

# Distributional National Accounts: Methods and Estimates for the United States Data Appendix\*

Thomas PIKETTY (Paris School of Economics)  
Emmanuel SAEZ (UC Berkeley and NBER)  
Gabriel ZUCMAN (UC Berkeley and NBER)

November 10, 2017

## **Abstract**

This Data Appendix supplements our paper “Distributional National Accounts: Methods and Estimates for the United States.” It provides complete details on the methodology, data, and programs.

---

\*Thomas Piketty: [piketty@psemail.eu](mailto:piketty@psemail.eu); Emmanuel Saez: [saez@econ.berkeley.edu](mailto:saez@econ.berkeley.edu); Gabriel Zucman: [zucman@berkeley.edu](mailto:zucman@berkeley.edu). We thank Tony Atkinson, Arthur Kennickell, Jean-Laurent Rosenthal, and numerous seminar and conference participants for helpful discussions and comments. Kaveh Danesh and Juliana Londono-Velez provided outstanding research assistance. We acknowledge financial support from the Center for Equitable Growth at UC Berkeley, the Institute for New Economic Thinking, the Laura and John Arnold foundation, NSF grant SES-1559014, and the Sandler foundation.

# Contents

<b>A</b>	<b>Organization of data files and computer code</b>	<b>3</b>
A.1	Raw data: macroeconomic aggregates, tax data, survey data . . . . .	3
A.1.1	Macroeconomic aggregates . . . . .	3
A.1.2	Tax data . . . . .	5
A.1.3	Survey data . . . . .	7
A.2	Programs creating US distributional national accounts . . . . .	7
A.2.1	Programs using survey data . . . . .	7
A.2.2	Programs building tax data . . . . .	8
A.2.3	Programs enriching public-use tax data . . . . .	9
A.2.4	Programs creating DINA micro-files . . . . .	9
A.2.5	Programs producing statistics and exporting . . . . .	10
A.3	Micro-files of U.S. distributional national accounts . . . . .	11
A.3.1	External-use DINA files . . . . .	11
A.3.2	Codebook . . . . .	11
A.3.3	Internal-use DINA files . . . . .	18
A.4	Results: Distributional summary statistics . . . . .	18
A.4.1	Main distributional results . . . . .	18
A.4.2	Detailed distributional results . . . . .	18
A.5	List of online files . . . . .	19
<b>B</b>	<b>Supplementary information on imputations methods</b>	<b>19</b>
B.1	Imputation of non-filers . . . . .	19
B.2	Imputation of income split among spouses . . . . .	21
B.3	Imputations of wealth and income from survey data . . . . .	22
B.3.1	General imputation method . . . . .	22
B.3.2	Wealth imputations . . . . .	23
B.3.3	Income imputations . . . . .	23
B.4	Imputation of taxes . . . . .	24
B.4.1	Individual income and payroll taxes . . . . .	24
B.4.2	Corporate income tax . . . . .	25
B.4.3	Property taxes . . . . .	26
B.4.4	Sales and excise taxes . . . . .	27

B.5	Collective consumption expenditure . . . . .	27
B.5.1	Health spending . . . . .	27
B.5.2	Education spending . . . . .	27
B.6	Imputations to match national income . . . . .	28
B.6.1	Income of non-profits . . . . .	28
B.6.2	Government deficit and interest payments on debt . . . . .	28
B.6.3	Surplus of pension system . . . . .	29
B.7	Other . . . . .	31
B.7.1	Unit of observation . . . . .	31
B.7.2	Price deflator . . . . .	32
<b>C</b>	<b>Comparison with Other Studies</b>	<b>33</b>
C.1	Comparison with Auten and Splinter (2017) . . . . .	33
C.2	Comparison with Smith et al. (2017) . . . . .	37

This data appendix supplements our paper “Distributional National Accounts: Methods and Estimates for the United States.” The appendix includes a large number of data files and computer codes, mostly in Excel and Stata formats. In section A, we describe the organization of these data files. In section B, we provide a number of supplementary methodological details on the imputations we make to create our U.S. distributional national accounts.

## A Organization of data files and computer code

Our data files and computer codes are organized into four parts. First are a number of raw data sources, which form the starting point of our project: macroeconomic aggregates from the national accounts, tax return micro-data, and survey micro-data. Second are the [Stata programs](#) that compute the distribution of U.S. national income starting from tax data, supplementing tax data using surveys, and making explicit assumptions for the categories of income which are not covered by tax or survey data. Third are our [Distributional National Accounts micro-files](#), i.e., micro-files of synthetic observations representative of the U.S. population containing the national accounts income and wealth variables. Fourth are a number of Excel files with summary distributional results which were produced using the synthetic files. We describe each of these four components in turn and conclude this section by providing a list of all the files used in this research.

### A.1 Raw data: macroeconomic aggregates, tax data, survey data

#### A.1.1 Macroeconomic aggregates

The first key data source used in this research is the national income and wealth accounts of the United States. All the macroeconomic series used in this research are gathered and presented in the Excel file [Appendix Tables I \(Macro\)](#). This file is built as follows.

First, [Appendix Tables I \(Macro\)](#) collects all the raw national income and wealth accounts published by U.S. statistical agencies. The National Income and Product Accounts (NIPA) of the United States, published by the Bureau of Economic Analysis (US Department of Commerce, Bureau of Economic Analysis, 2016), are collected in the sheet “nipa\_raw”. The Financial Accounts of the United States, published by the Federal Reserve Board, are collected in the sheet “ima\_raw”.<sup>1</sup> Both the NIPAs and Financial Accounts are frequently updated; we regularly

---

<sup>1</sup>The Financial Accounts of the United States, formerly known as the Flow of Funds, include data on the flow of funds and levels of financial assets and liabilities by sector and financial instrument; full balance sheets, including net worth, for households and nonprofit organizations, nonfinancial corporate businesses, and nonfinancial noncorporate businesses; the Integrated Macroeconomic Accounts (an attempt at bringing together the

fetch updated series to reflect the latest data revisions and ensure full consistency with the current macroeconomic aggregates of the United States. The national accounts series currently included in [Appendix Tables I \(Macro\)](#) were downloaded in October 2016 from the websites of the [Bureau of Economic Analysis](#) and the [Federal Reserve Board](#).<sup>2</sup>

Second, we have extended the U.S. income and wealth accounts backwards to 1913. The NIPAs start in 1929, while official Financial Accounts start in 1945. However, high-quality, well-documented historical accounts exist before the official series start. For income, we rely on the national income accounts of Kuznets (1941) for 1919–1929 and King (1930) for 1913–1919. For wealth, we combine balance sheets from Goldsmith, Brady, and Mendershausen (1956), Wolff (1989), and Kopczuk and Saez (2004) that are based on the same concepts and methods as the Financial Accounts.<sup>3</sup> We adjust the historical series so as to ensure continuity with the official statistics; all the adjustments are carefully documented in the sheets “DataIncome” and “DataWealth” of [Appendix Tables I \(Macro\)](#).

Third, we construct national income and wealth categories that are consistent with the 2008 System of National Accounts, the international standard for national accounting (United Nations, 2009). The U.S. macroeconomic statistics are generally broadly consistent with the SNA but differ in a number of ways. For instance, sectoral decompositions differ (U.S. accounts isolate a non-financial non-corporate business sector that does not exist in the SNA<sup>4</sup>); the NIPAs use income concepts that do not exist in the SNA (such as personal income and net interest); and on the contrary the SNA includes decompositions of income that do not exist in the NIPAs (such as the decomposition of national income by sector<sup>5</sup>). McCulla, Moses and Moulton (2015) provide a detailed comparison and reconciliation of the National Income and Product Accounts with the 2008 System of National Accounts. Because our ultimate goal is to be able to provide cross-country comparisons of inequality using similar methods and concepts, we have re-organized the U.S. national accounts so as to make them consistent with the SNA. This re-organization does not affect the level or growth of national income and wealth, but makes comparing the components of national income and wealth across countries more straightforward.

---

NIPAs and the Financial Accounts in an accounting framework founded on the System of National Accounts); and additional supplemental details. They are published quarterly in US Board of Governors of the Federal Reserve System (2016).

<sup>2</sup>Downloading the raw national accounts series and integrating them into [Appendix Tables I \(Macro\)](#) is done by the program `scrap_macro.do`.

<sup>3</sup>The same historical income and wealth series were used by Saez and Zucman (2016).

<sup>4</sup>this sector includes non-corporate businesses such as large partnerships that would be included in the corporate sector in the SNA. As a result the share of the corporate sector in total value-added is low in the United States compared to other countries; see, e.g., Table I-A2 in [Appendix Tables I \(Macro\)](#).

<sup>5</sup>See Table I-A3 in [Appendix Tables I \(Macro\)](#).

The harmonized income and wealth categories we use—founded on the SNA—are described in Alvaredo et al. (2016).

The file [Appendix Tables I \(Macro\)](#) is organized as follows. Tables I-A1 to I-A10 provide series of U.S. national income and its components covering the 1913-2015 period. Tables I-B1 to I-B7 provides similar series for wealth, and Tables I-D1 to I-D8 for saving & investment. Tables I-S.A1 to I-S.A13, I-S.B1 to I-S.B3, and I-S.D1 to I-S.D4 provide supplementary decompositions of income, wealth, and saving & investment respectively.<sup>6</sup>

### A.1.2 Tax data

The second key data source used in this research is micro tax data. Here we describe the raw tax data we used.

**Public-use microfiles.** Since 1962, there exists annual, high-quality public-use files (PUF) with samples of tax returns that have been created by the Statistics of Income (SOI) division of the Internal Revenue Service (IRS). The complete set of files and their documentation are maintained by Daniel Feenberg at the [NBER](#).<sup>7</sup> These files provide information for a large sample of taxpayers, with detailed income categories. We have made these files comparable over time by constructing consistent income categories; this is primarily done by the program `build_small.do` (see below). The most recent file available is for 2010. SOI is currently working on redesigning the PUFs. Once this redesign is completed, SOI plans to release the PUFs more timely and reduce the time lag (as had been the case in the past).

**SOI individual income tax files.** The PUF files are actually a subset of records and variables coming out of the SOI individual tax return files that are created annually. These internal files are used for publishing official SOI statistics on individual incomes (US Department of Treasury, Internal Revenue Service, annual). This *Statistics of Income: Individual Income Tax Returns* provides detail on the sampling. The files are also used inside government tax agencies (US Treasury, Joint Committee on Taxation, and Congressional Budget Office) as a base for evaluating and scoring tax policies. The annual files are maintained since 1979.<sup>8</sup> The most recent file available is for 2014. Typically, the file for year  $t$  becomes available around July of

---

<sup>6</sup>In addition to the national account aggregates, we present totals for the flow of income reported to the IRS (“fiscal income”) in Tables I-C1 to I-C6 (and I-S.C1 to I-S.C4) of [Appendix Tables I \(Macro\)](#).

<sup>7</sup>There are no files for 1963 and 1965. A file for 1960 exists but with much fewer variables.

<sup>8</sup>SOI files for 1960-1978 (except 1961, 1963, 1965) had also existed but have been lost. Only the public use version remains.

year  $t + 2$ .

The files contain a much larger number of variables than the PUF. In particular, they include basic demographic information: age, gender, date of birth (and date of death). They also include a more detailed set of variables. In particular, they provide the breakdown of self-employment income across spouses. Starting in 1999, they can also be merged internally to the population-wide file of information tax returns to obtain the breakdown of wage earnings across spouses for married joint filers. Saez (2016) uses these internal files to create a set of basic tabulations of tax filers by age, gender, and earnings splits that allows to create synthetic versions of these demographic variables in the PUFs. Saez (2016) shows that the calculations using internal data vs. using the PUFs enhanced with these synthetic variables delivers very close results. The results published in this paper are created using the superior internal data whenever possible. In order to have the same set of programs working both externally on the PUF and internally on the SOI files, we have extracted and renamed the variables in the SOI files using the variables names from the PUFs maintained at NBER and keeping a few extra variables such as gender, age, and earnings split within married couples that are needed in some of our series (these variables are added synthetically in the external PUFs in our programs).

**Pre-'62 tax data.** For the pre-1962 period, no micro-files are available so we rely instead on the [Piketty and Saez \(2003, updated to 2015\) series](#) of top income shares, which were constructed from annual tabulations of income and its composition by size of income (US Treasury Department, Internal Revenue Service, annual since 1916). We made minor adjustments to the Piketty-Saez estimates to fix an inconsistency in the composition (but not in the level) of top incomes early in the twentieth century. Between 1927 and 1936, the composition of income in the top 10% and top 5% as estimated by Piketty and Saez (2003) is sometimes inconsistent in the sense that dividends, interest, rents and royalties earned by the top 10% and top 5% can exceed total taxable dividends, interest, rents and royalties.<sup>9</sup> In order to address this issue, we assume that no group of tax units above the 90th percentile can earn more than 95% of any income category; this condition is imposed in the the program `pre62.do` (described below) that

---

<sup>9</sup>This is mainly due to the following issue. Piketty and Saez (2003) add returns to account for the fact that there are missing filers, as the exemption threshold for married couples is much higher than for singles. The added returns are added bracket by bracket (\$1,000-\$2,000, \$2,000-\$3,000, \$3,000-\$4,000, \$4,000-\$5,000) using extrapolation based on the ratio between the number of single men with no dependents and the number of married joint filers from other years. The problem is that composition tables for 1927-1936 do not break down brackets below \$5,000. Piketty and Saez (2003) assumed that the added returns have the same income composition as all returns below \$5,000, which is a problem because the composition of income changes significantly below \$5,000. To account for that problem, they shaved off a few points of dividend share in the composition tables from 1923-1937 for P90-95 and P95-99. They did, however, very little correction to rents/royalties and interest.

distributes national income and wealth before 1962.

### **A.1.3 Survey data**

The third key ingredient of our distributional national accounts is survey data. We use two surveys: the Survey of Consumer Finances (SCF) and the Current Population Survey (CPS).

**SCF.** The [Survey of Consumer Finances](#) is available on a triennial basis from 1989 to 2013. It is a high quality survey that over-samples wealthy individuals. The Federal Reserve Board disseminates two sets of SCF files: summary extract public data (containing a number of key summary variables such as net wealth, expressed in inflation-adjusted dollars) and the full public data set (with the raw data in current dollars). We use both the summary extract and full public files. We rely on the SCF to impute forms of wealth (and therefore economic capital income) that cannot be captured by capitalizing income tax returns, see program `use_uscf.do` described below.

**CPS.** We use the CPS March Supplement data, which are available since 1962 through the [NBER](#). We use the Stata programs available on the NBER website to convert the raw data files into Stata. For March 2014, we use the traditional CPS file rather than the re-designed survey. Complete documentation is available on the website of the [Center for Economic and Policy research \(CEPR\)](#). The CPS data are used to create a sample of non-filers (following the methodology developed by the Tax Policy Center for its tax simulator) and to impute a number of benefits that are not reported in tax data.

## **A.2 Programs creating US distributional national accounts**

Here we describe the organization of the [Stata programs](#) we use in this research. The master program is `runusdina.do`; this program defines the paths and calls all the other programs described below. The programs are designed to run both externally using solely public use sources and internally within IRS using internal data. The external programs also use a series of additional basic tabulations based on internal data that are gathered in Saez (2016).

### **A.2.1 Programs using survey data**

We start by constructing homogeneous SCF and CPS datasets, i.e., datasets that contain the same variables over time, and use them to compute the distributions of the income and wealth components that cannot be captured by tax data.



**Program use\_scf.do.** We use the SCF to estimate the distribution of housing wealth for non-itemizing tax units, mortgage debt for non-itemers, currency assets, and other debts. The program `use_scf.do` (i) constructs SCF datasets with the relevant income and wealth variables using consistent definitions over time and (ii) estimates the distributions of housing, mortgage debt, non-mortgage debt, and currency assets.

**Program use\_cps.do** We use the CPS to estimate the distribution of a number of benefits that cannot (or only imperfectly) be observed in tax data: employee fringe health and pension benefits; Social Security benefits, Supplemental Security income, food stamps/SNAP, Veterans' benefits, AFDC/TANF, and Medicaid. The program `use_cps.do` (i) constructs CPS datasets with the relevant income variables using consistent definitions over time and (ii) estimates the distribution of the above benefits. We also use the CPS to estimate how wage income is split among spouses in married couples for the years when we do not have information on this income split from tax data. See Section B.2 below.

### A.2.2 Programs building tax data

Next, we construct homogeneous tax dataset, i.e., dataset that contain the same tax variables over time.

**Program build\_small.do** This program constructs annual datasets of income tax returns micro-data using the public-use micro-files (PUF). It starts from the public-use tax files disseminated through the [NBER](#) and constructs consistent income variables over time.

**Program aggrecord.do** This program adds a synthetic record to the homogeneous PUF files constructed by `build_small.do`. The synthetic record insures that the total income in the PUF matches the total income reported to the IRS, component by component. Starting in 1996, the PUF have excluded extreme records from their sampling. From 1996 to 2008, the number of excluded records was small (between 13 and 191 in these years, as reported in the official PUF documentation), but large enough in income size to create significant discrepancies at the very top. Starting in 2009, the PUF excludes a larger number of extreme records (slightly over 1000) from its sampling but it aggregates all the excluded records into an aggregate record. We add a synthetic record to each PUF file over the 1996-2008 period; see Saez (2016) for more details.

**Programs `build_xdata.do` and `build_xfile.do`.** These programs are run on internal SOI tax files. They extract from the SOI files the variables available in the PUF and rename them following the PUF variables names used at NBER. They also merge in additional wage earnings split data from the population wide databank since 1999. These programs generate datasets that have the same structure as the PUFs but with additional variables such as demographic and earnings split information. The same program `build_small.do` can also be run on these datasets.

### A.2.3 Programs enriching public-use tax data

Next, we enrich the public-use micro-tax files by splitting income between married spouses, adding coarse demographic characteristics, and adding synthetic records for the income of non-filers.

**Programs `sharebuild.do` and `addsharef`.** These programs split income between spouses in married couples. They use tabulations on earnings split from internal data from Saez (2016) and CPS data to estimate how the share of total income earned by the secondary earner varies with income (see Section B.2 below) and impute the earning split in the homogeneous PUF files constructed by `build_small.do`.

**Program `impute.do`** This program imputes age and gender information in the post-1979 homogeneous PUF files constructed by `build_small.do` using tabulations of internal SOI tax files presented in Saez (2016). It also imputes gender of single filers over the period 1962-1978.

