

Top 1 Percent Income Shares: Comparing Estimates Using Tax Data

By Gerald Auten and David Splinter

Online Appendix

This online appendix provides additional details of the changes needed to move from tax return based fiscal income definitions to expanded income definitions in order to compare top 1% income share estimates for Piketty and Saez (2003 and online updates, hereafter PS), the Congressional Budget Office (2018, hereafter CBO) and our approaches. In addition, the last section discusses Piketty, Saez, and Zucman (2019) criticisms of our analysis and differences between our results. Our two Auten and Splinter papers are both referred to below as AS because in almost all cases we follow the same methodology in both papers. When a specific paper or appendix is being referenced, they are distinguished by including the date (2018 or 2019).

I. Piketty-Saez Fiscal Income

The baseline population in PS is the estimated number of tax units in the U.S. Census resident population age 20 and over. The number of non-filer tax units is the difference between the number of filed tax returns and the estimated total. Tax units include all individuals on a tax return, or who would be expected to file together in the case of non-filers. Tax units can differ from Census households. For example, some households have multiple tax units.

The income definition (called market income by PS and fiscal income in Piketty, Saez, and Zucman (2018) and in our paper) equals adjusted gross income (AGI) less Social Security benefits and unemployment compensation included in AGI. To account for repeal of the capital gains exclusion in 1987 and non-filers, fiscal income includes capital gains excluded from AGI before 1987 and non-filer income. PS assumes that non-filer income is 20% of the average income of filers. AS (2018) presents tabulations of non-filer income using information returns, which suggest that 20% is a conservative but reasonable assumption for fiscal income. AS follows this assumption.

AS estimates of fiscal income reflect our replication of the methodology of PS and therefore can differ by small amounts. The primary example is that PS estimates after 2001 are based on IRS published summary data, while our estimates are based on the internal data files corrected by dropping duplicate returns, some of which have very high incomes. For example, in 1979 and 2014, the most recent PS estimates of top 1% shares are 9.96 and 21.52 and our replications are 9.95 and 21.82. The 2014 estimates differ by about 0.3 percentage point because PS uses Pareto interpolations of summary data, while we use microdata.

Limitations of the PS approach include:

- Fiscal income excludes important income sources (Social Security and other social insurance benefits and income accrued inside corporations) and the ratio of fiscal income to national income has declined over time to about 60% in recent years.
- The income reported on individual income tax returns has changed over time due to statutory changes in tax laws and changes in incentives to report income (especially the Tax Reform Act of 1986 and the Economic Recovery Tax Act of 1981).
- A number of tax unit issues tend to overstate top income shares. There are many more tax units than there are households and as much of 10 percent of filed returns are not in the baseline tax unit population (under age 20 and non-resident filers). Some married couples file separate tax returns, but are counted as only one tax unit in the baseline population. While most dependent

filers are under age 20, several million are full-time students deriving over half their support from their parents or guardian who claim a personal exemption for them. These dependent filers are treated in PS as if they were independent economic units with very low incomes.

II. CBO Income

CBO adds a number of income sources to fiscal income (including capital gains) to estimate pre-tax/pre-transfer income. For our CBO replication, we match these totals for each source. Table A1 shows the 2014 totals of these sources, as well as the share of each earned by the top 1%, as estimated by CBO. Note that these totals are less than the national income totals used by AS. This is because CBO focuses on federal taxes (excluding state and local taxes, such as sales and property taxes) and limits their sample to the civilian non-institutional population, which appears to reduce Social Security and Medicare benefits by about 10 percent in recent years.

Table A1 also shows that only 8.8 percent of these additional income sources are earned by the top 1%, significantly less than the 21.8 percent of fiscal income earned by the top 1%. Therefore, adding these more equally distributed sources tends to lower top income shares. For purposes of replicating CBO estimates, we generally follow AS allocation approaches when adding these additional income sources. Since the allocation of corporate tax differs significantly, we follow the CBO approach.

While both CBO and AS allocate 25 percent of the corporate tax by wages, AS allocates a share of the non-wage component (i.e., the capital component) of the corporate tax to the owners of retirement accounts. Between 1960 and 2015, this share increased from 4 to 50 percent of corporate ownership. CBO allocates the entire non-wage component by the more highly concentrated taxable capital income reported on individual tax returns.

