

# Presence and Persistence of Poverty in U.S. Tax Data

## Appendix

Jeff Larrimore  
Federal Reserve Board

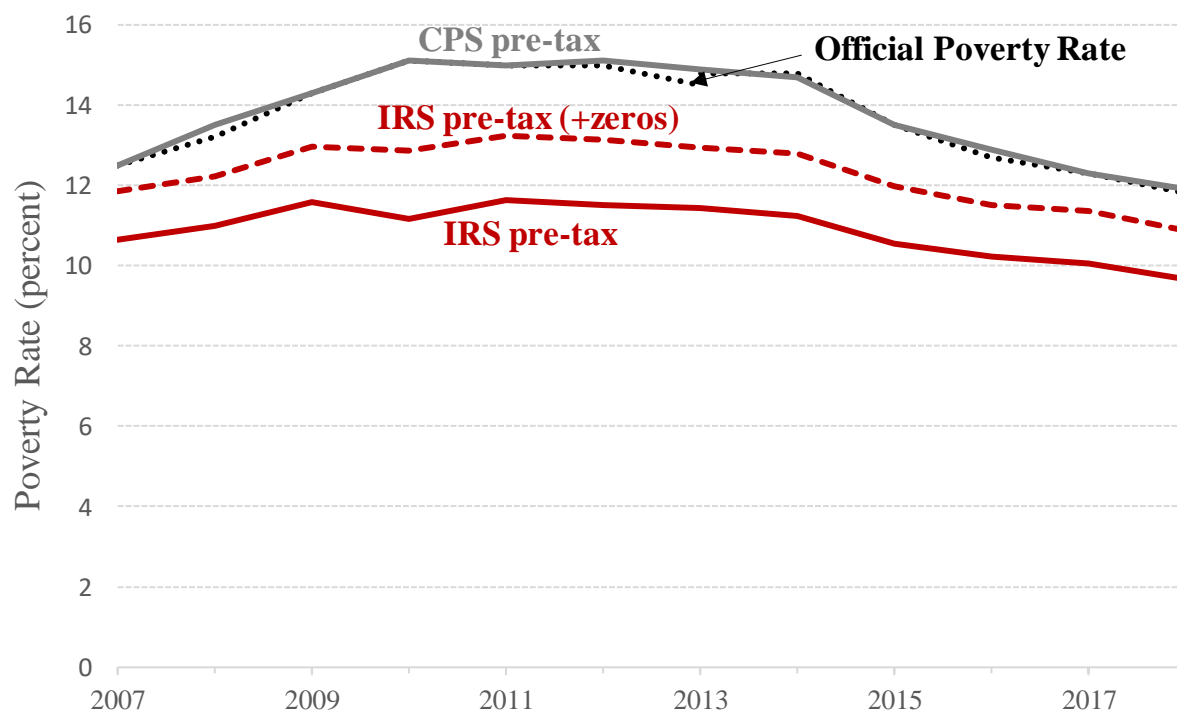
Jacob Mortenson  
Joint Committee on Taxation

David Splinter  
Joint Committee on Taxation

Main paper: Larrimore, Jeff, David Splinter, and Jacob Mortenson. 2021. “[Presence and Persistence of Poverty in U.S. Tax Data](#).” In Raj Chetty, John N. Friedman, Janet C. Gornick, Barry Johnson, and Arthur Kennickell (eds.), *Measuring and Understanding the Distribution and Intra/Inter-Generational Mobility of Income and Wealth*. Cambridge, MA: NBER.

## Appendix Figures

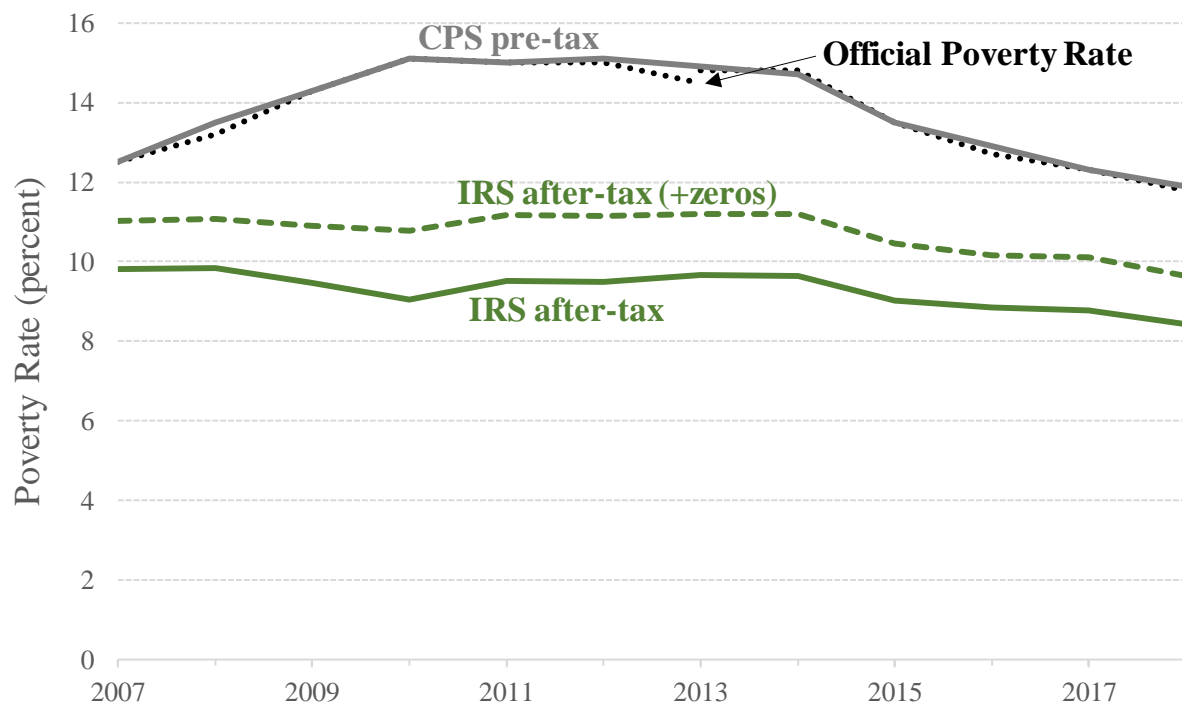
**Appendix Figure A1. Pre-tax poverty rates without anchoring IRS data, with and without imputed individuals living in zero-income households.**