**Program `nonfilerappend.do`** This program adds records representing non-filing households in the homogeneous PUF files constructed by `build_small.do`. Non-filers are created using CPS data but adjusted to match statistics of non-filers for the period 1999-2014 presented in Saez (2016). See Section B.1 below for more details on the imputation of non-filers.

### A.2.4 Programs creating DINA micro-files

Next, we combine the enriched micro tax data with survey data and national accounts data to create our distributional national accounts micro-files.

**Program `build_usdina.do`** This program creates one distributional national accounts micro-file per year since 1962 (except in 1963 and 1965 when there are no publicly available samples of individual income tax returns). It uses the enriched PUF files constructed above, and from there

creates wealth variables matching Financial Accounts totals and income variables matching National Income and Product accounts totals. To do so, it calls the file `parameters.csv` that contains the national accounts totals constructed in [Appendix Tables I \(Macro\)](#). It splits income between spouses and imputes forms of income that are not captured by the tax data using the CPS and SCF distributions constructed in `use_cps.do` and `use_scf.do`

**Program `pre62.do`** This program construct top income and wealth shares matching national accounts macroeconomic totals for the period 1913 to 1961, when no micro tax data are available. It starts from the [Piketty and Saez \(2003, updated to 2015\) series](#) of top income shares, which were constructed from annual tabulations of income and its composition by size of income (US Treasury Department, Internal Revenue Service, annual since 1916).<sup>10</sup> It then constructs factor income, pre-tax income and post-tax income matching national income, making adjustments in 1962 to ensure consistency with post-1962 estimates obtained from micro-data. While our post-1962 distributional national accounts micro-file contain the full distribution (including for the bottom groups) of a wide range of national accounts variables, before 1962 we only focus on top groups (top 10% and above) and on a selected set of core national account variables. This selected set includes pre-tax income, post-tax income, wealth, taxes paid, transfers, and their components.

### A.2.5 Programs producing statistics and exporting

**Program `outsheet_dina.do`** This program computes a large set of summary distributional statistics using the 1962-onward annual distributional national accounts micro-files constructed by `build_usdina.do`. These statistics are computed both on the internal data and on the external data. Saez (2016) compares the internal vs. external series. All series are gathered in the companion excel file of Saez (2016). In this paper, we use the internal series whenever they are available as they are based on more comprehensive data requiring fewer imputations.

**Program `graph_dina.do`** This program combines the post-1962 summary distributional statistics created by `build_usdina.do` with the pre-1962 top shares constructed by `pre62.do` to assemble homogeneous, long-run series of income and wealth share series and their components matching macro totals. It also export results to the Excel file [Appendix Tables II \(Distributions\)](#).

---

<sup>10</sup>All the issues since 1916 have been digitized by the Statistics of Income division of the IRS and posted online at <https://www.irs.gov/uac/soi-tax-stats-archive> making this very rich and valuable data source easily accessible.

Programs **taxrates.do**, **export\_dina.do** and **export-to-excel.do** These programs make additional exports of the results.

## A.3 Micro-files of U.S. distributional national accounts

### A.3.1 External-use DINA files

The main output of this research is a set of annual micro-files representative of the U.S. economy, where each line is a synthetic individual created by combining tax, survey, and national account data, and each column is a variable of the national accounts. There is one file per year since 1962 (except in 1963 and 1965 when there are no publicly available samples of individual income tax returns). These [Distributional National Accounts micro-files](#) are available online and structured as follows. The files are at the adult individual (aged 20 and above) level, so the sum of weights (variable `dweight`) adds up to the adult population, 226 million in 2010. The variable “id” identifies tax units, which makes it possible to compute statistics at the tax-unit level rather than adult level. The files also include socio-demographic information: age (restricted to three age categories, 20 to 44 years old, 45 to 64, and above 65), gender, marital status, and number of dependent children.

### A.3.2 Codebook

Our [Distributional National Accounts micro-files](#) contains the following 140 variables each year:

#### 1. Socio-demographic characteristics

- **id**: Tax unit ID
- **dweight**: Population weight  $\times$  100,000
- **dweighttaxu**: Population weight to recover the Piketty-Saez total tax units
- **female**: Dummy for being female (PUF year 1969, 1974 and non-MFJ only, 2009 for all)
- **ageprim**: Imputed age of primary filer (husband if married)
- **agesec**: Imputed age of wife in married filers
- **age**: Imputed age
- **oldexm**: Dummy for primary filer being 65+
- **oldexf**: Dummy for secondary filer being 65+

- **old**: Aged 65+
- **oldmar**: Interaction of married and old
- **married**: Dummy for being married joint return (filing status)
- **second**: Dummy for being secondary earner
- **xkidspop**: Number of children (imputed to match population total j20)
- **filer**: Tax filer dummy

## 2. Core income and wealth series

- **finc**: Fiscal income (incl. capital gains) =  $fiwag + fibus + firen + fiint + fidiv + fikgi$
- **fninc**: Fiscal income (excl. capital gains) =  $fiwag + fibus + firen + fiint + fidiv$
- **fainc**: Personal factor income =  $finc + fkinc$
- **finc**: Personal factor labor income =  $flemp + flmil + flprl$
- **fkinc**: Personal factor capital income =  $fkhou + fkequ + fkfix + fkbus + fkpen + fkdeb$
- **ptinc**: Personal pre-tax income =  $plinc + pkinc$
- **plinc**: Personal pre-tax labor income =  $finc + plcon + plbel$
- **pkinc**: Personal pre-tax capital income =  $fkinc + pkpen + pkbek$
- **diinc**: Extended disposable income =  $dicsh + inkindinc + colexp$
- **princ**: Factor national income (matching national income) =  $fainc + govin + npinc$
- **peinc**: Pre-tax national income (matching national income) =  $ptinc + govin + npinc + prisupen + invpen$
- **poinc**: Post-tax national income (matching national income) =  $diinc + govin + npinc + prisupenprivate + invpen + prisupgov$
- **hweal**: Net personal wealth =  $hwequ + hwfix + hwhou + hwbus + hwpen + hwdeb$

## 3. Detailed income and wealth series

- **fiwag**: Fiscal income, wages and pensions

- **fibus**: Fiscal income, business income
- **firen**: Fiscal income, rents
- **fiint**: Fiscal income, interest
- **fidiv**: Fiscal income, dividends
- **fkgi**: Fiscal income, capital gains
- **fnps**: Fiscal income (excl. KG), flat income for non-filers, matching PS shares
- **peninc**: total taxable pension income (=DB+DC+IRA withdrawals, but not Social Security)
- **schcinc**: schedule net income
- **scorinc**: S corp net income
- **partinc**: partnership net income
- **rentinc**: net rental income from Schedule E = rentincp-rentincl
- **estinc**: estate and trust net income
- **rylinc**: royalties net income
- **othinc**: other income in AGI
- **flemp**: Compensation of employees
- **fmil**: Labor share of net mixed income
- **flprl**: Sales and excise taxes falling on labor (prop. to  $Y^*(1-s)$ )
- **fkhou**: Housing asset income
- **fkequ**: Equity asset income
- **fkfix**: Interest income
- **fkbus**: Business asset income
- **fkpen**: Pension and insurance asset income

- **fkdeb**: Interest payments
- **plcon**: (Minus) social contributions (pensions + DI + UI, employers + employees + self-e
- **plbel**: Labor share of social insurance income (pensions + DI + UI)
- **pkpen**: (Minus) Investment income payable to pension funds (DB + DC + IRA, but excluding
- **pkbek**: Capital share of social insurance income (pensions + DI + UI)
- **hwequ**: Equity assets (div+KG capitalized)
- **hwfix**: Currency, deposits, bonds and loans of households
- **hwhou**: Housing assets
- **hwbus**: Business assets
- **hwpen**: Pension and life-insurance assets
- **hwdeb**: Liabilities of households
- **flwag**: Taxable wages of filers + non-filers
- **flsup**: Supplements to taxable wages
- **waghealth**: Health insurance contributions
- **wagpen**: Pension contributions (employer + employee)
- **fkhoumain**: Main housing asset income
- **fkhourent**: Rental housing asset income
- **fkmor**: Mortgage interest payments
- **fkmmo**: Non-mortgage interest payments
- **fkprk**: Sales and excise taxes falling on capital (prop. to  $Y^*(1-s)$ )
- **proprestax**: Residential property tax (prop. to housing assets)
- **propbustax**: Business property tax (prop. to wealth excl. housing)

- **rental**: Tenant-occupied housing wealth, net of mortgage debt
- **rentalhome**: Gross tenant-occupied housing
- **rentalmort**: Mortgages on tenant-occupied houses
- **ownerhome**: Gross owner-occupied housing wealth
- **ownermort**: Mortgages on owner-occupied houses
- **housing**: Housing wealth, net of mortgage debt
- **partw**: Partnership wealth
- **soleprop**: Sole proprietorship wealth
- **scorw**: S-corporations equities
- **equity**: Equity assets (only div. capitalized)
- **taxbond**: Taxable fixed-income claims
- **muni**: Tax-exempt municipal bonds
- **currency**: Currency and non-interest bearing deposits
- **nonmort**: Non-mortgage debt
- **hwealnokg**: Net personal wealth (KG not capitalized)
- **hwfin**: Financial assets of households
- **hwnfa**: Non-financial assets of households
- **plpco**: (Minus) pension contributions (employer + employee + self-employed, SS + non-SS)
- **ploco**: (Minus) DI and UI contributions (employer + employee + self-employed)
- **plpbe**: Pension benefits (SS + non-SS)
- **plobe**: UI and DI benefits
- **plben**: Social insurance income (pensions + DI + UI)



- **plpbl**: Labor share of pension benefits
- **plnin**: Personal pre-tax labor income (narrow definition: pensions only)
- **pkpbk**: Capital share of pension benefits
- **pknin**: Personal pre-tax capital income (narrow definition: pensions only)
- **ptnin**: Personal pre-tax income (narrow definition: pensions only)
- **dicsh**: Disposable cash income
- **inkindinc**: Social transfers in kind
- **colexp**: Collective consumption expenditure
- **govin**: (Minus) Net property income paid by gov. (allotted 50% prop. to taxes, 50% to be
- **npinc**: Net primary income of non-profit institutions (prop. to disposable income)
- **prisupen**: Primary surplus (= contrib - distrib) of pension system
- **invpen**: Investment income payable to pension funds (DB + DC + IRA, but excluding life in
- **prisupenprivate**: Primary surplus (= contrib - distrib) of private pension system (al-
- **prisupgov**: Government primary surplus (= taxes - benefits) (allotted 50% prop. to
- **educ**: Education collective consumption expenditure
- **colexp2**: Collective consumption exp. (with lump sum educ.)
- **poinc2**:
- **tax**: Total taxes and social contributions paid
- **ditax**: Current personal taxes on income and wealth
- **ditaf**: Federal personal income tax (gross of EITC, additional CTC, and other refundable

- **ditas**: State personal income tax
- **salestax**: Sales and excise taxes (prop. to  $Y^*(1-s)$ )
- **corptax**: Corporate tax (prop. to wealth excl. housing)
- **estatetax**: Estate tax
- **govcontrib**: Total contributions for government social insurance
- **ssuicontrib**: Contributions for government social insurance: pensions, UI, DI
- **othercontrib**: Contributions for government social insurance other than pension, UI, DI
- **ssinc\_oa**: Social Security income (old age)
- **ssinc\_di**: Social Security income (disability)
- **uiinc**: Unemployment insurance benefits
- **ben**: Total benefits (cash + kind + coll, excl. pensions, UI, DI)
- **dicab**: Social assistance benefits in cash
- **dicred**: Refundable tax credits
- **difoo**: Food stamps (SNAP)
- **disup**: Supplemental security income
- **divet**: Veteran benefits
- **diwco**: Workers' compensation benefits
- **dicao**: Other social assistance benefits in cash
- **tanfinc**: TANF / AFDC benefits
- **othben**: Other cash benefits (State and local benefits similar to SNAP, etc.)
- **medicare**: Medicare = capitation for 65+ individuals
- **medicaid**: Amount of medicaid benefits received by tax unit
- **otherkin**: Other in-kind transfers (pell grants + state schships + veterans' health care +

- **pell**: Pell grants received
- **vethealth**: Veteran in-kind health benefits

### A.3.3 Internal-use DINA files

We have constructed similar distributional national accounts micro-files but starting from the internal use SOI files of individual income tax returns rather than the PUF tax files disseminated through the [NBER](#). The internal-use files go from 1979 to 2014 and they contain exact age and gender information, exact self-employment earnings split, and exact wage earnings split since 1999. Before 1979, there are no internal files so external and internal computations are identical before 1979. The same set of programs builds both the external and internal dina files (and outputs the same statistical series). We use adjustments specific to external vs. internal data embedded in the programs with the use of a global macro `data` (equal to 0 for external and equal to 1 for internal).

## A.4 Results: Distributional summary statistics

### A.4.1 Main distributional results

The main distributional results are presented in our [working paper](#) (Piketty, Saez and Zucman 2016) and the accompanying Excel file of [Main Tables and Figures](#). This file includes the 13 Figures (with 2 panels each) and 2 tables included in the [working paper](#), as well as a limited set of supplementary figures and tables (numbered FS.1, FS.2, TS.1, TS.2, etc.), which are also printed at the end of this appendix document.

### A.4.2 Detailed distributional results

In addition to the main distributional results gathered in [Main Tables and Figures](#), we present detailed appendix distributional results in the Excel file [Appendix Tables II \(Distributions\)](#). This file shows the distribution of factor income (Tables II-A1 to II-A14), pre-tax income (Tables II-B1 to II-B15), post-tax income (Tables II-C1 to II-C13), fiscal income (Tables II-D1 to II-D13), and wealth (Tables II-E1 to II-E13) with detailed decompositions. In addition, Tables II-F1 to II-F3 provide series on the gender labor income gap, average age by pre-tax income group, and the age-profile profile, and Tables II-G1 to II-G4 provide series on the distribution of taxes and benefits, with detailed decompositions.

## A.5 List of online files

- [Working paper](#): Text of the main paper, November 2016.
- [Main Tables and Figures](#): Excel file with all the tables and figures included in the working paper and a small set of supplementary tables and figures.
- [Appendix Tables I \(Macro\)](#): Excel file with all the macroeconomic series used in this research (national income accounts, Financial Accounts, and historical national account series).
- [Appendix Tables II \(Distributions\)](#): Excel file with detailed distributional summary statistics.
- [Distributional National Accounts micro-files](#): Stata micro-files of synthetic observations representative of the U.S., with detailed income and wealth components matching national accounts totals.
- [Stata programs](#): Complete set of programs used to compute our distributional national accounts microfiles by combining tax, survey, and national accounts data.
- [Data Appendix](#): Text of the Data Appendix, October 2016.
- [Supplementary information on internal SOI computations](#): See “Improving the Individual Tax Return Public Use Files for Tax and Distributional Analysis” (Saez, 2016)

## B Supplementary information on imputations methods

All the details on the imputations methods we use are provided in the online computer code described above. Here we simply provide additional information on a number of issues.

### B.1 Imputation of non-filers

About 10-15% of adults aged 20 and above do not file tax returns. To supplement tax data, we start by adding synthetic observations representing non-filing tax units using the Current Population Survey (CPS). We identify non-filers in the CPS based on their taxable income

following the methodology developed by the Tax Policy Center for their tax simulator.<sup>11</sup> We proceed in three steps.

First, we create tax units in the CPS data, then we apply filing thresholds to the tax units. As tax units with children are eligible for refundable credits, we assume that all tax units with children file tax returns whenever they have positive labor income. All tax units that fall below the filing thresholds are considered to be non-filers. Second, we re-weight these observations such that the total number of adults in our final dataset matches the total number of adults living in the United States, for both the working age population (aged 20-65) and the elderly (above age 65). The weights are fairly close to one in general. Third, we have used the detailed tabulations on statistics for non-filers from IRS data for 1999-2014 presented in Saez (2016) to adjust the CPS based non-filers. Social security benefits, the major income category for non-filers is very close and does not need adjustment. However, there are more wage earners and more wage income per wage earner in the IRS non-filers statistics (perhaps due to the facts that very small wage earners may report zero wage income in CPS and a fraction of individuals with wages may fail to file income tax returns). Over the period 1999-2014, there about twice as many wage earners in the IRS non-filer sample than in our CPS built sample of non-filers, and conditional on being a wage earner, they earn about twice as much as in the IRS non-filer sample than in the CPS. Hence, when building the CPS sample (and before adjusting weights for age), we simply re-weight wage earners in the CPS by a factor 2 and we uniformly increase their wages by a factor 2. With these adjustments, the CPS based sample of non-filers has about as many wage earners and the same aggregate wage amount as the IRS non-filer sample for the period 1999-2014. We apply the same adjustment to wages retro-actively over the full period 1962-2014.<sup>12</sup>

---

<sup>11</sup>Tax simulators typically rely on tax return data. However, for some tax simulations (such as reducing exemptions or deductions), it is important to capture the population of non-filers as some current non-filers might become taxpayers. Hence, the Tax Policy Center developed a method to model non-filers using CPS data. Their methodology is presented in detail in Rohaly, Carasso, and Adeel Saleem (2015). We thank the Tax Policy Center for sharing their programs with us allowing us to build on their work.

<sup>12</sup>In the future, it might be possible to construct a file of non-filers using exclusively tax data as done in Saez (2016). We have not followed this route because non-filers can be obtained from tax data only for the period 1999-2014. Furthermore, there is no simple route to determine the marital status and link spouses together in the non-filer IRS data. The spousal link is important as we split income equally across spouses in our benchmark series. A fairly large numbers of non-filers are retirees with only Social Security Income, a significant fraction of whom are married.

## B.2 Imputation of income split among spouses

Because the US income tax is family based, one needs to supplement income tax returns by other data sources to estimate individual income. The CPS contains information on individual income but not in top groups because of top coding. Individual income is available from W2 tax forms, but only since 1999 and a few years isolated years before 1999 (namely 1969, 1974, and 1982-6). We therefore proceed as follows.

First, we always split capital income and wealth 50/50 because of the lack of data on property regimes. Next, we estimate individual wage income by combining CPS and tax data. We use the CPS to estimate the wage split among couples in the bottom 95% of the wage income distribution among married couples with positive wage income. We use tax data to estimate the wage split in the top 5%. More precisely, for 1969 and 1974, wages from W2 tax forms are included in the public-use micro-files of tax returns. Between 1982 and 1986, secondary earners had a tax credit and therefore wage split information is also included in the PUF. Since 1999, we rely on tabulations of share of female wages in total family wages by fractiles of family wage income (among married joint filers with positive wage income) computed in the internal-use SOI samples files (presented in Saez, 2016).<sup>13</sup> For the years with no direct information, we linearly interpolate the wage split using the closest years with available data. We have checked that this imputation method delivers results that are close to the results obtained for years when we have exact wage split (1969, 1974, and 1999-2014).

