conducted), there is no direct way to evaluate our approach. Table 1 shows a fairly tight relationship between the total number of simulated and actual filings by year. Additional data shown here suggests the model provides reliable estimates returns by individual filing status as well.

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

Source: Survey of Consumer Finances and the IRS Statistics of Income

The next step maps 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. Table 2 reports values of actual and simulated taxable income across the survey years.

TABLE 2

Estimated (SCF) and Actual (SOI) Taxable Income (Billions of Dollars)

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

Source: Survey of Consumer Finances and the IRS Statistics of Income

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. Table 3 shows differences in actual and imputed business income.

TABLE 3

Estimated (SCF) and Actual (SOI) Taxable Business Income (Billions of Dollars)

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

Source: Survey of Consumer Finances and the IRS Statistics of Income

The next step concerns 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. Adjusted gross income less deductions and (when they existed) personal exemptions generates taxable income.

TABLE 4

Tax Liabilities Before Credits (Billions of Dollars)

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

Source: Survey of Consumer Finances and the IRS Statistics of Income

TABLE 5

Tax Liabilities After Credits (Billions of Dollars)

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

Source: Survey of Consumer Finances and the IRS Statistics of Income

Finally, we generate baseline tax liability estimates, before and after credits, using the NBER TAXSIM model and benchmark those against published SOI values (Tables 4 and 5). 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.

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. 2022b) that builds on the methodology developed here. More generally, adding tax information to a data set like the SCF that already contains information on race, family structure, income, and wealth, and oversamples affluent households, could lead to a wide variety of new studies.

REFERENCES

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.

Gale, William G., Swati Joshi, Christopher Pulliam, and John Sabelhaus. 2022a. Simulating Income Tax Liabilities in the Survey of Consumer Finances. The Brookings Institution. (February).

Gale, William G., Swati Joshi, Christopher Pulliam, and John Sabelhaus. 2022b. "Taxing Business Incomes," The Brookings Institution. (February).

ACKNOWLEDGMENTS

Gale is Senior Fellow, Joshi is Research Assistant, Pulliam is Research Analyst, and Sabelhaus is Nonresident Senior Fellow, all at the Brookings Institution in Washington, DC. Email contact: jsabelhaus@gmail.com. We thank the Peter G. Peterson Foundation and Arnold Ventures for generous financial support and Janet Holtzblatt and Alice Henriques Volz for helpful comments. The analysis and conclusions set forth are those of the authors and do not necessarily reflect those of any other individual or organization.

The views expressed are those of the authors and should not be attributed to the Urban-Brookings Tax Policy Center, the Urban Institute, the Brookings Institution, their trustees, or their funders.