To more closely approximate the household-based CBO approach, we need to make a number of adjustments to the tax unit-based sample. CBO creates synthetic households by statistically merging tax returns and non-filers based on CPS data. This accounts for households containing multiple tax units. For the sample adjustment step, we remove filers younger than 20 years old (and increase the number of non-filing tax units an equal amount), rather than also removing older dependent filers or non-resident filers, as in AS. These older dependent filers are primarily full-time students age 20 to 23. Census may count such students as separate households in some cases but not others. Such issues complicate distribution estimates and comparisons among studies. In addition, we reduce the number of non-filing adults by the institutional population (following the Census estimate of 4 million in 2014) to account for the CBO sample being limited to the civilian non-institutional population. The institutional population includes the prison population and individuals residing in retirement homes, of which there were far fewer in earlier decades so this adjustment is not made for 1979. The estimated CBO non-filer income may also differ from the Piketty and Saez (2003) assumption of 20% of average filer income, but we do not make an adjustment for this possibility.

For the purposes of ranking only, we account for top households having larger household sizes than tax unit sizes. We increase the number of individuals in top tax units, following the estimates of Larrimore, Mortenson, and Splinter (2017): 22 percent of top 1% households have one additional tax unit (i.e., the household consists of two or more tax units), 10 percent have two more tax units, and 6 percent have three more tax units. These adjustments are not made in 1979 because high-income tax units and households should have been much more similar in earlier decades, in part because of the lower prevalence of cohabitation. Note that this approach only provides an approximation of the CBO household-based approach for top 1% income shares and further adjustments would be needed to consider the rest of the income distribution.

III. Auten-Splinter Income

The Auten-Splinter income series shown in AS (2019) is the AS (2018) definition of pre-tax national income plus social insurance benefits (Social Security, including disability insurance, unemployment insurance, and Medicare). Note that AS (2018) pre-tax national income excludes transfers so as to be consistent with the conceptual basis of national income. As a result, the total of the modified income definition used in AS (2019) exceeds national income, but more closely matches the CBO definition of including both payroll taxes (including the employer share) and social insurance benefits. The addition of these benefits is important because instead of many retirees having little income, as in the AS (2018) pre-tax national income-based definition, their Social Security benefits are included.

Medicare benefits are based on allocating national income totals. A tenth of benefits is allocated to non-filers and remaining amounts to filers, with a share proportional to the number of adult individual tax filers aged 65 or older and with wages below \$150,000 (2015 dollars and indexed), assuming that if the primary filer is aged 65 or older then the secondary is also. The wage limitation is intended to account for the fact high wage-earners over age 65 are likely to be receiving health insurance through their employers. We subtract Medicare premium payments based on totals from Medicare Trustees Reports. We allocate Part D premiums and Part B premiums before 2007 proportionally to Medicare benefits. CBO makes similar adjustments in all years of their estimates. Since 2007, however, we add an additional step of allocating Part B premiums according to AGI-based progressive rates (and allocate only 7 percent, rather than 10 percent to non-filers).

AS incorporate adjustments that account for technical features of how income is reported on tax returns. See the AS (2018) online appendix for details. These adjustments include:

- Net operating loss carryover deductions are added back because these reflect prior year rather than current year income.
- Tax exempt interest and exempt combat pay are added to income.
- Itemized deductions for gambling losses are subtracted up to the amount of gambling winnings included in other income.
- State and local tax refunds included in AGI are subtracted because these are an adjustment for claiming too large a deduction in the prior year rather than income.
- Ordinary gains reported on Form 1040 are treated the same as Schedule D capital gains because tax reforms such as increasing the recapture of depreciation deductions changed previous capital gains into ordinary income. Tax reforms also tightened the rules for deducting losses.
- To account for the effects of the Tax Reform Act of 1986, post-TRA treatment of losses is imposed on income prior to 1987 by limiting business losses to post-TRA levels. Prior to 1987, excluded dividends are added to income.
- Dependent filers and other filers under age 20 are dropped from the sample. The income of dependent filers is allocated to returns claiming dependents. In addition to correcting for the large increase in the number of dependent filers due to the Tax Reform Act of 1986, this also better reflects the incomes of tax units with dependents.

IV. Responses to PSZ (2019)

Piketty, Saez, and Zucman (2019), hereafter PSZ (2019), makes four critiques of the AS approach. We summarize and respond to each these critiques, showing that their computations and claims suffer from several major problems.