Regarding self-employment income, we assume that 70% of self-employment income is labor income and 30% is capital income. We split the capital portion 50/50 (just like other forms of capital income). We split the labor portion of self-employment income using tax based information on the split of self-employment income across spouses (this information is reported on tax returns individually for each spouse on the Schedule SEs). In external tax data, the self-employment individual earnings variables are capped (and start in 1984) but they are fully uncapped and start in 1979 in internal data. In internal computations, we use the exact split for 1979 on. In external computations, we use the exact split when the variables are not capped. When the variables are capped, we rely on tabulations of the share of self-employment income earned by the female spouse by fractiles to family self-employment income in married couples with positive self-employment income. These tabulations are presented in Saez (2016). Before 1979, we assume that the split of the labor portion of self-employment income is the same as in 1979. Last, we split benefits and pension income 50/50.

---

<sup>13</sup>In the internal DINA files, we use the true wage of each spouse as reported on W2 forms since 1999.

## B.3 Imputations of wealth and income from survey data

### B.3.1 General imputation method

Generally speaking, our strategy to impute income or wealth components seen in surveys but not in tax data is the following. Take an income or wealth component  $y$ , for instance Veterans' benefits or non-mortgage debt. First, we use survey data to estimate the probability that a tax unit earns or owns  $y$  in each of the 40 following bins: decile of the taxable income distribution  $\times$  marital status  $\times$  primary earner above or below 65 years old. We then compute the average amount of  $y$  (conditional on earning/owning  $y$ ) in each bin. Next, in the enriched tax data created by the program `build_small` (see above), we randomly impute  $y$  in each bin of taxable income decile  $\times$  marital status  $\times$  above 65 using the probability distributions estimated in the surveys and assigning the bin-specific average amount of  $y$  to each recipient. We adjust the number of recipients and/or the average amount of  $y$  proportionally in each bin in order to match macroeconomic totals. For instance, we adjust the number of imputed Medicaid beneficiaries so as to match the true number of beneficiaries from administrative records; we adjust the average amount of imputed non-mortgage debt such that total non-mortgage debt adds up to the total recorded in the Financial Accounts.

For some types of income, we implement variants of this imputation method which can be either more sophisticated (e.g., using more detailed bins, such as for Medicaid benefits) or less sophisticated depending on data availability. In a number of important cases we have also checked the consistency between the distribution of our imputed income components and information available from other sources.

The advantage of our imputation method is that it is simple to implement and overcomes automatically issues of under reporting in the Current Population Survey that are well known (see Meyer, Mok, and Sullivan 2009, 2015 for detailed analyzes and discussion). Our method does not require developing a sophisticated and granular benefit calculator either.<sup>14</sup> Our imputation is designed to respect the correlation of benefits reciprocity with income, marital, and elderly status.

The drawback of our imputation method is that it fails to capture correlations in benefits reciprocity across programs (for example, TANF recipients are automatically eligible for Medicaid, etc.), or the fact that reciprocity is often correlated with observable characteristics (such as age or presence of children in the household). Importantly, our methods should be seen as a

---

<sup>14</sup>The most extensive effort to create a benefit simulator in the United States is the Transfer Income Model (TRIM) of the Urban Institute. See Zedlewski and Giannarelli (2015) for a general presentation.

simple initial benchmark which will could be further refined and improved down the road.

### B.3.2 Wealth imputations

We construct wealth by capitalizing income tax returns and accounting for the forms of wealth that do not generate taxable income, as in Saez and Zucman (2016). The only methodological difference with Saez and Zucman (2016) is that we provide better imputations of housing wealth for non-itemizers, mortgage debt for non-itemizers, currency assets, and other debts. Namely, we compute the distribution of these assets in the SCF and randomly impute them to tax units following the method described in Section B.3.1 above; see program `use_scf.do`.

### B.3.3 Income imputations

**Employee fringe benefits** Pension contributions are only imperfectly captured in the CPS. The CPS provides information on who is covered by an employee pension plan, but does not provide data on the amounts contributed, which can be hard to know for respondents, in particular in the case of defined benefit plans. Therefore, we use the CPS to impute the probability to be covered by a pension plan by income decile  $\times$  marital status  $\times$  above 65; within each bin we then assume that the pension contributions are proportional to wages winsorized at the 99th percentile. We winsorize because there are annual caps on defined contributions such as 401(k)s. Defined benefit plan tend to be proportional to wages (often with a cap). We have checked using IRS Statistics of Income tabulation of elective pension contributions and retirement plan indicators by brackets of wage income that this is a reasonable assumption.<sup>15</sup>

Health benefits are better captured in the CPS as there is information on the cost of health insurance plans. We therefore apply our standard imputation method described in Section B.3.1 above. Since 2012, health benefits have been reported on W2 forms. We have checked that the distribution of health benefits in the CPS and our imputed series is consistent with the high-quality information available in internal-use SOI sample tax files. In the future, it should be possible to use these data to do better imputations of employer based health insurance benefits.

**Individualized transfers** We use the CPS to estimate the distribution of the following individualized transfers: Social Security benefits, Supplemental Security income, food stamps/SNAP, Veterans' benefits, AFDC/TANF, and Medicaid. We follow the imputation strategy described

---

<sup>15</sup>These statistics are available online for years 2008-2010 at <https://www.irs.gov/uac/soi-tax-stats-individual-information-return-form-w2-statistics>. In the future, it should be possible to use these data to do better imputations of pension contributions.



in Section B.3.1 above; that is, we compute the joint distribution of taxable income and Social Security benefits (resp. Supplement Security income, SNAP, etc.) by estimating the probability to receive Social Security in 40 income bins (decile of the taxable income distribution  $\times$  marital status  $\times$  above or below 65 years old), and we then compute the average amounts of Social Security benefits (resp. Supplement Security income, SNAP, etc.) in each bin conditional on receiving Social Security income (resp. Supplement Security income, SNAP, etc.). For Medicaid, we use more detailed bins (8 income groups  $\times$  marital status  $\times$  number of kids  $\times$  number of Medicaid recipients).

## B.4 Imputation of taxes

Computing pre-tax income shares requires making tax incidence assumptions. As is well known in tax incidence theory, the burden of taxes is not necessarily borne by the economic agent who nominally pays them. Behavioral responses to taxes can affect the relative prices of factors hereby shifting the tax burden on a given factor to other factors. Taxes, on top of being shifted, also generate deadweight burden (see Fullerton and Metcalf, 2002 for a survey of the economic literature). In this paper, we do not attempt to measure the complete effects of taxes and transfers on economic behavior and ultimate money-metric welfare of each individual. There is a long tradition in federal government agencies to present distributional tables of the Federal tax burden which naturally require making tax incidence assumptions (see e.g., US Congressional Budget Office (CBO), 2016). We discuss how we depart from the official CBO tax incidence assumptions. In contrast to CBO statistics, we compute taxes not only at the federal level but also include all taxes at the State and Local level.<sup>16</sup> We make the following simple assumptions regarding tax incidence.

### B.4.1 Individual income and payroll taxes

We assume that individual income taxes are paid by individual taxpayers with no further incidence. We also assume that payroll taxes are paid entirely by labor, regardless of whether they are nominally paid by employers or employees. Because we expect labor demand to be more elastic than labor supply, standard tax incidence theory predicts that payroll taxes should primarily fall on workers (see e.g. Hamermesh 1993). This has been the standard assumption in most analyzes of the distributional effects of taxes and is the CBO assumption as well.

---

<sup>16</sup>In contrast to federal taxes, there are no systematic or official statistics on the tax incidence of state and local taxes, owing to the decentralized nature of US governments, each local government having its specific tax system.

## B.4.2 Corporate income tax

We assume that the corporate income tax falls on all capital except housing (net of mortgage debt). In practice, this means that equity owners bear only about 45% of corporate taxes, since equity wealth (including equities held through pension plans) is about 45% of all non-housing wealth today. The owners of fixed income claims (deposits, bonds, etc., including held through pension plans) bear about 40% of the corporate tax, and the owners of non-corporate business assets the remaining 15%; see [Appendix Table I-S.A9](#). Because the effective tax rate on U.S. corporate profits (i.e., the amount of corporate tax paid divided by the flow of corporate profits accruing to U.S. residents, either from domestic or foreign-owned corporations) has been around 25% in recent years, our assumptions imply that all equity owners pay around  $25\% \times 45\% = 11.25\%$  in corporate taxes out of the corporate profits that accrue to them. One implication is that in our distributional national accounts, by construction billionaires—for whom almost all wealth is equity and all income is corporate profits—cannot have an overall tax rate below 11.25%. This is true including for the owners of companies that in practice pay zero or very small amounts of corporate taxes.

Note that in the U.S. national accounts, the profits made by the Federal Reserve System are treated as corporate taxes received by the Federal government. These profits have been large since the financial crisis of 2008-09 and the subsequent expansion of the Fed's balance sheet. They amounted to about \$80 billion in 2013, almost 20% of all corporate income taxes received by U.S. governments, Federal and States. We have followed the U.S. national accounts convention of treating the Federal Reserve Systems's profits as corporate tax revenue for the government, although it seems more meaningful to view these profits as dividends paid to the government rather than taxes. We may change our treatment of Fed's profits in the future as we learn more about how other countries proceed.<sup>17</sup>

CBO assumes that corporate taxes fall 75% on capital and 25% on labor (CBO earlier assumed that corporate taxes fell 100% on capital).<sup>18</sup> We think that the assumption of 100% incidence on capital is a more reasonable benchmark as there is no compelling empirical evidence that corporate taxes depress wages in a large economy like the United States. If the corporate tax

---

<sup>17</sup>Zucman (2014) removed Fed's profits for corporate tax revenue to compute effective corporate tax rate on U.S. corporate profits, which explains why he finds effective corporate tax rates around 20% rather than the 25% implied by a naive computation that does not remove Fed's profits.

<sup>18</sup>The Tax Policy Center (TPC) (an external non-governmental think-tank) assigns 80% of the corporate tax burden on capital and 20% on labor. However, in practice, CBO and TPC do not assign corporate taxes on all forms of capital, but solely on capital income reported on individual tax returns, hereby excluding pension funds from the corporate tax burden.

is partly shifting on labor income, then the overall tax system would be slightly less progressive than what we obtain here. It is fairly easy to modify our programs to change our assumption on corporate tax incidence.

In contrast to CBO, we do not assume that the corporate tax falls on housing capital because we think the housing market vs. the corporate equity market are not sufficiently integrated as we discuss just below.

### **B.4.3 Property taxes**

About 60% of property taxes are business property taxes; the remaining 40% correspond to residential property taxes. We assume that business property taxes are borne by all capital excluding housing, just like the corporate tax, and that residential property taxes are borne by the owners of housing assets. These assumptions are consistent with the following model that has two forms of capital: corporate assets vs. housing assets (in the spirit of Harberger, 1962).

With an infinite elasticity of substitution between different forms of capital, the after-tax rate of return must be the same for all forms of capital assets. In such a model, any tax on a capital asset is effectively shifted on all capital assets (this is the CBO assumption for corporate taxes). While the elasticity of substitution between different forms of financial assets is probably very high, there seems to be relatively little substitutability between housing and non-housing. Typically, many families own real estate as a place to live and own corporate assets through their pension funds without arbitraging between assets as in a standard frictionless asset pricing model. This motivates our assumption to treat housing and non-housing assets as two separate sectors, and within the non-housing sector to assume that taxes fall broadly on all forms of non-housing assets. In the housing sector, the incidence of the residential property tax then depends on how elastic housing supply and housing demand are. If the supply of housing is not elastic (e.g., it is impossible to build new housing units due to lack of land or excessive regulations) relatively to the demand of housing, then the incidence of the property tax falls on the owners of housing; if the supply of housing is very elastic (land is plentiful so that new housing can be built at a fixed marginal cost), then the incidence falls on tenants. There is evidence that in the short run, housing supply is not very elastic and therefore that housing subsidies benefit landlords and property taxes are paid by landlords. Fack (2006) finds that in France, one additional euro of housing benefit leads to an increase of 78 cents in the rent paid by new benefit claimants. There is more uncertainty regarding the long-run incidence of residential property taxes; in the long run, housing supply may become more elastic, shifting part or all of

the burden of residential property taxes to tenants. Using our micro-files, it is straightforward to assume that a fraction of residential property taxes is shifted to prices rather than borne by capital owners. We have tested the extreme scenario where housing supply is infinitely elastic and residential property taxes are entirely shifted to prices. Because the residential property tax is modest in the U.S. (\$166 billion in 2015, only 3% of total tax revenue), this makes only a small difference on our pre-tax income series.

#### **B.4.4 Sales and excise taxes**

We assume that excise and sales taxes are paid proportional to disposable income less savings.

### **B.5 Collective consumption expenditure**

Computing post-tax income by adding back to individuals all forms of government spending requires making assumptions on who benefits from non-individualized government transfers, i.e., collective consumption expenditure for education, defense, etc.

#### **B.5.1 Health spending**

We assign Medicare and Medicaid benefits to beneficiaries on a lump sum basis; i.e., every Medicare-eligible adult gets the same benefit (about \$10,000 today) and every Medicaid-eligible individual (adult or child) gets the same benefit (about \$5,000 today). The benefits received by children are added to their parents' income.

For comparability with countries where there is no information on the age profile of government health benefits (which is typically the case in single-payer systems where almost all health spending is paid for by the government), we also provide additional series where we assign all Medicare and Medicaid beneficiaries the same average transfer (equal to total Medicare plus Medicaid spending divided by the total number of beneficiaries). One drawback of this procedure is that the Medicare lump sum transfer is significantly higher than the Medicaid lump sum transfer (about twice higher in recent years) and has been growing faster. Disregarding this heterogeneity artificially inflates the post-tax income and income growth of Medicaid-eligible working-age Americans.

#### **B.5.2 Education spending**

In our benchmark series, we allocate public education spending proportional to disposable income as well. This can be justified from a lifetime perspective where everybody benefits from

education, and where higher earners attend better schools and for longer. We have also considered a polar alternative where we consider the current parents' perspective and attribute education spending as a fix lump sum per child. In married couples, we attribute each child 50/50 to each parent. Note that children going to college and supported by parents are typically claimed as children dependents so that our lump sum measure gives more to families supporting children through college. This slightly increases the level of bottom 50% post tax incomes but without affecting the trend

## **B.6 Imputations to match national income**

### **B.6.1 Income of non-profits**

### **B.6.2 Government deficit and interest payments on debt**

In general, a government deficit means either higher taxes, lower government transfers or both in the future. Taxes can be implicit, e.g., financial repression (lowers returns on assets; may also lower bankers' wages). Transfer reduction: case of 19th century UK with arguably lower educational spending compared to counterfactual without Napoleonic war debts. Whether taxes or transfers adjust depends on regime (democracy vs. oligarchy), maybe the type of event generating the deficit (war vs. financial crisis; cyclical vs. structural), so no general rule on whether deficit should be seen as affecting future taxes more or less than future transfers. So to allocate the deficit to the existing population, we reduce everybody's transfers and increase everybody's taxes proportionally and with equal weights on taxes and transfers.<sup>19</sup>

Interest income paid on government debt is included in individuals pre-tax income but is not part of national income (as it is a transfer from government to debt holders). Hence, in our pre-tax measure, we also deduct interest income paid by the government in proportion to taxes paid and benefits received (50/50) in the same way we treat government deficit.

---

<sup>19</sup>A more sophisticated possibility would be to allocate the deficit based on age  $\times$  taxes paid  $\times$  transfers received (rather than taxes paid  $\times$  transfers received only). For instance, if government deficit is due to public transfers to elderly being bigger than taxes, and if people have life-cycle saving motive only, then currently young people are those who will ultimately have to pay more taxes/receive less transfers  $\rightarrow$  deficit could be seen as reducing young people's current transfers and increasing their current tax liabilities, leaving the disposable income of the elderly unchanged. Generational accounting is heavily influenced by this life-cycle saving view and so tends to see government deficit as affecting young people disproportionately. But if people have dynastic preferences, then the deficit will trigger increased saving by current pensioners, reducing their income available for consumption; in that case it's less clear that the deficit should only be allocated to young people.

### B.6.3 Surplus of pension system

We offer two treatments for pensions. In our “factor income” series ([Appendix Tables II-A1 to II-A14](#)), we treat pensions on a contribution basis. That is, we include all pension contributions and payroll taxes in our factor income flow, and we exclude all pension distributions from it. One problem with the concept of factor income is that it allocates little income to retirees who mostly live off of pension distributions, and therefore it tends to overestimate inequality in aging societies (although one can always look at inequality within the working-age population). That is why we prefer to use “pre-tax income” as our benchmark concept for the distribution of income before government intervention. In our pre-tax income series ([Appendix Tables II-B1 to II-B15](#)), pensions are treated on a distribution basis. That is, compared to factor income, pre-tax income deducts contributions and adds back pension benefits.

In a given year there is usually an imbalance between pension contributions and distributions. In the United States, as shown by [Appendix Table I-A7](#), pension contributions have been significantly larger than distributions over the last decades. Pension contributions include all Social Security contributions (the old-age portion of payroll taxes paid by employers, employees, and the self-employed, \$683 billion in 2015), non-Social Security contributions paid out of labor income (payments made by employers and employees to defined contributions plans; payments made by employers to defined benefit plans; contributions made by workers to their individual retirement accounts out of pre-tax income—all of which add up to \$921 billion in 2015), and the capital income earned by pension plans that is immediately reinvested in those plans (\$1,316 billion in 2015). In total, pension contributions amounted to \$2,920 billion in 2015; see [Appendix Table I-A7 and I-A.S10](#) for a detailed decomposition. By contrast, distributions amounted to \$1,706 billion only.

The excess of pension contributions over distributions can be decomposed into three terms: the surplus of the pay-as-you-go public pension system (Social Security), the primary surplus of private pensions (the excess of private pension contributions excluding reinvested capital income over private pension distributions), and the capital income reinvested in private pension plans. As shown by [Appendix Table I-A10](#), Social Security had a small surplus from 1984 to 2008, and has had a small deficit since 2009 (about \$100 billion in 2015). The primary surplus of private pensions is roughly zero. So as a first approximation the primary surplus of the overall pension system (Social Security + private pensions) is close to zero, i.e., pension distributions are roughly equal to the sum of all contributions paid out labor income. Almost all the gap between total pension contributions and distributions comes from the large amount of capital

income reinvested in pension plans.