*Sources:* Semega et al. (2019), authors' calculations using IPUMS CPS (Flood et al. 2018) and the IRS Tax Household Data.

*Notes:* The break in the official poverty rate reflects the change in the CPS questionnaire with a split sample partially receiving the original questions, as described in Semega et al. (2019). All series except the Official Poverty Measure use a poverty threshold for a 4-person household of \$20,994.

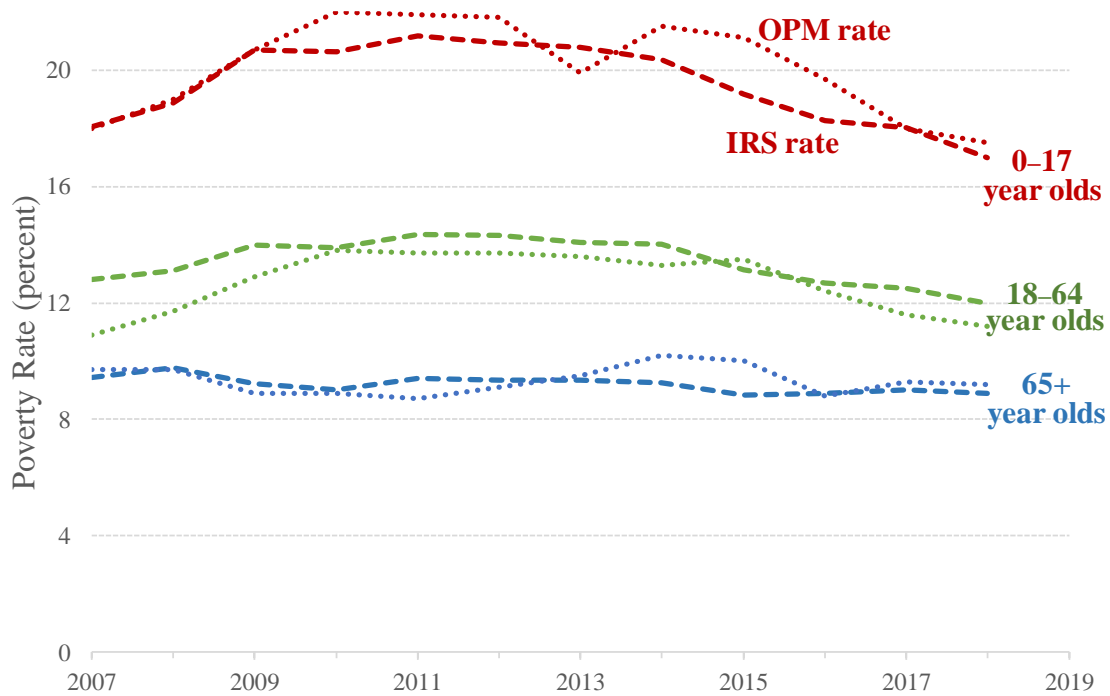
**Appendix Figure A2. After-tax poverty rates without anchoring IRS data, with and without imputed individuals living in zero-income households.**



*Sources:* Semega et al. (2019), authors' calculations using IPUMS CPS (Flood et al. 2018) and the IRS Tax Household Data.

*Notes:* The break in the official poverty rate reflects the change in the CPS questionnaire with a split sample partially receiving the original questions, as described in Semega et al. (2019). All series except the Official Poverty Measure use a poverty threshold for a 4-person household of \$20,994.

**Appendix Figure A3: Pre-tax poverty rates by age in IRS data compared to Official Poverty Rates**



*Sources:* Semega et al. (2019) and authors' calculations using the IRS Tax Household Data.

*Notes:* Poverty in the IRS data is based on pre-tax income using 2007 anchored poverty thresholds.