A. Non-taxable/non-retirement capital income share to the top 1%

PSZ (2019) propose a “simplified” method for allocating non-taxable income and imply that in 2015 we only allocate about 10% of non-taxable/non-retirement capital income to the top 1%. This implication is incorrect. The PSZ (2019) analysis doesn’t take into account our adjustments to obtain an improved measure of the distribution of fiscal income in AS (2018). In particular, we define income groups based on the number of individuals (including dependents) and for ranking purposes adjust for tax unit size. This removes the upward bias and trend from the PSZ (2019) use of tax units by accounting for declining marriage rates outside the top of the income distribution. It also accounts for the effects of the increasing proportion of single-parent households. We drop filers under age 20 and non-resident filers, who are not included in the baseline estimate of the number of total tax units. Other dependent filers are also dropped because most are full-time students who are supported by their parents or guardians and thus not independent economic units.

As shown in Table 1 of AS (2019), these corrections reduce the top 1% income share by over 2 percentage points in 2014. In addition, we make the technical corrections to the definition of income listed in the prior section and allocate underreported income based on detailed IRS audits.

If these adjustments were accounted for, the PSZ (2019) computation of the share of non-taxable/non-retirement income that we allocate to the top 1% would be much larger than 10%. In fact, our top 1% receives about 40% of this income (excluding imputed rents from owner-occupied housing). This includes 40% of tax-exempt interest and 30% of business property taxes. These are easily estimated from the AS (2018) online spreadsheet tab C1. For private non-retirement portions, they receive about 40% of corporate income taxes and 50% of retained earnings.¹ Thus, the PSZ simplified method mischaracterizes our actual allocations of non-taxable capital income. Specific limitations of the PSZ simplified method are discussed in the next section.

B. Amount of income earned by the top 1% and limitations of the PSZ simplified method

PSZ (2019, p. 6) claim that the AS (2018) top 1% estimate of pre-tax national income is less than the PS estimate of fiscal income in 2015. In fact, our top 1% receives \$2.23 trillion of pre-tax national income which is considerably more than the \$1.78 trillion of fiscal income excluding capital gains (see AS (2018) online spreadsheet tab C1).

AS (2018) Table 1 and online spreadsheet show the step-by-step effects of adjustments to arrive at our estimates. Here, we summarize that detailed approach to provide a simple analysis for 2015.

First, converting from using tax units to individuals to set income groups, ranking by size-adjusted incomes, and sample corrections (such as removing non-resident filers) reduces the top 1% income share by about 2.3 percentage points.

Second, non-taxable income sources are roughly allocated to the top 1% as follows: 2% of employer-sponsored insurance, 4% of employer-paid payroll taxes, 5% of other untaxed income (underreported income, sales taxes, non-taxable income from non-profit/government corporate ownership, and effects of re-ranking and our income corrections; ignoring income corrections this would be about 7%), about 10% of non-taxable retirement account income (including both defined benefit and defined contribution), 9% of imputed rent (including property taxes), and about 40%

¹ The AS (2018) online spreadsheet only reports distributions of corporate taxes and retained earnings that are aggregated with more equally distributed retirement and non-profit/government portions. For example, total corporate retained earnings distributions result from allocating about half to retirement accounts (top 1% receives about 15%), a tenth to non-profits/government (top 1% is allocated about 7%), and the rest by private capital income (top 1% earns about 50% of weighted dividends and capital gains). In summary: $0.50 \cdot 15\% + 0.10 \cdot 7\% + 0.40 \cdot 50\% = 28\%$ (as seen in the AS (2018) online spreadsheet tab C6, cell AK61).

of private non-retirement capital income (tax-exempt interest, business property taxes, non-taxable fiduciary income, portions of corporate income taxes and retained earnings).²

Hence, starting from fiscal income excluding capital gains (of which the top 1% earns 18.6%), top 1% AS pre-tax national income can be estimated as follows:

$$(18.6\% - 2.3\%) \cdot \$9.54 + 2\% \cdot \$0.8 + 4\% \cdot \$0.5 + 5\% \cdot \$2.1 + 10\% \cdot \$0.9 + 9\% \cdot \$0.7 + 40\% \cdot \$1.2 = \$2.23 \text{ trillion}$$

Fiscal indivs. ESI payroll other retirement imp. rent private capital

One can change the equation above in three ways to approximate the PSZ simplified method and PSZ estimate: remove the individual/size-adjustment/sample correction effect, increase the private capital allocation to 53% and all other allocations to 16%,³ and treat imputed rental income as private capital income.