An economy with funded pensions naturally tends to have contributions greater than distributions (see Alvarado et al. 2016). To see this, consider an economy in steady-state growth (fixed demographic and productivity growth rates, with a stable age structure) with a total growth rate equal to  $g = n + h$  (the sum of demographic and productivity growth), and an average return to capital equal to  $r$ . With a pay-as-you-go pension system, contributions are equal to pensions (disregarding small accounting surpluses or deficits), so that personal pre-tax income is equal to personal factor income in aggregate. However with a funded pension system with total steady-state pension wealth equal to  $W_{Pt} = \beta_P \cdot Y_t$  (where  $Y_t$  is personal factor income, growing at rate  $g$ ;  $W_{Pt}$  is pension wealth, also growing at rate  $g$ ; and  $\beta_P$  is the steady-state pension wealth-factor income ratio), one can immediately see that contributions (including accrued investment income) exceed pension distributions by  $g \cdot W_{Pt}$ .<sup>20</sup> So for instance if  $g = 2\%$  and steady-state pension wealth represents 300% of personal factor income, then in steady-state pretax income will be equal to 96% of personal factor income in a country with funded pensions (and 100% in a country with pay-as-you-go pensions). In the United States, as shown by [Appendix Figure S.24](#), personal pre-tax income has fallen from 100% of personal factor income in the 1960s to 90% in the mid-1990s, and has slightly rebounded since then to 92.5% in 2015. This ratio is low because pension wealth is still accumulating and has not reached a steady state yet: pension wealth is 160% of national income is still rising; see [Appendix Table I-B2](#)). As we approach a steady-state, the ratio of personal pre-tax to personal factor income should rise above 95%.

Because we want to match national income in all our series—whether looking at factor income, pre-tax income, or post-tax income—we allocate to individuals the excess of pension contributions over distributions. This procedure ensures that our pre-tax income total is the same as our factor income total (and equal to national income), which makes comparing growth rates straightforward. We proceed as follows. First, we allocate the surplus of the Social Security pension system proportionally to taxes paid and benefits received, just like we do for the general government deficit. This assumes that Social Security is fungible with other government revenue and expenditure, i.e., that any deficit will eventually have to be addressed by raising taxes or cutting government transfers.<sup>21</sup> Second, we allocate the (very small) primary surplus of the private pensions system proportionally to wages. Last, we re-attribute the capital

---

<sup>20</sup>Denote by  $c_t$  contributions out of labor income and  $d_t$  distributions, we have  $W_{Pt+1} = (1 + g)W_{Pt} = (1 + r)W_{Pt} + c_t - d_t$  so  $gW_{Pt} = rW_{Pt} + c_t - d_t$ , which is the excess of total contributions over distributions.

<sup>21</sup>Smetters (2004) provides supporting evidence that indeed what matters for decisions is the combined deficit.

income reinvested on pension plans to the owners of pension wealth. If the pension system was in a steady-state, it would be preferable to reattribute this income to retirees only for pre-tax income (in contrast to factor income). But since the U.S. private pension system is still in accumulation, we choose instead to reattribute investment income to both pensioners and retirees in proportion to their share of pension wealth, so as to avoid artificially impoverishing working-age adults. In recent years, about 60% of pension wealth belongs to retirees and 40% to wage-earners.

## B.7 Other

### B.7.1 Unit of observation

Our unit of observation is the adult, i.e., we allocate total national income to all US residents aged over 20.<sup>22</sup> This seems preferable to possible alternatives.

First, while national accounts are often expressed per capita (and Kuznets 1953 used per capita unit to compute his top income shares), children do not directly earn income, so it makes more sense not to include them when we look at the distribution of pre-tax income. Second, the family tax unit (as used, e.g., in Piketty and Saez 2003, and Saez and Zucman 2016) has the problem that it is sensitive to the fraction of the population that is married, which can affect growth trends: fewer marriages lead to lower income growth per family.<sup>23</sup> Third, survey-based inequality studies often use equivalence scales (such as income divided by the square root of the number of household members). In effect, this would lead to series that would be intermediate between tax-unit series and our equal-split individual series. However, these equivalence scales are somewhat arbitrary and introduce non-linearities in our growth decomposition exercises that are difficult to justify.

In our view, however, statistics based on equal-split adults, tax units, or individualized adults all have their merits and shed valuable light on income concentration and its evolution. There is a long tradition of computing inequality across households, which are conceptually close to tax units.<sup>24</sup> However, because the size of households changes over time, inequality between

---

<sup>22</sup>We include the institutionalized population in our base population. This includes prison inmates (about 1% of adult population in the US); population living in old age institutions (about 0.6% of adult population) and mental institutions; and the homeless population. Institutionalized population is generally not covered by surveys, so BEA removes income of institutionalized households from NIPA aggregates to construct their distributional accounts (see Furlong, 2014). We prefer to take everybody into account and allocate zero or small pre-tax income to institutionalized households (see on-line files).

<sup>23</sup>Also note that over the 1913-1947 period the US income tax system was de facto individualized as couples where both spouses had income were better off filing separate returns. In other words looking at tax units can bias a number of historical evolutions (as well as international comparisons).

<sup>24</sup>A household can include several tax units like two adult roommates sharing meals, or a grandparent living



households can rise or fall for purely demographic reasons. In the United States, the number of households has been growing faster than the number of adults over the last decades, because of the decline of marriage and the rise of single-headed households. Computing inequality across equal-split adults neutralizes this demographic trend and, as Appendix Figure S.15b shows, leads to a smaller increase in inequality than computing inequality across tax units. To compare inequality over time, using the equal-split adult as unit of observation is therefore a meaningful benchmark, as it abstracts from confounding trends in household size and gender inequality. There is no silver bullet, however. To measure the inequality of living standards in the cross-section, one might want to use the household unit, maybe with adjustments to capture economies of scale within the household as done for example in the US Congressional Budget Office (2016) official statistics.<sup>25</sup> To measure the inequality of monetary power, one might favor fully individualized series—where each spouse is assigned her own income—such as those discussed in Section 5 of the main paper. None of these approaches alone offers a comprehensive view; all provide valuable vantage points on the current evolutions of income inequality and can be studied using our distributional national accounts.

### B.7.2 Price deflator

We use the national income deflator (NID), i.e. the price index for net national product (from NIPA table 1.7.4). It is better to use the NID than the Consumer Price Index (CPI) for two reasons. First, the NID corresponds to what we conceptually want to use, as we are interested in overall national income, and not only consumption. Second, the CPI is not chained and hence tends to over-state inflation. The CPI under-states real growth by 0.5% per year as compared to NID, which has large effects over long periods.<sup>26</sup> The NID closely follows the GDP (and GNP) deflators, which are also chained. The CPI deflator misses investment (whose price increases less than consumption, especially once quality improvements have been accounted for) and public consumption. In the US, the price of public goods seems to have grown a bit more than the price of private goods.

---

with her kid and grandkids (see US Census Bureau, 2016 for the exact definition of households).

<sup>25</sup>Equal-split series under-estimate economies of scale within the household. John who earns \$10,000 gets the same income as Felix and Maria who as a couple earn \$20,000 in total, while in reality John probably has a lower living standards due to economies of scale—it may be harder for him, for instance, to pay his rent. Household (or tax-unit)-based series, in contrast, over-estimate economies of scale, as Felix and Maria count as one unit, just as Felix. The right equivalence scale probably lies in between the tax unit and the equal-split adult.

<sup>26</sup>The Census Bureau still uses the CPI for official distributional and poverty statistics (US Census, 2016) as Piketty and Saez (2003) do, thus under-stating real growth compared to what national accounts aggregates show.

## C Comparison with Other Studies

### C.1 Comparison with Auten and Splinter (2017)

In a recent and valuable paper, Auten and Splinter (July 2017) (hereafter AS) compute top income shares (for the top decile, top percentile, and top .1%) since 1960 using the same individual tax data source as we do. They start from the fiscal income definition used in Piketty and Saez (2003) and broaden this income definition by adding specific components but without matching (at least not yet) total national income. They consider (1) AS market income (similar to our pre-tax income definition but including only about 76% of National income in 2013), (2) AS pre-tax income (their preferred series, not the same as our pre-tax income) which is AS market income plus government individualized transfers, and (3) AS post-tax income defined as AS pre-tax and subtracting all taxes.

The AS series show substantially lower top income shares in recent years and much less increase in top income shares in recent decades. The AS pre-tax income series show almost no increase in the top 1% income share from 11.0% in 1960 to 12.2% in 2013 although their series still show a 56% increase from 7.8% in 1980 to 12.2% in 2013. In contrast, our pre-tax series show an increase from 12.6% in 1960 to 19.6% in 2013 and a 83% increase from 10.7% in 1980 to 19.6% in 2013. In this appendix, we explain the key reasons behind the discrepancies outlining the largest factors. We thank Gerald Auten and David Splinter for having generously shared their series and answered our queries to make this comparison possible. Both our work and theirs have benefitted from mutually constructive discussions. This online appendix will be revised when AS finalize their series.

The key methodological difference between AS and our paper is that we systematically match national income, component by component, while they start from fiscal income and add some but not all tax-exempt income. Currently (i.e., in the July 2017 version of their paper), their pre-tax and pre-transfer “consistent market income” series capture only about three quarters of national income in recent years. We believe that starting from the national accounts is preferable, because it allows by construction to match 100% of national income. By contrast, the bottom-up approach of AS is likely to always retain a large residual discrepancy with national income, particularly for capital income where the gap between fiscal income and national income is largest. This gap arises for many reasons: (a) capital income earned indirectly through pension and insurance funds, undistributed income from trusts, fiduciaries, etc., (b) under-reporting due to tax evasion, (c) a large number of differences in definitions and concepts between fiscal income

as defined by the IRS and national income as internationally defined by the System of National Accounts (different definitions of what constitutes costs deductible from income, of depreciation, of bad debt expense; timing adjustments, etc.) We do not claim that the national accounts are perfect—indeed one of our goals is to contribute to the improvement of the national accounts by using them extensively and pinpointing a number of areas where improvements could be made. But we think that our approach that matches the national accounts totals component by component is likely to be most robust for our purposes of decomposing macroeconomic growth and applicable to other countries. We list below the main sources of discrepancy between our and AS series (focusing on their preferred pre-tax income series).

**Income definitions.** AS consider three income definitions.

(1) AS market income is defined as fiscal income adding retained earnings of corporations owned directly by individuals, the corporate income tax, the business property tax, imputed rents of homeowners, employer payroll taxes, and employer provided health insurance. This market income definition is about 87% of National income in 1960 and declines to about 76% of National income in 2013. Indeed, it misses significant components of income such as capital income earned on pension and insurance funds, residential property taxes, sales and indirect taxes, undistributed capital income from trusts, and a large number of definitional discrepancies between national accounts and fiscal data.<sup>27</sup>

(2) AS pre-tax income is defined as market income and adding government individualized transfers (social security benefits, unemployment insurance benefits, medicare, medicaid, SNAP, etc.). This is AS preferred income definition. In our view and as we discussed in the main text, adding government transfers (funded by taxes) but without subtracting taxes is not conceptually fully meaningful. As individualized transfers have grown by more than 10 points of national income since 1960, using this definition depresses the top 1% income share increase by about 1.5 points since 1960.

(3) AS after-tax income which is AS pre-tax income minus all taxes.<sup>28</sup>

**Pensions and corporate tax.** AS treat pensions on a distribution basis, disregarding the large flow of capital income earned by pension funds that is not distributed as pensions to current retirees. As a result, their series distribute less than 50% of the macro flow of corporate

---

<sup>27</sup>AS are planning to revise their series to include some of these components.

<sup>28</sup>They currently deduct sales and excise taxes and residential property taxes that are not included in income in the first place so that they effectively deduct these taxes twice relative to national income conventions.

retained earnings (as more than 50% of it goes to pension funds), while we distribute 100% of the macro flow of retained earnings (to individuals who own equity wealth directly and through pension funds).<sup>29</sup> AS also allocate all of the corporate tax paid by pension funds to current retirees, who are typically relatively low income. As a result, the top 1% pays only 17% of the corporate tax in 2013 in their series (down from 33% in 1960) while in our estimates, the top 1% pays about 45% of the corporate tax (with no significant downward trend). AS assumption that the corporate tax is regressive because it is borne mostly by low income retirees is non standard and does not seem the most realistic to us. CBO official numbers for 2013 assume that the top 1% pays 47% of the corporate tax, very close to our 45% number. Treasury and JCT have comparable estimates where the top 1% pays more than 40% of the corporate tax, again in line with our estimates. With a more standard incidence assumption, AS top 1% income share would increase by about 1 point in 2013 and would barely rise in 1960.

**Under-reported income due to tax evasion.** The national accounts make an explicit allowance for misreported taxable income for two income categories: unincorporated business profits (\$538 billion in 2013) and wages (\$80 billion in 2013). In total, this unreported taxable income adds up to 4.3 points of national income in 2013. We distribute unreported wages proportionally to reported wages, and unreported business profits proportionally to reported business profits (just as we do for other residual discrepancies between fiscal and national income). By contrast, AS plan to distribute unreported taxable income proportionally to overall taxable income. Because most of the misreported taxable income is business income, and business income is more concentrated than overall income, this would depress the top 1% income share by about 1.2 point compared to our estimates.<sup>30</sup> AS argue that their allocation is preferable because it is consistent with the distribution of tax evasion estimated by IRS random audit studies (see Johns and Slemrod, 2010). We believe, however, that our treatment is better justified, because random audit data severely under-estimate tax evasion at the top-end of the distribution. Random audits are able to uncover unreported self-employment income (and more broadly relatively simple and coarse forms of tax evasion) but fail to capture sophisticated forms of evasion involving legal and financial intermediaries, the detection of which would require much more resources than available to tax authorities for their random audit programs. We refer to

---

<sup>29</sup>In 2013, they distribute only 37% of retained earnings in national income (\$236bn out of the \$635bn) while in 1960 they distribute close to 100% of retained earnings.

<sup>30</sup>In our estimates, 47% of business income and 11% of wages is earned by the top 1%. Hence, we allocate \$262bn of under-reported income to the top 1% ( $=538 \times 47\% + 80 \times 11\%$ ) while AS will attribute only about \$93bn  $= 15\% \times (538+80)$  (as the top 1% income share of AGI is about 15%). Hence, we attribute \$169bn extra to the top 1%, or 1.17% of the \$14.445tr national income in 2013.

the recent paper by Alstadsæter, Johannesen and Zucman (2017) for extensive evidence of this key issue. In the U.S. context, a particularly acute problem is that 30% of partnership income (which is highly concentrated) is not traceable on individual tax returns (Cooper et al., 2016). In future research we plan to improve U.S. estimates of the tax gap and its distribution, and we will include the results of this research in future updates of our distributional national accounts.

**Various other income discrepancies.** There are many discrepancies between fiscal and national income beyond tax evasion, due to indirect ownership (for example, trusts do not distribute all their income to individuals), differences in definitions, timing, etc. AS for now ignore these discrepancies while we always distribute them component by component (i.e., for wage income proportionally to reported wages, etc.). As these discrepancies are most sizable for capital income (which is highly concentrated), they increase the top 1% income share by around 1 point.

**Ranking of individuals.** When we compute a top 1% income share, we naturally rank all adults by the relevant income definition and select the top 1% highest earners. In contrast, AS define the top 1% by ranking individuals based on their *total family* income (i.e., summing the incomes of married spouses) while they define income based on *average family* income (i.e., averaging the incomes of married spouses). Hence, if the top 1% income threshold is \$400K. A couple with family income of \$500K would count as two adults in the top 1% each assigned \$250K. In contrast, a single individual with income of \$350K would not be assigned in the top 1% yet would be assigned an income of \$350K higher than the incomes of the couple. This has a large downward effect on the top 1% income of 1.8 points in 2013 (AS, Table 1) relative to the Piketty and Saez (2003) series based on tax units. But it has a more modest downward effect of about 0.3 points relative to our benchmark series based on individual adults (with equal split among couples) as the majority of top 1% earners are married.<sup>31</sup>

Summing these various effects closes most of the gap between our pre-tax income series and their preferred pre-tax series. AS are planning to further revise their series. We hope they will consider adding fully pre-tax and pre-transfers series that match national income as closely as possible so that we can do a more systematic side by side comparison. We will accordingly update this appendix when their series are finalized.

---

<sup>31</sup>As married filers qualify for the top 1% based on their family income, the AS method greatly increases the fraction married in the top 1%, from around 78% in our series to about 92% in AS series in recent years.

## C.2 Comparison with Smith et al. (2017)

We have conducted a detailed reconciliation of our findings with the important contribution by Smith et al. (2017). It is important to note at the outset that Smith et al. focus mostly on a person-level analysis: how does the typical person in the top 1% or 0.1% earn his or her income. We, in contrast, focus on income distribution: how much of national income accrue to the top 1% or 0.1% and in what forms. We therefore focus here on the implications of Smith et al.'s findings for our question of interest.

In contrast to Smith et al. (2017), we find that capital income is overall the key driver of the rise of the top 1% income share since 2000. Our conclusions differ from Smith et al.'s (2017) for the following reasons:

**Missing tax-exempt capital income.** Smith et al. (2017) use individual income tax data (and exclude realized capital gains) to study the capital/labor split. While their approach does not require assumptions on the allocation of retained earnings and other income that does not appear on individual income tax returns, their focus on directly observed ordinary capital income (excluding realized capital gains) misses two-thirds of economic capital income. By definition, pass-through business income shows up on individual income tax data, but a great deal of pure capital income does not: corporate retained earnings; imputed rents; capital income paid to pension funds; dividends and interest retained in trusts, estates, fiduciaries; etc. Since the vast majority of capital income does not show up as ordinary capital income on individual tax returns, the capital/labor split cannot in our view be studied compellingly with individual income tax data only.

**Rising retained earnings since 2000.** An important macroeconomic development since 2000 has been the rise of retained earnings of C-corporations, which are concentrated at the top. Smith et al. (2017) miss all corporate retained earnings (because these earnings are not visible in individual income tax data), while we include retained earnings (along with all other forms of tax-exempt capital and labor income) in our analysis. The macro flow of retained earnings amounts to 5.1% of national income (on average over 2010–2014), twice as much as the flow of S-corporation profits (2.6% of national income) that Smith et al. (2017) focus on. It is the rise of C-corporation retained earnings that has been the main driver of the rise in the top 1% pre-tax income share since 2000. Out of the 1.9 points increase in the top 1% share since 2000, we find that 1.2 points owe to the rise of C-corporation retained earnings.