## **Data Appendix**

### *A.I. Sample Construction*

This appendix describes the sample construction and income definitions. The data used in this paper is based on the Tax Household Data, whose target population is the annual resident population of the U.S. Currently, these data are available from 2007 to 2018. Because these are population-level files (not samples), individuals can be linked over time (even if they move between households). Individuals not in the data in all years might enter the data because they are born, immigrate to the United States, or have a tax record for the first time. Similarly, individuals might exit the panel due to death, emigration, or not having a tax record in a given year. Birth (death) and immigration (emigration) are much more likely causes for entrance (exit) than a first-time tax record. For example, Cilke (2014) and others find that roughly 99 percent of all individuals appear on a tax filing in a given year.

However, we make different adjustments based on our estimate of the reason an individual entered (exited). If the cause was birth (death), we do not adjust, and the panel is naturally unbalanced to remain representative of the U.S. population in a given year (which also has births and deaths). However, we make a special adjustment for individuals entering (exiting) the United States. To identify individuals missing in a given year because they are non-residents, we append non-resident individuals that are observed in the tax data, including excluded dependents with Individual Taxpayer Identification Numbers (ITINs). Note that all non-residents are removed for the analysis in this paper. For each year, we then extract an individual-level sample based on the last four digits of Taxpayer Identification Numbers (TIN), mostly Social Security Numbers. This is a standard approach for selecting a random sample of individuals in the tax data. Our sample includes 100 of 9,999 of these TIN endings, as none end in all zeros, resulting in a 1 percent sample that includes about 3 million individuals each year. These samples include household level income sources and some individual and household level characteristics.

Pre-tax income is mostly defined as in Larrimore, Mortenson, and Splinter (2019):

For annual tax returns, this definition starts with the total income from line 22 of IRS Form 1040—which includes income from wages, salaries, taxable interest, dividends, alimony, business income, rents and royalties, taxable Social Security, taxable private retirement income, and unemployment compensation. Five adjustments are made to this income from the Form 1040: (1) non-taxable interest reported on Form 1040 is added, (2) realized capital gains (from Schedule D) are removed, (3) taxable Social Security benefits are replaced by total Social Security benefits reported on Form SSA-1099, (4) taxable private retirement income is replaced with gross private retirement income, which reflects retirement savings distributions less rollovers from Forms 5498 and 1099-R, and (5) incomes are bottom-coded at zero to limit the effect of business losses. This income measure is broader than the tax return income definition used by Piketty and Saez (2003), since it includes Social Security income and unemployment compensation. It also comes as close as possible to the pretax income measure from the CPS and used by the Census Bureau for their official income statistics. The primary difference between our income measure and the income measures used by the Census Bureau for their official income statistics is that we are not able to observe nontaxable cash transfer income such as public assistance and supplemental security income...these non-

taxable cash transfers represent approximately 2.5 percent of income reported by the Census Bureau...

For non-filers, pre-tax cash income is calculated as the sum of income reported on information returns that would be included in the income definition for filers were they to file a tax return. Following Mortenson et al. (2009), who also derive income for non-filers based on information returns, we include income from wages and salaries reported on Form W-2, unemployment benefits from Form 1099-G, Social Security and disability benefits from Form SSA-1099, interest income from Form 1099-INT, dividends from Form 1099-DIV, gross private retirement income as retirement savings distributions less rollovers from Forms 5498 and 1099-R, and self-employment income from Form 1099-MISC...However, for Form 1099-MISC we offset reported income by 70 percent to reflect that gross income from self-employment activities appear on the 1099-MISC information returns but the associated business expenses do not. This offset is necessary to convert gross self-employment income to net self-employment income.

We broaden the income sources for non-filers, adding income from Forms 1099-K (reduced by 70 percent to convert to net income and excluded if net greater than \$10,000) and Schedule K-1 for Forms 1065 and 1120S to capture partnership and S-corporation shares of income, as well as accounting for unemployment income excluded from 2009 tax returns.

After-tax income deducts federal individual income tax and deducts federal payroll taxes. Federal refundable credits are added, including the EITC, the Additional Child Tax Credit (ACTC), and small amounts of other refundable credits (excluding 2008 recovery payments). Income taxes for filers are the total tax amount (line 60 on the 2010 Form 1040) less self-employment taxes and for non-filers the amounts withheld on Forms W-2 and 1099-MISC. Federal payroll taxes include Social Security and Medicare taxes withheld on Form W-2 and self-employment taxes less excess Social Security taxes withheld as reported on Form 1040.

## References

- Cilke, James. 2014. "The Case of the Missing Strangers: What we Know and Don't Know about Non-Filers." *Proceedings of the 107th Annual Conference of the National Tax Association*.
- Larrimore, Jeff, Jacob Mortenson, and David Splinter. 2019. "Household Incomes in Tax data: Using Addresses to Move from Tax Unit to Household Income Distributions." *Journal of Human Resources*. Advance online publication. doi: [10.3368/jhr.56.2.0718-9647R1](https://doi.org/10.3368/jhr.56.2.0718-9647R1)
- Semega, Jessica, Melissa Kollar, John Creamer, and Abinash Mohanty. 2019. "Income and Poverty in the United States: 2018." Current Population Reports. United States Census Bureau.