These changes reveal three limitations of the PSZ simplified method: (1) going back to using tax units to set income groups rather than individuals or adults retains the upward bias from failing to account for lower marriage rates outside the top of the income distribution, (2) allocating 16% of employer-sponsored insurance and payroll taxes (or alternatively Social Security benefits) to the top 1% rather than the more appropriate and equal shares of 2% and 4%, (3) treating owner-occupied imputed rent like other concentrated capital income, hence allocating 53% of imputed rent to the top 1% rather than the more appropriate and equal share of about 9%. Thus, it is inappropriate to use their simplified method to make comparisons with our results.

C. Size-adjusted incomes for ranking purposes

Following the CBO (2018) approach, we rank tax returns by size-adjusted incomes. The reason to adjust for tax unit (or household) size is to determine which tax units belong in each relative income centile after accounting for differences in the number of individuals in each tax unit. Once that unit is assigned into a centile, the full amount of income is counted in that centile such that incomes sum to national income.

PSZ (2019, p. 3) state that size-adjusting for ranking purposes “mechanically biases downward the top 1 percent income share.” In fact, our size-adjustment for ranking purposes actually *increases* top 1% income shares. For example, in 2015 it increases the top 1% income share by 0.9 percentage point (see AS (2019) online spreadsheet). This occurs because size-adjusting for ranking purposes pushes tax units with many dependents out of the top 1%, and as these tax units have many individuals this allows more tax units with fewer individuals per tax unit to take their place. Hence, the number of tax units and amount of income in the top 1% increase results in a higher top 1% income share. Note that we estimate this ranking effect after grouping by the number of individuals. Essentially, size-adjusting for ranking purposes offsets part of the decrease in top 1% income shares from setting income groups by individuals rather than tax units.

Size-adjusting incomes accounts for sharing and economies of scale. Our approach and that of CBO is to divide tax unit income by the square-root of the number of individuals in the tax units. This size-adjustment is similar to that implied by Census Bureau poverty thresholds and is a standard approach in the inequality literature dealing with households and families (Gottschalk and Smeeding, 1997; Atkinson and Brandolini, 2001). The OECD and U.S. Census both use similar equivalence scales (d’Ercole and Förster, 2012; Fontenot, Semega, and Kollar, 2018).

² These shares are generally calculated from AS (2018) online spreadsheet tab C1 by dividing the change in top 1% income by the change in total income from adding each non-taxable income source. These additions can be sensitive to the order in which different sources are added and can result in some re-ranking of tax units. Re-ranking is generally limited, except for underreported income (discussed below).

³ These values come from the PSZ (2019) online spreadsheet: <https://eml.berkeley.edu/~saez/PSZ2019datafile.xlsx>

In the description of goals for national income distributions, Alvaredo et al. (2016, p. 10) state that the equal-split between spouses approach distributes income “equally between its members.” But the equal-split approach, used by PSZ (2018), only accounts for adults and ignores effects from other members of tax units or households—including children, student dependents, elderly dependents, and possibly extended family members or cohabiting partners. Thus, while the equal-split approach accounts for effects from declining marriage rates, it does not account for changes in household structure, such as the increase in single-parent households.

D. Allocating underreported income

PSZ (2019, p. 7) state that “...Auten and Splinter (2018)...attribute a disproportionately large and growing fraction of under-reported income to the bottom 90 percent.” In the main paper and AS (2018) we point out that PSZ (2018) allocate underreported business income only by positive reported business income (none to business losses). In comparison, we allocate underreported income based on detailed IRS audit studies. Relative to what these IRS studies show, the PSZ approach allocates too much to the top income groups.

In 2001, the year of the most recently published distributional estimates from an audit study, our allocation of the filer portion of underreported income has a negligible effect on top income shares (an increase of less than two-hundredths of a percentage point). This is consistent with the Table 4 estimates of Johns and Slemrod (2010), which show that top 1% income shares are unchanged when moving from underreporting-exclusive (reported) incomes to underreporting-inclusive (true) incomes. The Johns and Slemrod (2010) distribution tables only include filers, but their Table 1 shows that an additional portion of the tax gap is due to non-filers. In 2001, our allocation of 15% of underreported income to non-filers results in a small decrease in top income shares from allocating underreported income: only 0.15 percentage point.