**Tax-induced shifting goes in conflicting directions.** We agree that our capital/labor split can be affected by tax-induced income shifting across bases, but the net effect of all forms of shifting could actually lead us to *understate* the capital share in recent years rather than overstate it. Smith et al. (2017) focus on one specific type of shifting: in S-corporations with active owners only, owners have an incentive to pay themselves artificially low wages so as to avoid Medicare taxes of 3.8%. But other forms of shifting go in the opposite direction. In all businesses with passive owners (i.e., as long as there is at least one outside non-active investor), active owners have incentives to pay themselves high wages, as any residual pass-through business profit is split across all owners including passive investors. Moreover, in all closely-held C corporations, owners have incentives to pay themselves high wages too, because the top marginal labor income tax rate (39.6% + 3.8% Medicare taxes since 2013) is below the top rate on distributed profits (tax rate over 50% when combining the 35% corporate income tax, the 20% tax on dividends and realized capital gains and the 3.8% investment income tax). Taking into account all forms of tax-induced shifting, it is therefore not clear whether we overstate or under-state the capital share. In this context, we believe it is more justified to start by following the standard national accounts labor/capital split (which is also followed in other countries), instead of selectively correcting it for only one form of shifting but not others.<sup>32</sup>

**Rising incentives to report labor income.** Because incentives to avoid individual income taxes on labor income were much higher in the 1970s–1980s than today, it is likely that we actually under-estimate labor income at the top in the 1970s–1980s, hence *under-estimate* the rise of the capital income share at the top over recent decades. Before the Tax Reform Act of 1986, top marginal rates on labor income and distributed profits were both high (70% or more up to 1981), so that owners of closely held businesses were organized as C-corporations and had incentives to pay themselves low wages, low dividends, retain earnings and either consume within firms or eventually sell and realize capital gains. Part of the high retained earnings we observe from 1960 to 1985 (3.8% of national income on average) therefore probably correspond to disguised labor income of business owners and should be subtracted from our capital income share.

**Ranking and definition of income.** Smith et al. (2017) rank tax units by their fiscal income excluding capital gains. By contrast, we rank equal-split adults by their pre-tax national income.

---

<sup>32</sup>Note that we always split mixed non-corporate business income (which includes all partnership income) 70% labor and 30% capital, so any tax-induced shifting that is stable over time does not affect our estimates.

More precisely, we rank adults as follows. We compute equity wealth by capitalizing positive S-corporation income, and we compute C-corporation wealth by capitalizing dividends and positive realized capital gains. We then apply the same pre-tax rate of return to S-corporation and C-corporation wealth, namely the macroeconomic average pre-tax return on equity wealth; this is a pre-tax return gross of any tax, in particular gross of the corporate income tax. We choose to apply the same pre-tax return to C- and S-corporation equities when ranking equal-split adults for three reasons. First, this is a natural benchmark consistent with perfect capital mobility between the S- and C-corporate sectors. Second, this is consistent with the treatment of equity income in other countries where distributional national accounts are being constructed. Third, we do not have systematic information on how pre-tax rates of returns on equities may vary across the distribution. S-corporation income earners may have higher rates of return than the average pre-tax return on equity wealth, which could lead us to under-estimate the importance of S-corporation income at the top of the income distribution. But on the other hand, wealthy C-corporation owners may have higher rates of return than the average pre-tax return on equity wealth, which would lead us to under-estimate the importance of C-corporation income at the top. Note that we assume that corporate taxes fall on S-corporation equities just like on C-corporation equities. If in reality a lot of S-corporation income corresponds to disguised wages, then a much smaller fraction of the corporate tax should be allocated to S-corporations, reducing the pre-tax return that should be applied to S-corporation equity.

To compute the total S-corporation vs. C-corporation income of the top 1%, we proceed as follows. We compute the total S-corporation wealth and C-corporation wealth of the top 1%. We then compute  $a$  = the post-corporate-tax C-corporation equity income of the top 1% by applying the post-corporate-tax average rate of return on C-corporation wealth (i.e., C-corporation dividends plus retained earnings divided by C-corporation wealth) to the C-corporation wealth of the top 1%. We compute  $b$  = the post-corporate-tax S-corporation equity income of the top 1% by applying the post-corporate-tax average rate of return on S-corporation wealth (i.e., S-corporation dividends divided by S-corporation wealth) to the S-corporation wealth of the top 1%. We divide the total pre-tax equity income of the top 1% across C-corporation and S-corporation proportionally to the share of  $a$  and  $b$  in  $a + b$ . Note that this procedure takes into account the fact that post-corporate-tax returns are significantly higher in S-corporations than C-corporations.



**Magnitude of the effects.** The magnitude of the various effects described above is illustrated in the Online Appendix. Figure S.34 shows that the S-corporation profits of our top 1% pre-tax earners amount to 1.6% of pre-tax national income in 2014 (up from 1.4% in 2000).<sup>33</sup> Figure S.35 shows that the flow of retained earnings of C-corporations (not considered in Smith et al. 2017) is large, and has increased a lot since the early 2000s. Figure S.36 takes the finding in Smith et al. (2017) that 54% of S-corporation profits are disguised wages, and shows that the impact on the composition of income for the top 1% is small. Figure S.37 assumes that before 1986, 10% of retained earnings were disguised wages; and Figure S.38 assumes that 25% of retained earnings were disguised wages. This increases the labor income of top 1% earners—substantially so for the second scenario. In both cases, we find that the rise of the top 1% income share has first been a labor income phenomenon in the 1980s and 1990s, and second mostly a capital income phenomenon since 2000. One would need to assume that the vast majority of retained earnings pre-1985 were disguised wages to overturn the first result. There are just not enough S-corporate profits to overturn the second result.

In the end, in DINA the top 1% pre-tax earners get about 50% of all pre-tax S-corporation income. In contrast, in the fiscal income data used by Smith et al. (2017), the top 1% fiscal income earners have about 70% of all S-corporation income. The difference owes to the different ranking of the units of observation and the different definition of income: Smith et al. (2017) look at observable S-corporation dividends, distributed across tax units ranked by fiscal income excluding capital gains. We look at pre-tax S-corporation profits (gross of imputed corporate and other taxes), distributed across equal-split adults ranked by pre-tax national income.

---

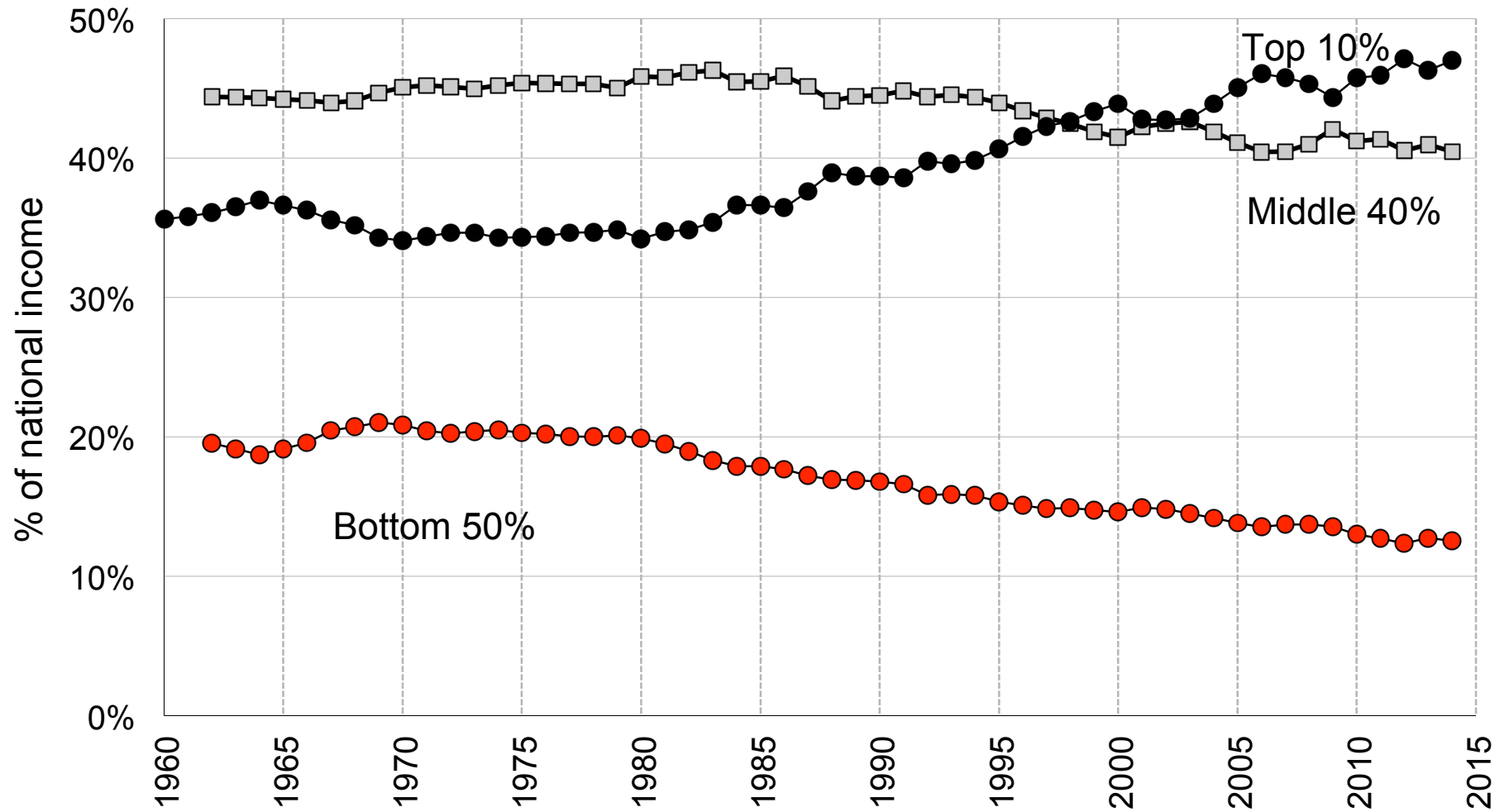
<sup>33</sup>The Piketty-Saez (2003) series where tax units are ranked by taxable income excluding capital gains (as used by Smith et al. 2017) over-state the importance of S-corporation income for the top 1%, for two reasons. First, they miss 2/3 of the flows of economic capital income not visible on tax returns. Second, ranking tax units by taxable income excluding capital gains removes by construction many high capital income earners from the top 1%.

## Appendix References

- Alstadsæter, Annette, Niels Johannesen and Gabriel Zucman.** 2017. “Tax Evasion and Inequality”, NBER working paper No. 23772.
- Alvaredo, Facundo, Anthony Atkinson, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman.** 2011-present. *The World Wealth and Incomes Database*. online at <http://www.wid.world>.
- Alvaredo, Facundo, Anthony Atkinson, Thomas Piketty, and Emmanuel Saez, and Gabriel Zucman.** 2016. “Distributional National Accounts (DINA) Guidelines: Concepts and Methods used in the World Wealth and Income Database.” World Wealth and Income Database document online at <http://www.wid.world>.
- Auten, Gerald and David Splinter.** 2017. “Income Inequality in the United States: Using Tax Data to Measure Long-term Trends,” U.S. Department of the Treasury, Office of Tax Analysis, unpublished mimeo.
- Cooper, Michael, John McClelland, James Pearce, Richard Prinszano, Joseph Sullivan, Danny Yagan, Owen Zidar, and Eric Zwick.** 2016. “Business in the United States: Who Owns It, and How Much Tax Do They Pay?,” *Tax Policy and the Economy* 30, 91–128.
- Fack, Gabrielle.** 2006. “Are Housing Benefits an Effective Way to Redistribute Income? Evidence From a Natural Experiment in France,” *Labour Economics* 13(6), 747–771.
- Fullerton, Don and Gilbert E. Metcalf.** 2002. “Tax incidence,” in: A. J. Auerbach and M. Feldstein (eds.), *Handbook of Public Economics*, Volume 4, chapter 26, 1787-1872 (North-Holland: Amsterdam).
- Goldsmith, Raymond, Dorothy Brady, and Horst Mendershausen.** 1956. *A Study of Saving in the United States*, Volume III, (Princeton University Press: Princeton).
- Hamermesh, Daniel S.** 1996. *Labor Demand*, (Princeton University Press: Princeton, NJ).
- Harberger, Arnold C.** 1962. “The Incidence of the Corporation Income Tax,” *Journal of Political Economy*, 70(3), 215–240.
- Johns, Andrew and Joel Slemrod.** 2010. “The Distribution of Income Tax Noncompliance,” *National Tax Journal* 63(3), 397–418.
- King, Willford I.** 1930. *The National Income and Its Purchasing Power*, (New York: National Bureau of Economic Research).
- Kopczuk, Wojciech and Emmanuel Saez.** 2004. “Top Wealth Shares in the United States, 1916-2000: Evidence from Estate Tax Returns,” *National Tax Journal*, 57(2), Part 2, 445-487.
- Kuznets, Simon.** 1941. *National Income and Its Composition, 1919–1938* (New York: National Bureau of Economic Research).
- Kuznets, Simon.** 1953. *Shares of Upper Income Groups in Income and Savings* (New York: National Bureau of Economic Research).
- McCulla, Stephanie H., Karin E. Moses, and Brent R. Moulton.** 2015. “The National Income and Product Accounts and the System of National Accounts 2008: Comparison and Research Plans,” *Survey of Current Business*, Bureau of Economic Analysis 95(6): 2-17.
- Meyer, Bruce, Wallace K.C. Mok, and James X. Sullivan.** 2009. “The Under-Reporting of Transfers in Household Surveys: Its Nature and Consequences,” NBER Working Paper No. 15181.

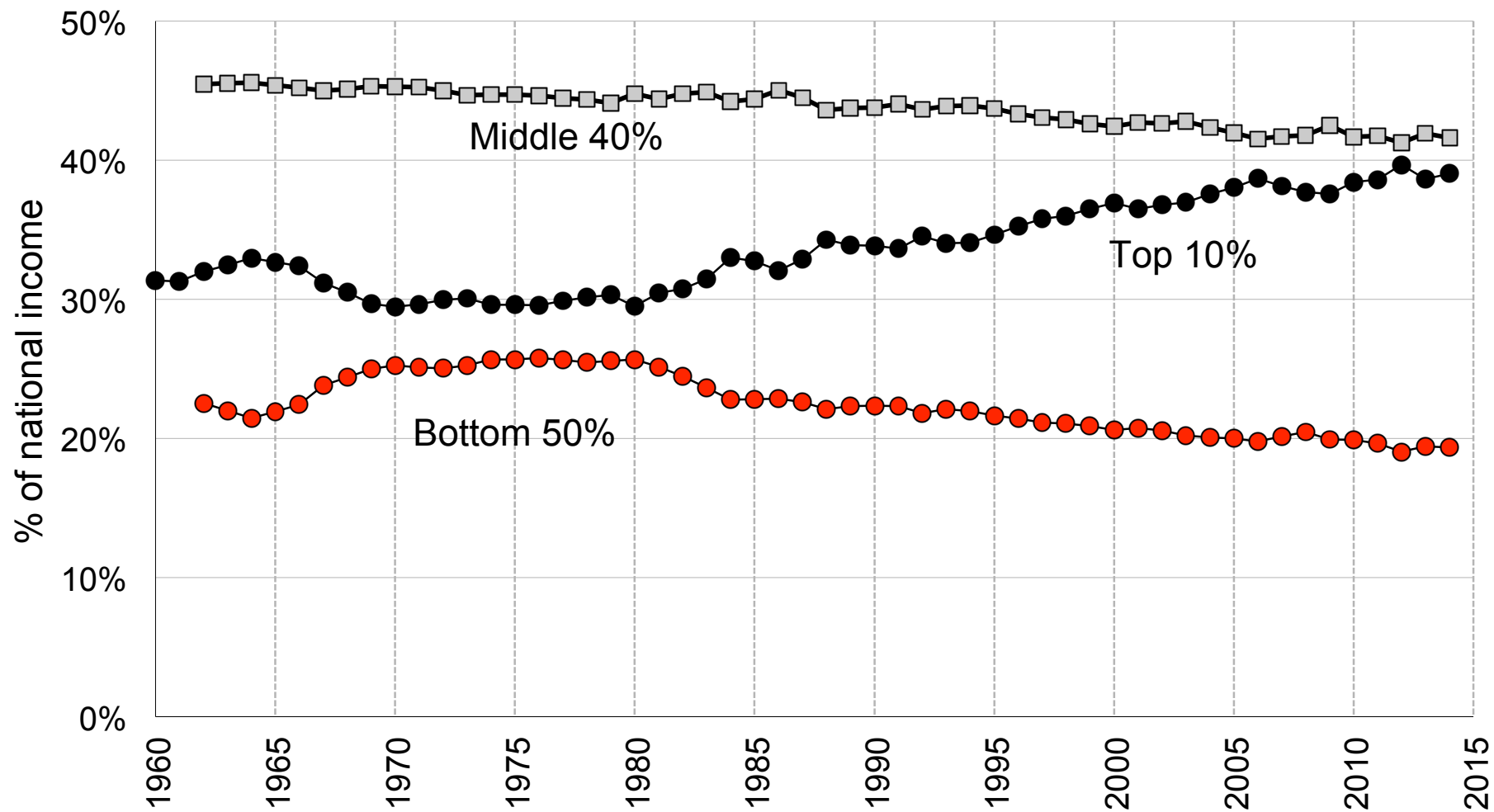
- Meyer, Bruce, Wallace K.C. Mok, and James X. Sullivan.** 2015. "Household Surveys in Crisis," *Journal of Economic Perspectives*, Fall 2015, 199-226.
- Piketty, Thomas, and Emmanuel Saez.** 2003. "Income Inequality in the United States, 1913-1998," *Quarterly Journal of Economics* 118(1), 1-39, series updated to 2015 in June 2016.
- Rohaly, Jeffrey, Adam Carasso, Mohammed Adeel Saleem.** 2005. "The Urban-Brookings Tax Policy Center Microsimulation Model: Documentation and Methodology for Version 0304", Tax Policy Center Working Paper.
- Saez, Emmanuel.** 2016. "Improving the Individual Tax Return Public Use Files for Tax and Distributional Analysis," SOI Working Paper.
- 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.
- Smetters, Kent.** 2004. "Is the Social Security Trust Fund a Store of Value ?" *American Economic Review, Papers and Proceedings*, May.
- United Nations.** 2009. "System of National Accounts 2008," European Communities, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations and World Bank. online at <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>
- US Board of Governors of the Federal Reserve System.** 2014. "Z.1 Financial Accounts of the United States , Fourth Quarter 2013." downloaded, March 6, 2014.
- US Census Bureau.** 2016. "Income and Poverty in the United States: 2015 Current Population Reports", Publication P60-252, US Census Bureau, Washington DC.
- US Congressional Budget Office.** 2016. *The Distribution of Household Income and Federal Taxes, 2013*. Washington, DC: Government Printing Press. (Annual updates available online).
- US Department of Commerce, Bureau of Economic Analysis.** 2016. *National Income and Product Accounts of the United States, 1929-2015*, (Washington, DC).
- US Treasury Department, Internal Revenue Service.** 1916-present. *Statistics of Income: Individual Income Tax Returns*, annual since 1916, Washington, D.C. online at <https://www.irs.gov/uac/soi-tax-stats-archive>
- Wolff, Edward.** 1989. "Trends in Aggregate Household Wealth in the United States, 1900-1983," *Review of Income and Wealth*, 35(1), 1-29.
- Zedlewski, Sheila and Linda Giannarelli.** 2015. "TRIM: A Tool for Social Policy Analysis", Urban Institute Working Paper.

