I. program-sas-auten-splinter-AEA-PP-2019.txt

This is a SAS program saved in txt format to make more accessible. It is used to generate estimates, which are then entered in the Excel file in tabs Decomp1979, Decomp2014, TRA86, and AS.

This file uses confidential tax data. The program could be adapted to use public use tax files. For more details, see below.

Contact David Splinter at david.splinter@jct.gov for questions.

To run:
1) Name folders after libname based on where your tax microdata files are located
2) Update tax file names to your names. They are currently puindiv&yr (1960-1978 public use files) and indiv.soi&yr (1979-2015 confidential files).
3) Uncomment macro statements at bottom of program to estimate income shares with various step-by-step changes from fiscal income (each calls MultiTopShares).
4) Output goes to the temporary output folder. The main output is TOTAL_INCxxx, where xxx is the income type is added to the end of the output name. Values are in nominal terms.

Income Groups are as follows:

1 = P0-90
2 = P90-95
3 = P95-99
4 = P99-99.5
5 = P99.5-99.9
6 = P99.9-99.99
7 = Top 0.01%
8 = ALL INCOME

II. AutenSplinter-TopIncomes-AEA-PP-2019.xlsx
This is an Excel file containing results and tabulations of variables used in the analysis.

Data Background: Information, Accessibility, and Persistence:

The SAS program is based on tax return microdata files for the U.S. between 1960 and 2015. We use public use data from 1960 to 1978 (there are no microdata files in 1961, 1963, or 1965, and confidential data between 1979 and 2015, also known as INSOLE files. We also use the DM-1 file, which links taxpayer identification numbers (usually SSNs) to years of birth. To access these data, a researcher would generally need to apply to the IRS Statistics of Income Joint Statistical Research Program, see www.irs.gov/uac/soi-tax-stats-joint-statistical-research-program.

The public use versions of the data are available to certain researchers, and should give similar results as the confidential data we use since 1979. Public use files do not have taxpayer identification numbers (usually SSNs) and therefore age cannot be determined (using the confidential DM-1 file). The files do include an indicator of dependent status. Most dependent filers are under age 20 and most of the rest are full-time students under age 24. For most purposes, this is a good proxy for our age restriction. In

addition, age 65 and over can often be determined using the presence of the extra standard exemption for those over age 65 and, for years before 1987, age 65 exemptions. The state variable is not available for returns with AGI over $200,000 or a high sampling rate. This means that non-residents with high incomes cannot be removed from the sample, but the number of such returns is relatively small and should not significantly affect results. If running the code with public use files, set state=1 when this variable is missing, and yob_dm1= &yr – 40 for non-dependent returns (so that these returns are not removed based on the residency or age restrictions).

Using the public use data or other versions of the confidential data requires recoding the JCT-specific variables in some years. This can be done using variable names from files from NBER: https://users.nber.org/~taxsim/gdb/ or http://users.nber.org/~taxsim/concordance79_96.pdf or http://users.nber.org/~taxsim/88-2004.pdf.

Totals for each variable of interest in all years are shown in the Auten and Splinter (2018) online data in tab 'TOTALS', which can be used to validate any updated variables, see http://davidsplinter.com/AutenSplinter-IncomeIneq.xlsx This program includes values from this spreadsheet (see tab T-T1, as well as other T tabs). These data are from the Bureau of Economic Analysis (BEA) National Income and Product Accounts (NIPA) and are used for sources of income not reported on tax returns, as well as control totals for income items only partially reported on tax returns. These data were taken from the NIPA website (bea.gov and under tools go to interactive data, currently the link is https://apps.bea.gov/iTable/index_nipa.cfm). Note that these values are regularly updated by the BEA and so can change slightly but should be permanently available.