In order to capture the effect of changing patterns of underreporting across the income distribution over time, we perform calculations similar to Johns and Slemrod (2010) using earlier random audit data. These estimates show higher rates of underreporting among those with negative adjusted gross incomes. This is likely due to the effects of tax shelters prior to and immediately following the base-broadening reforms of the Tax Reform of 1986. In these years, allocating a higher share of underreported income to returns with negative incomes results in a larger increase top income shares; this added income re-ranks more individuals from the bottom of the distribution by reported income to the top of the distribution by true income. Hence, audit data suggest that reforms limiting tax shelters resulted in a growing fraction of underreported income for the bottom 90 percent of the distribution by reported income.

References

Alvaredo, Facundo, Tony Atkinson, Lucas Chancel, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman. 2016. “Distributional National Accounts (DINA) Guidelines: Concepts and Methods used in WID.world.” WID.world Working Paper 2016/1, December.

Atkinson, Anthony B., and Brandolini, Andrea. 2001. “Promises and Pitfalls in the Use of Secondary Data Sets: Income Inequality in OECD Countries as a Case Study.” *Journal of Economic Literature* 39(3): 771–799.

Auten, Gerald, and David Splinter. 2018. “Income Inequality in the United States: Using Tax Data to Measure Long-Term Trends.” Working paper.

Auten, Gerald, and David Splinter. 2018. “Online spreadsheet.” Posted at davidsplinter.com along with associated computer code.

Auten, Gerald, and David Splinter. 2019. “Online spreadsheet.” posted at davidsplinter.com along with associated computer code.

Congressional Budget Office. 2018. “The Distribution of Household Income, 2015.” (supplemental data) Congressional Budget Office.

Fontenot, Kayla, Jessica Semega, and Melissa Kollar. 2018. “Income and Poverty in the United States: 2017.” Current Population Reports, P60-263. Washington: U.S. Census Bureau.

d’Ercole, Marco M., and Förster, Michael. 2012. “The OECD Approach to Measuring Income Distribution and Poverty: Strengths, Limits and Statistical Issues.” In *European Measures of Income and Poverty: Lessons for the U.S.*, D.J. Besharov and K.A. Couch, eds. Oxford University Press, New York, NY.

Johns, Andres, and Joel Slemrod. 2010. “The Distribution of Income Tax Noncompliance.” *National Tax Journal* 63(3): 397–418.

Gottschalk, Peter, and Smeeding, Timothy M. 1997. “Cross-National Comparisons of Earnings and Income Inequality.” *Journal of Economic Literature* 35(2), 633–687.

Larrimore, Jeff, Jacob Mortenson, and David Splinter. 2017. “Household Incomes in Tax Data: Using Addresses to Move from Tax Unit to Household Income Distributions.” Finance and Economics Discussion Series 2017-002. Washington: Board of Governors of the Federal Reserve.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman. 2018. “Distributional National Accounts: Methods and Estimates for the United States.” *The Quarterly Journal of Economics* 131(2): 519–578.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman. 2019. “Simplified Distributional National Accounts.” WID working paper 2019/1 (forthcoming *AEA Papers and Proceedings*).

Table A1: CBO additional income totals, 2014 (millions of dollars)

	NIPA	CBO		
	Totals	Total	Top 1%	Top 1% Share
Employee cont. to deferred comp. plans	---	261,215	14,305	5.5
Contributions to health insurance	754,366	435,358	9,080	2.1
Employer payroll taxes	547,972	460,236	17,539	3.8
Corporate taxes	408,416	335,848	161,953	48.2
Tax-exempt interest	62,103	62,194	33,087	53.2
Social Security benefits	846,595	771,206	7,588	1.0
Medicare benefits	600,907	447,797	3,980	0.9
Unemployment insurance benefits	35,449	37,316	124	0.3
Workers' compensation	---	12,439	124	1.0
Total		2,823,608	247,781	8.8

Notes: NIPA totals only show sources included in the AS (2019) income definition. Underlying CBO values are rounded to the nearest hundred dollar per household amount.

Sources: U.S. Bureau of Economic Analysis, Congressional Budget Office (2018) supplemental data, and authors’ calculations.