**Figure S.1: Pre-tax income shares (equal-split individuals)**



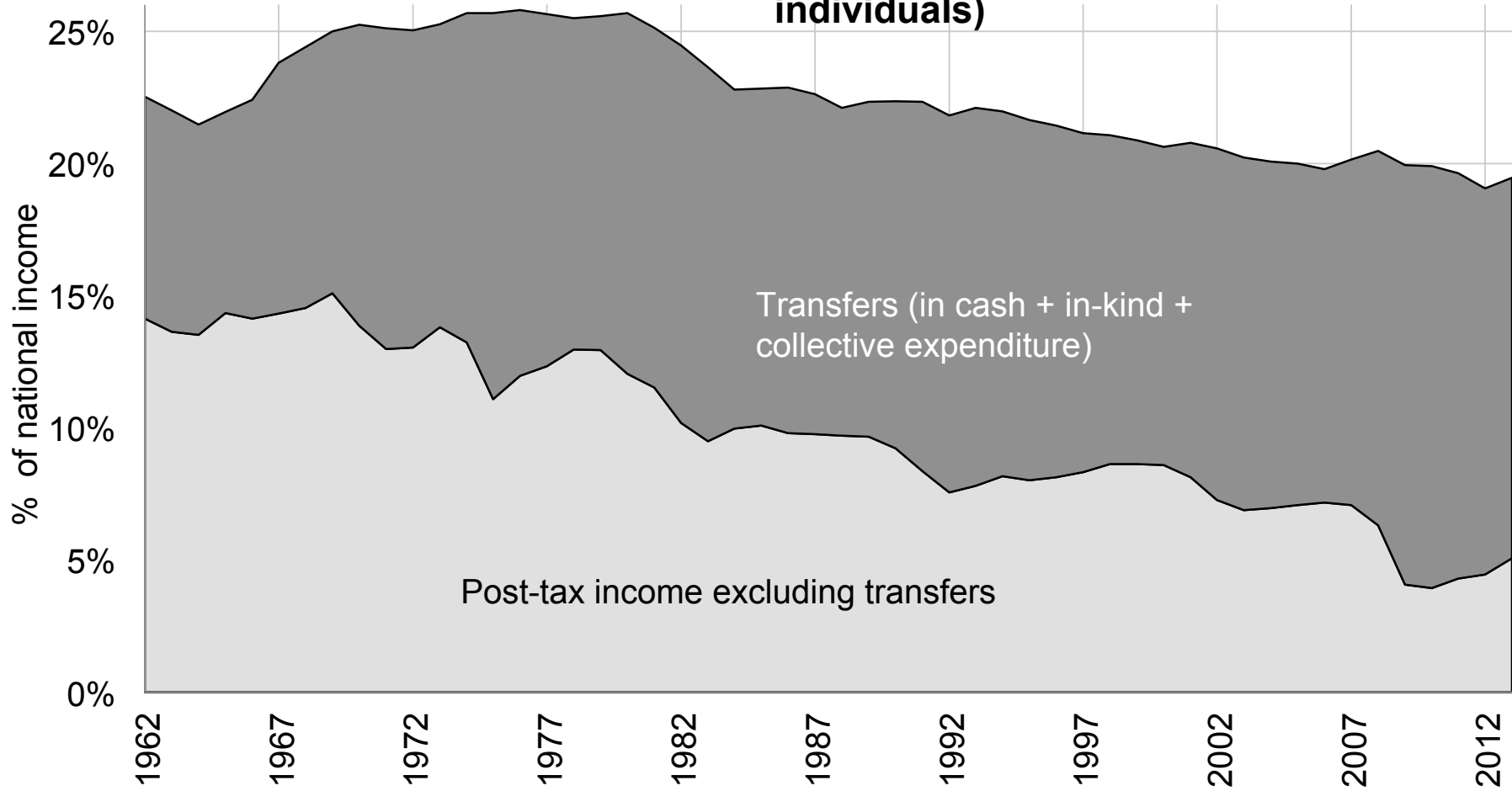
Source: Table II-B1

**Figure S.2: Post-tax income shares (equal-split individuals)**



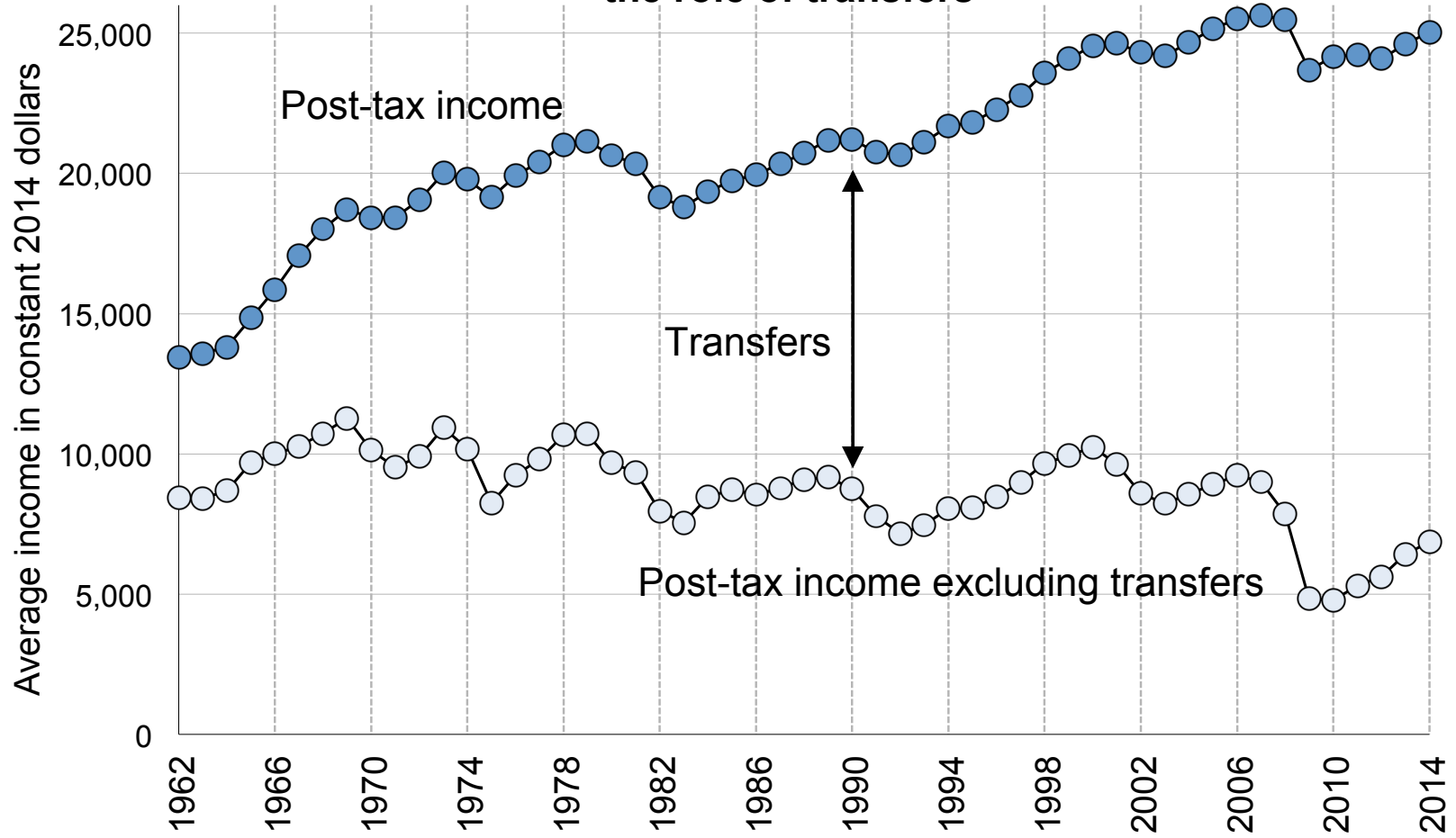
Source: Appendix Table II-C1

**Figure S.3: Post-tax income share of bottom 50% (equal split individuals)**



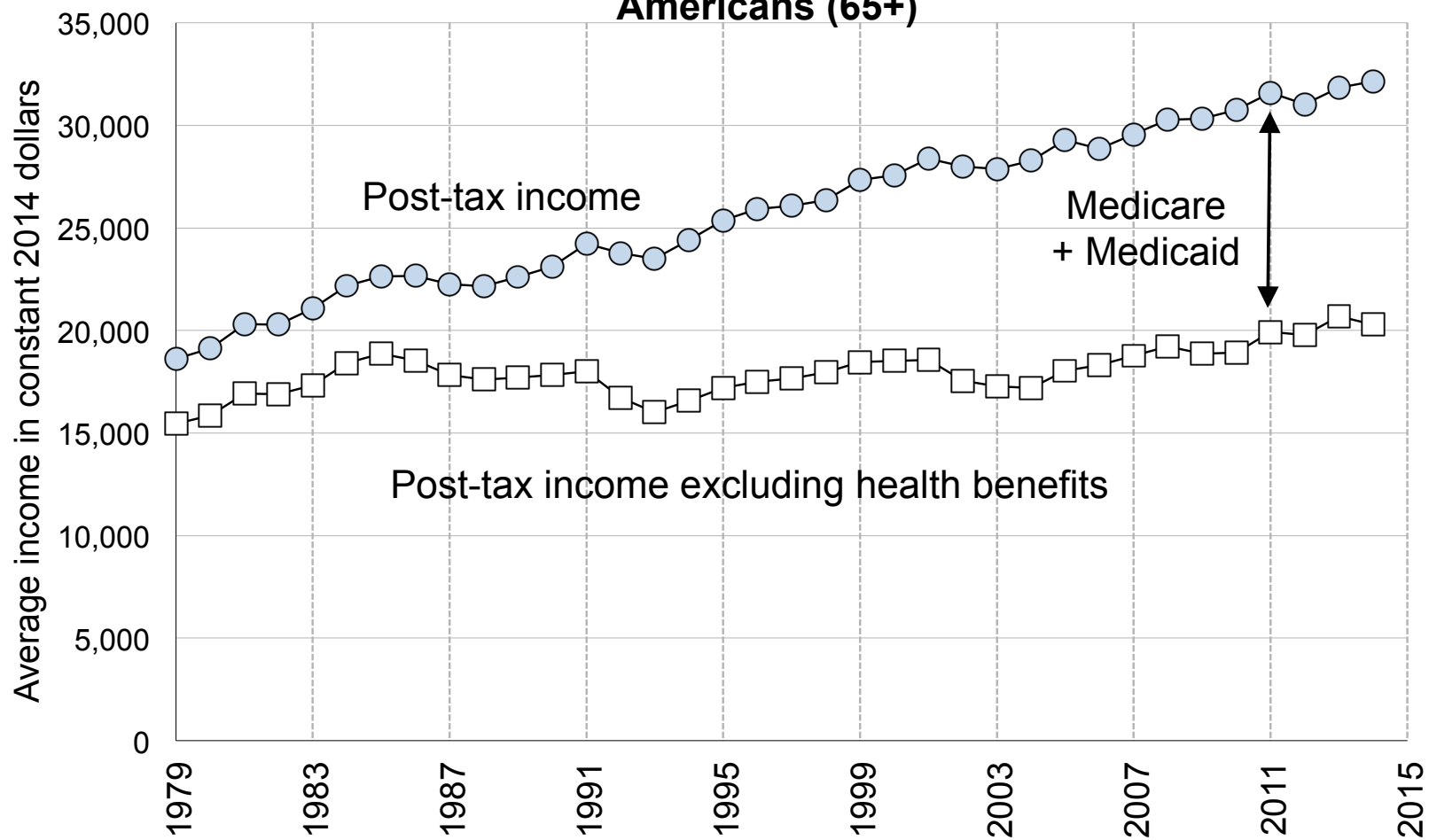
Source: Appendix Table II-C2

**Figure S.4: Real post-tax income of the bottom 50%:  
the role of transfers**



Source: Table II-C3b.

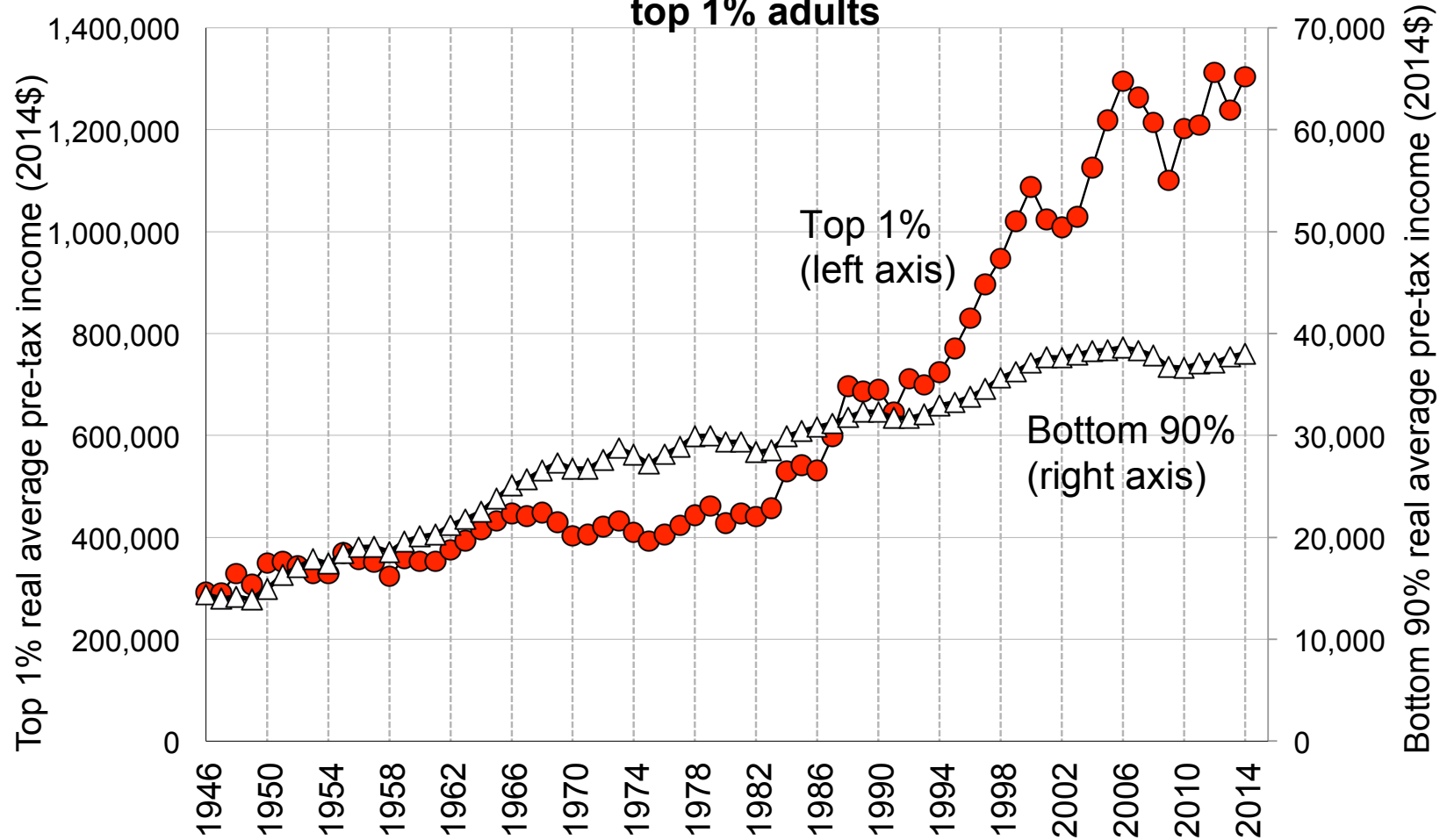
**Figure S.5: Post-tax income of the bottom 50% of elderly Americans (65+)**



Source: Appendix Table II-C7c.

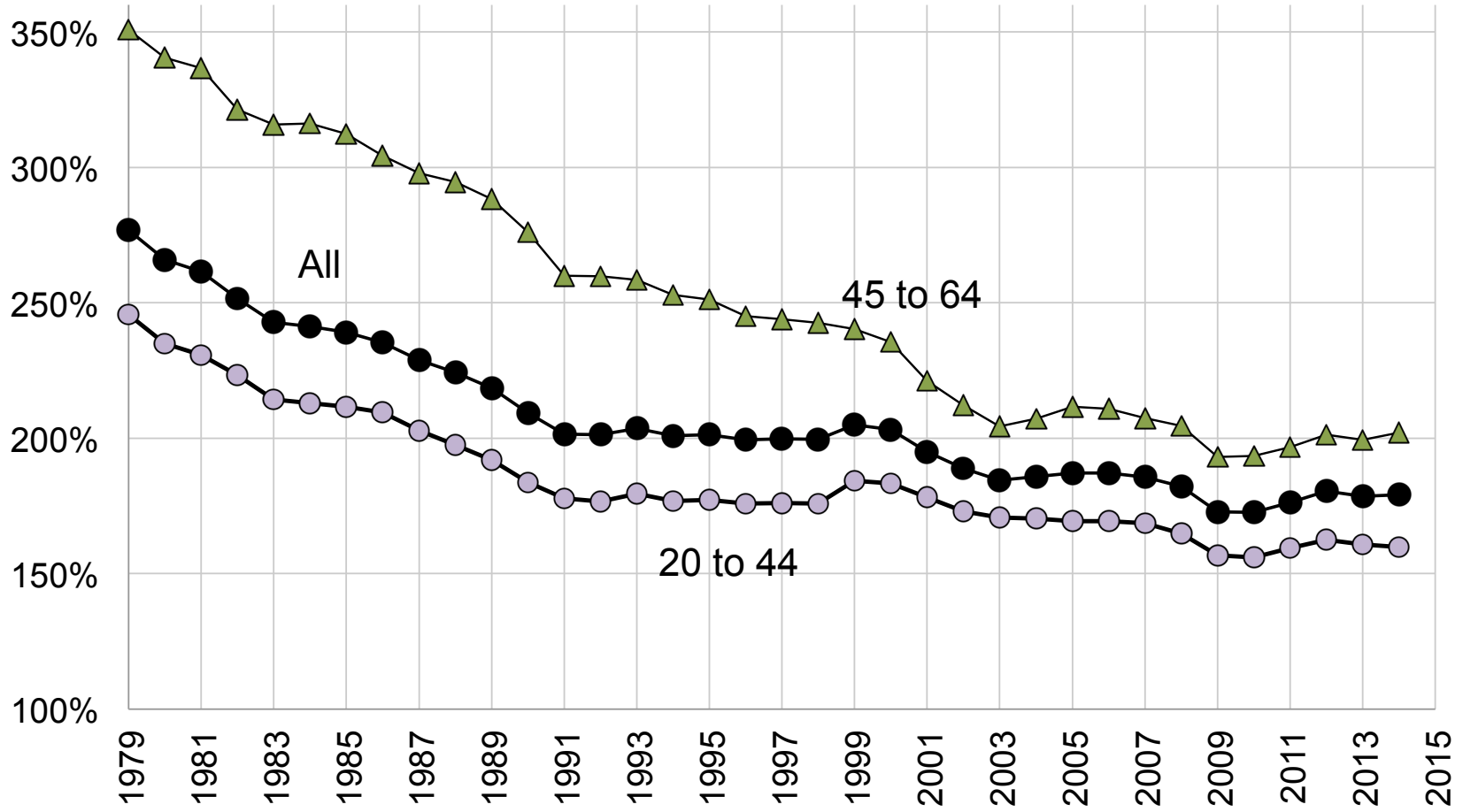


**Figure S.6: Real average pre-tax income of bottom 90% and top 1% adults**



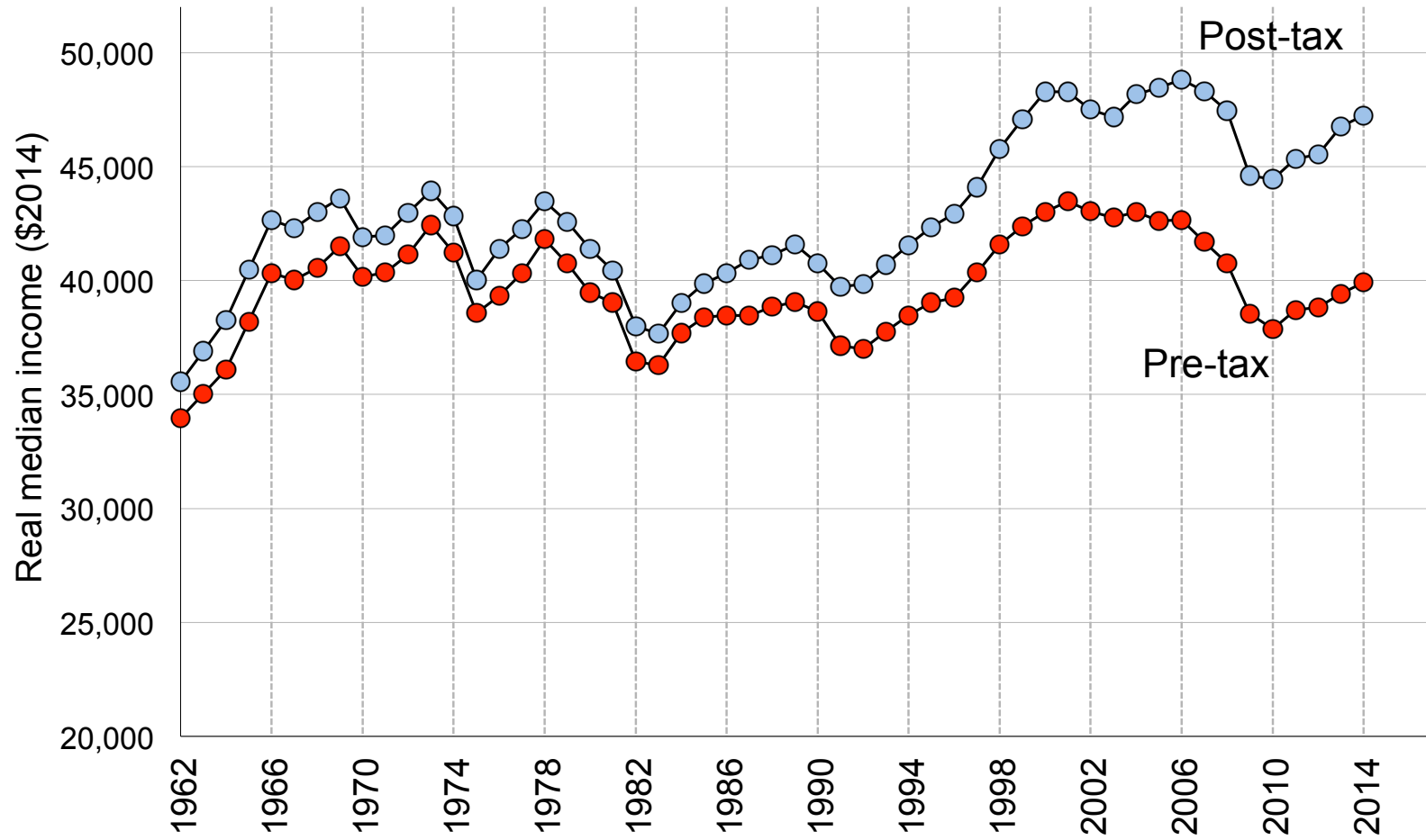
Source: Appendix Table II-B3

**Figure S.7: Average labor income of all men aged 20-64 / all women aged 20-64, by age group**



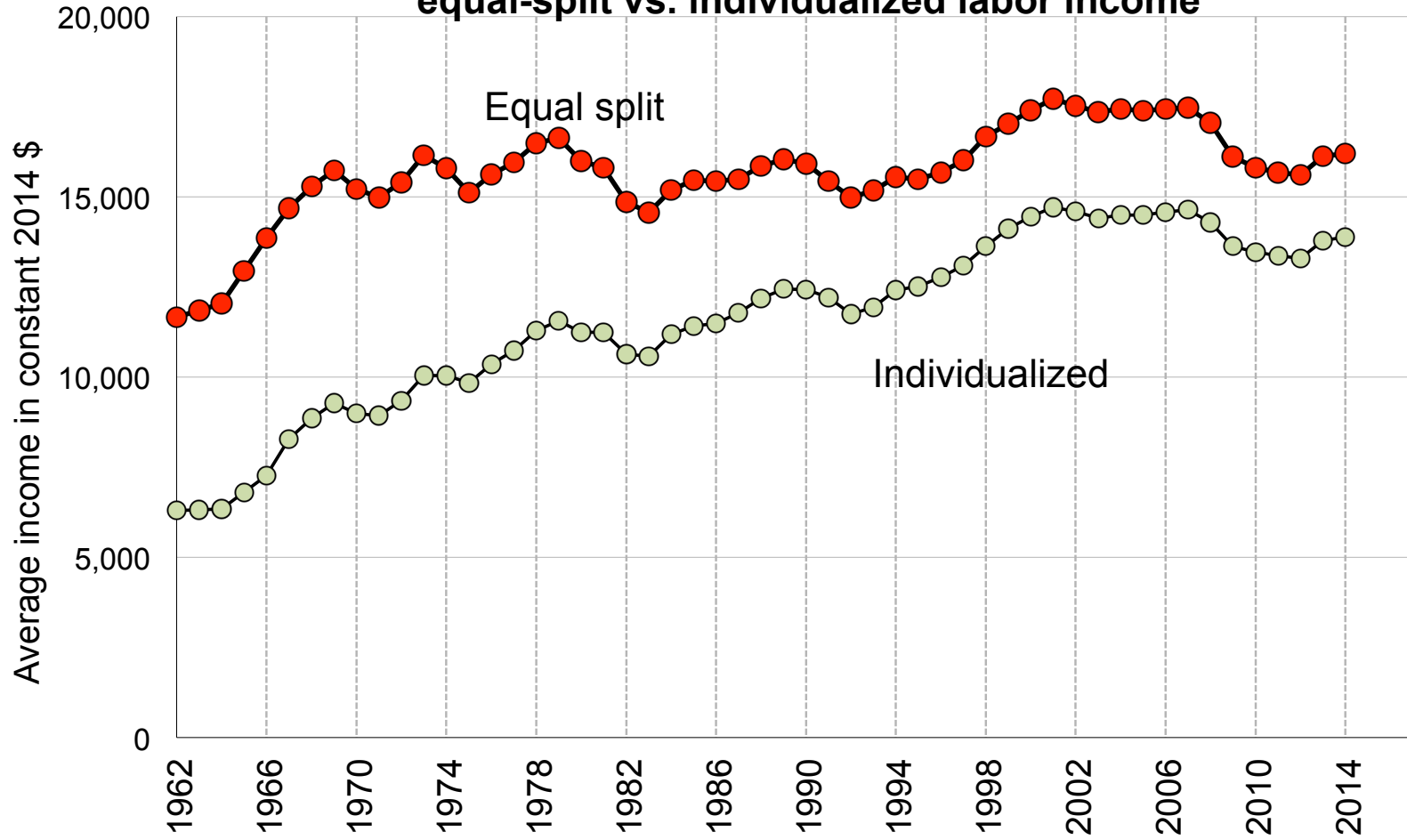
Source: Appendix Table F1.

**Figure S.8: Median pre-tax vs. post-tax income of men**



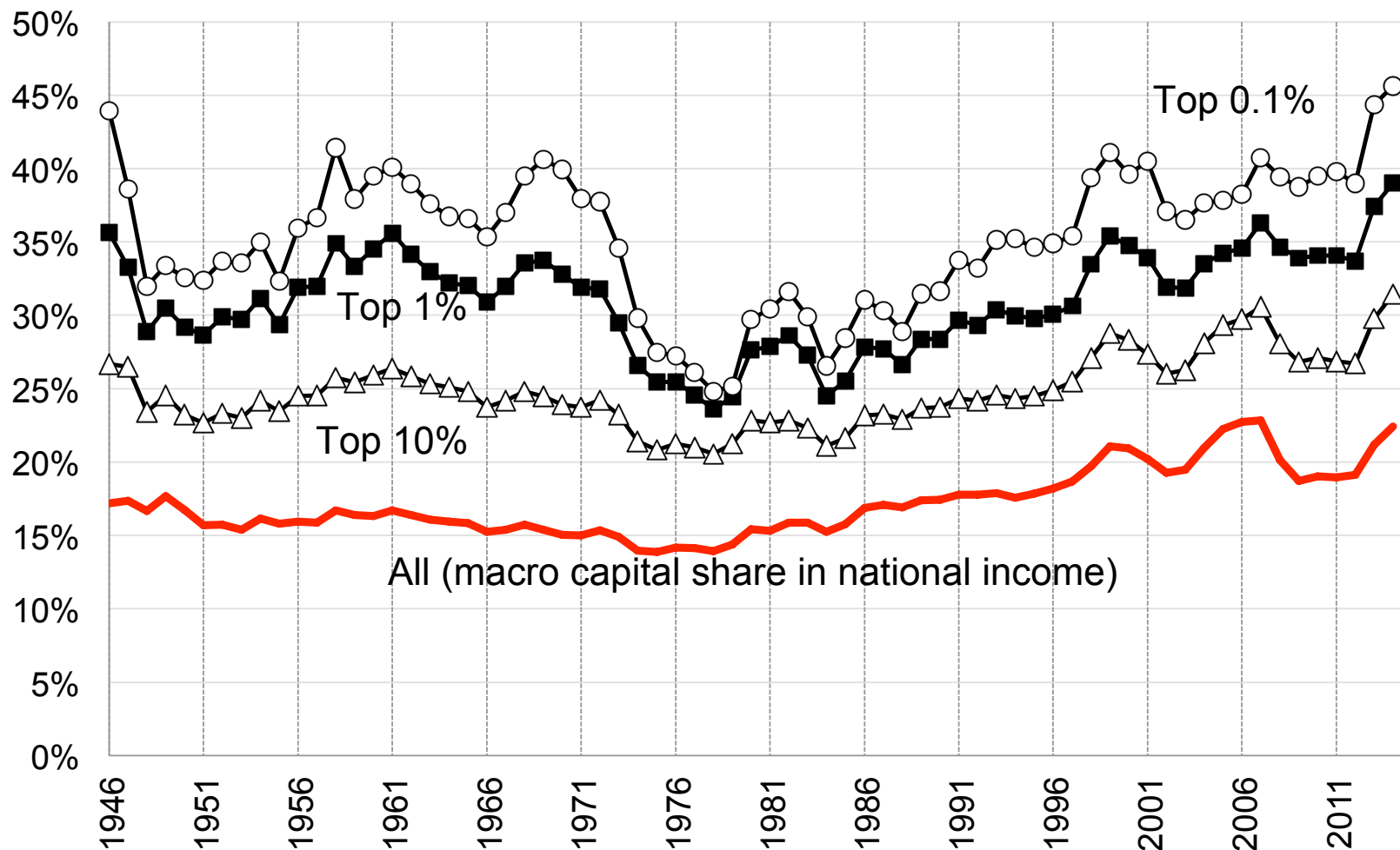
Source: Appendix Table II-B13 and II-C13.

**Figure S.9: Average income of bottom 50%:  
equal-split vs. individualized labor income**

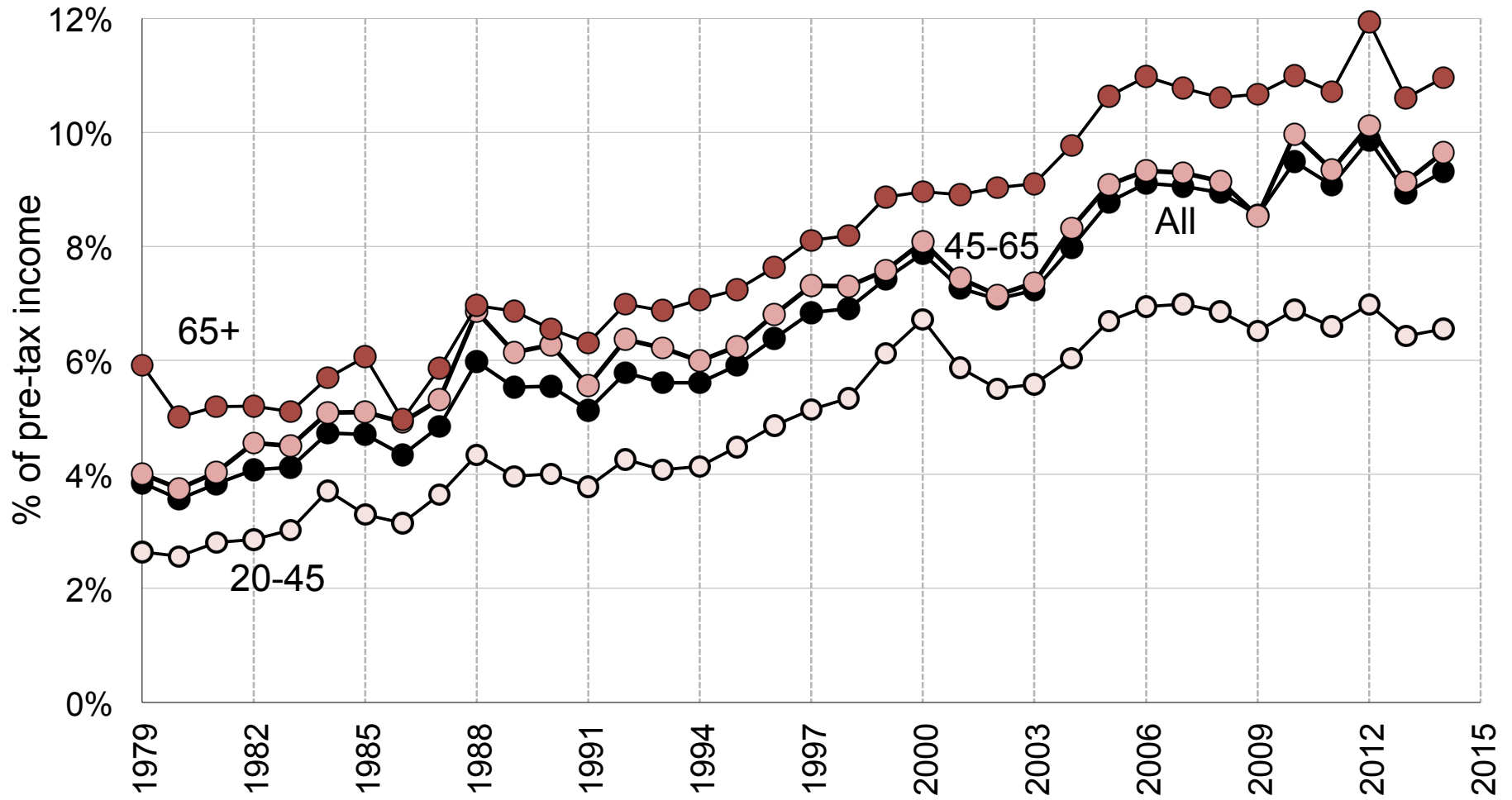


Source: Appendix Tables II-B7.

**Figure S.10: The share of capital in pre-tax income, assuming constant 5% return on capital**

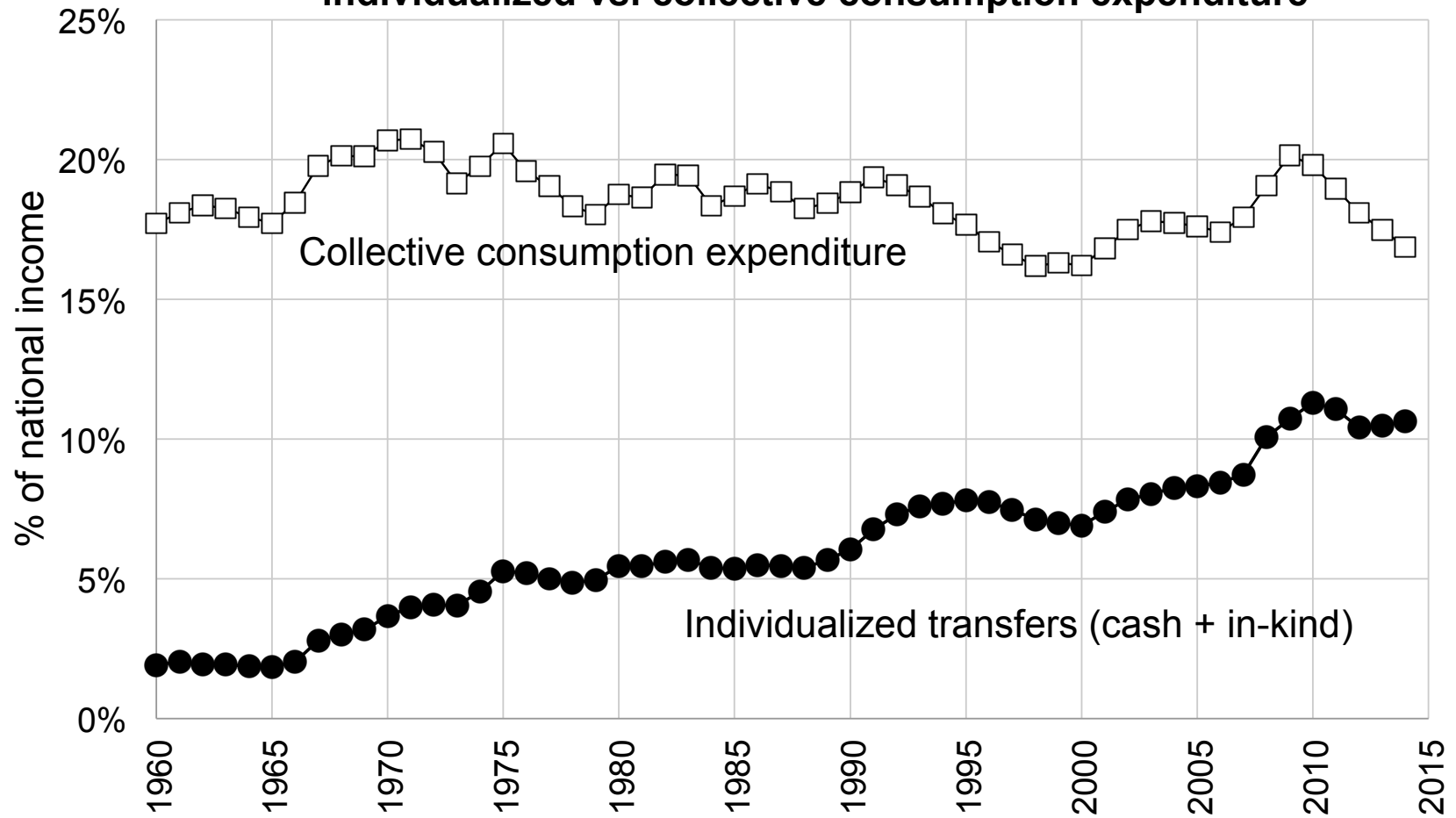


**Figure S.11: Top 0.1% pre-tax national income share by age group**



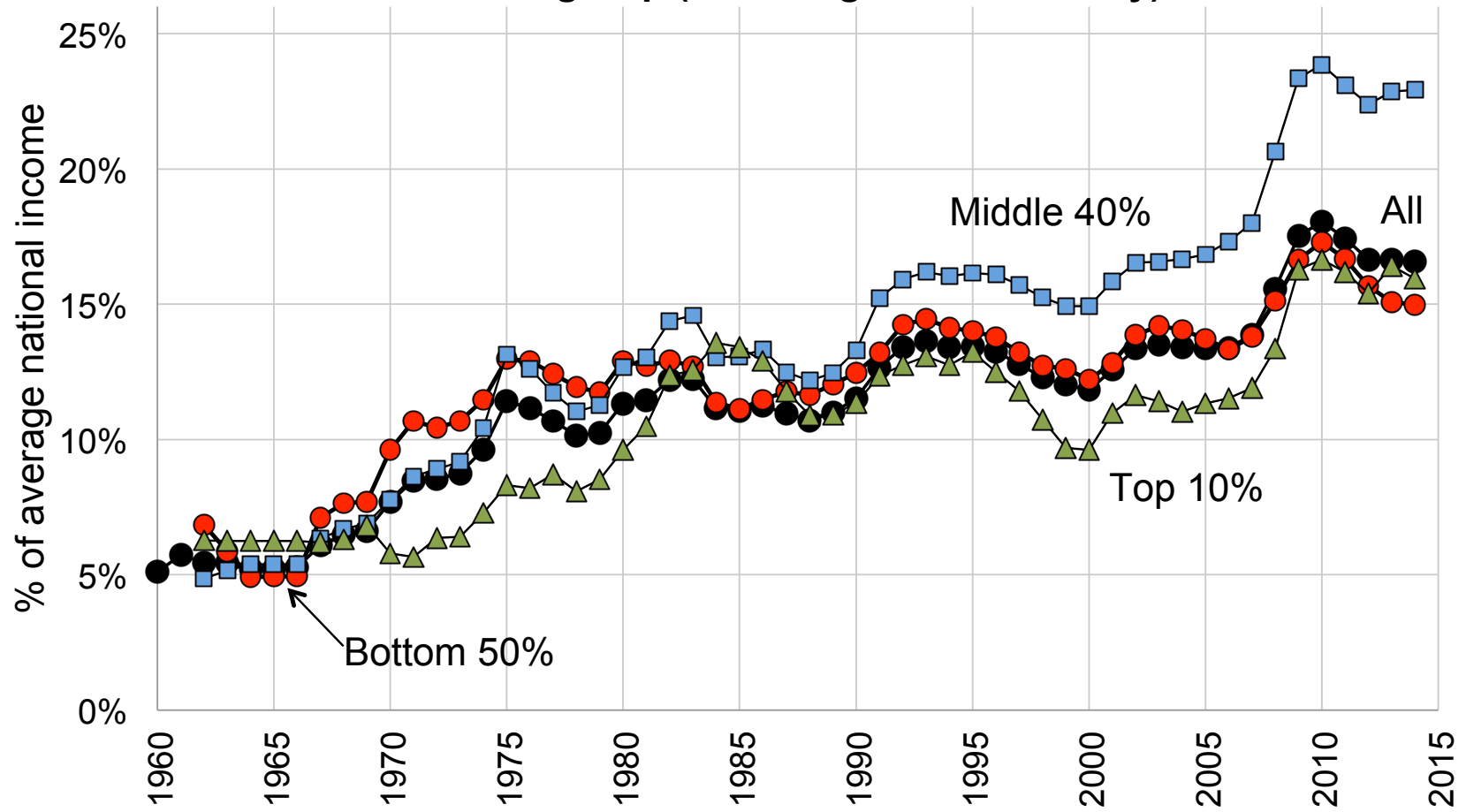
Source: Appendix Table II-B11b.

**Figure S.12: Average transfers:  
individualized vs. collective consumption expenditure**



Source: Appendix Table G4

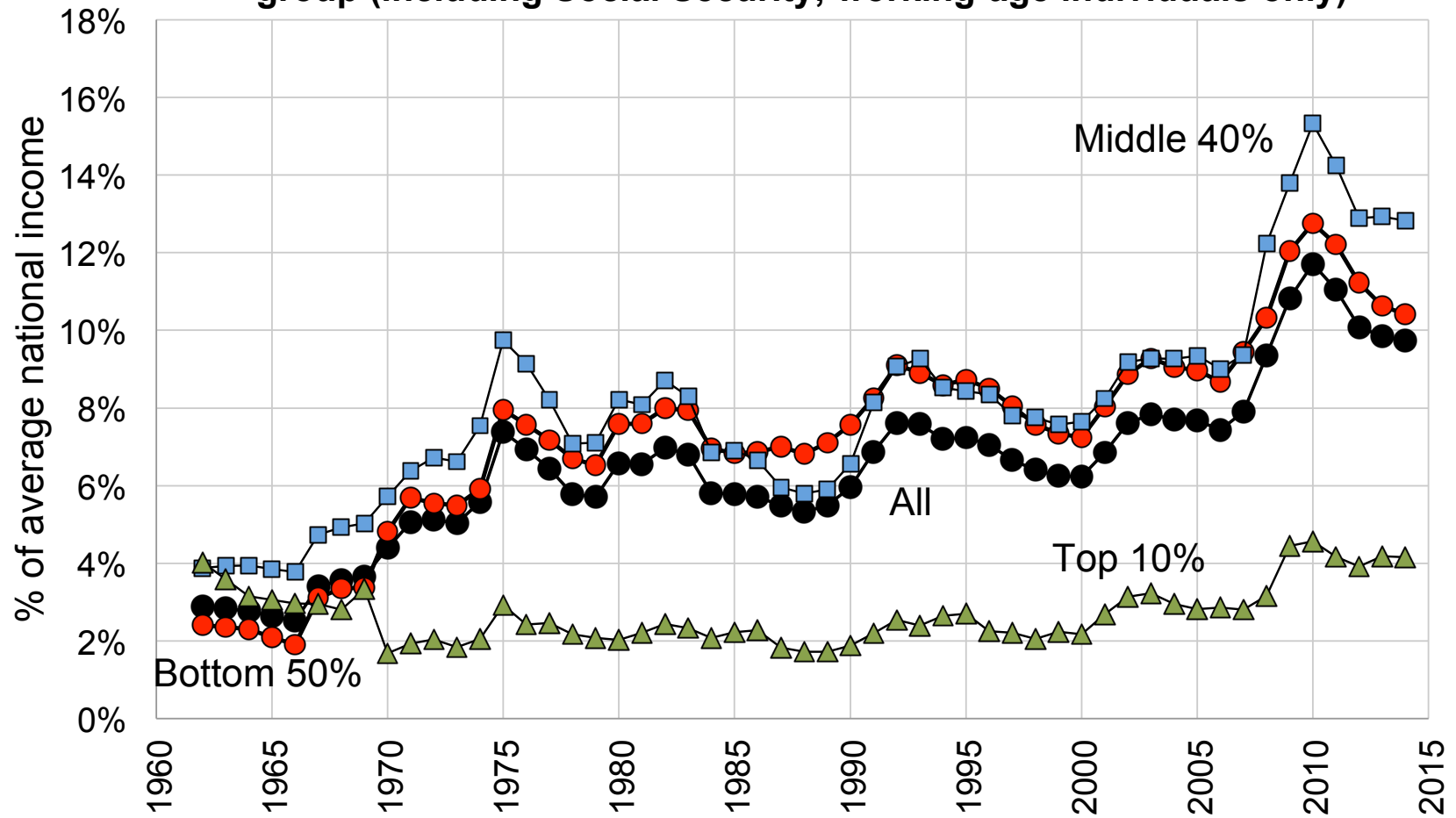
**Figure S.13: Average individualized transfer by post-tax income group (including Social Security)**



Source: Appendix Table II-G4b.

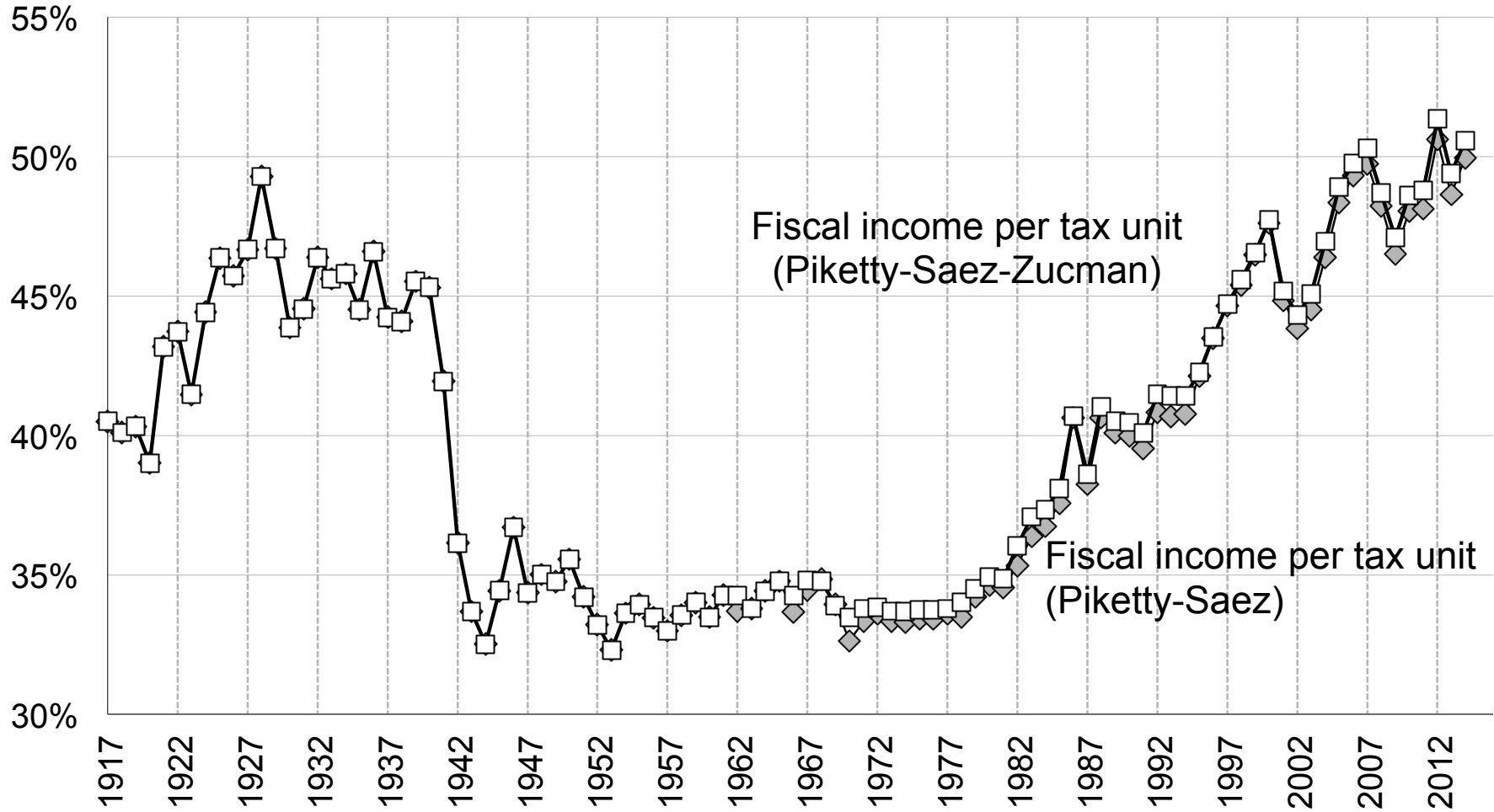


**Figure S.14: Average individualized transfer by post-tax income group (including Social Security; working-age individuals only)**



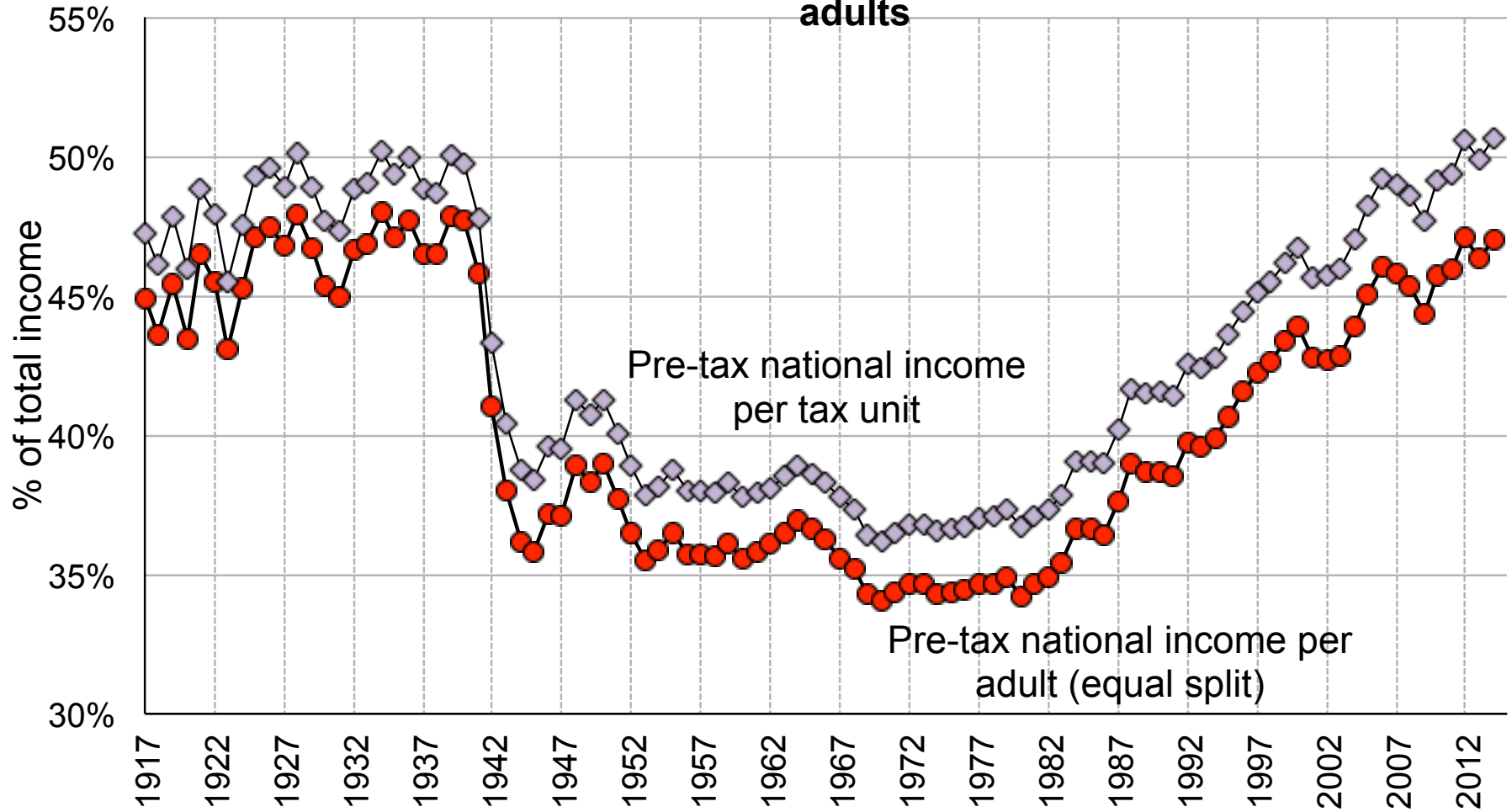
Source: Appendix Table II-G4c.

**Figure S.15a: Top 10% income share: comparison of estimates**



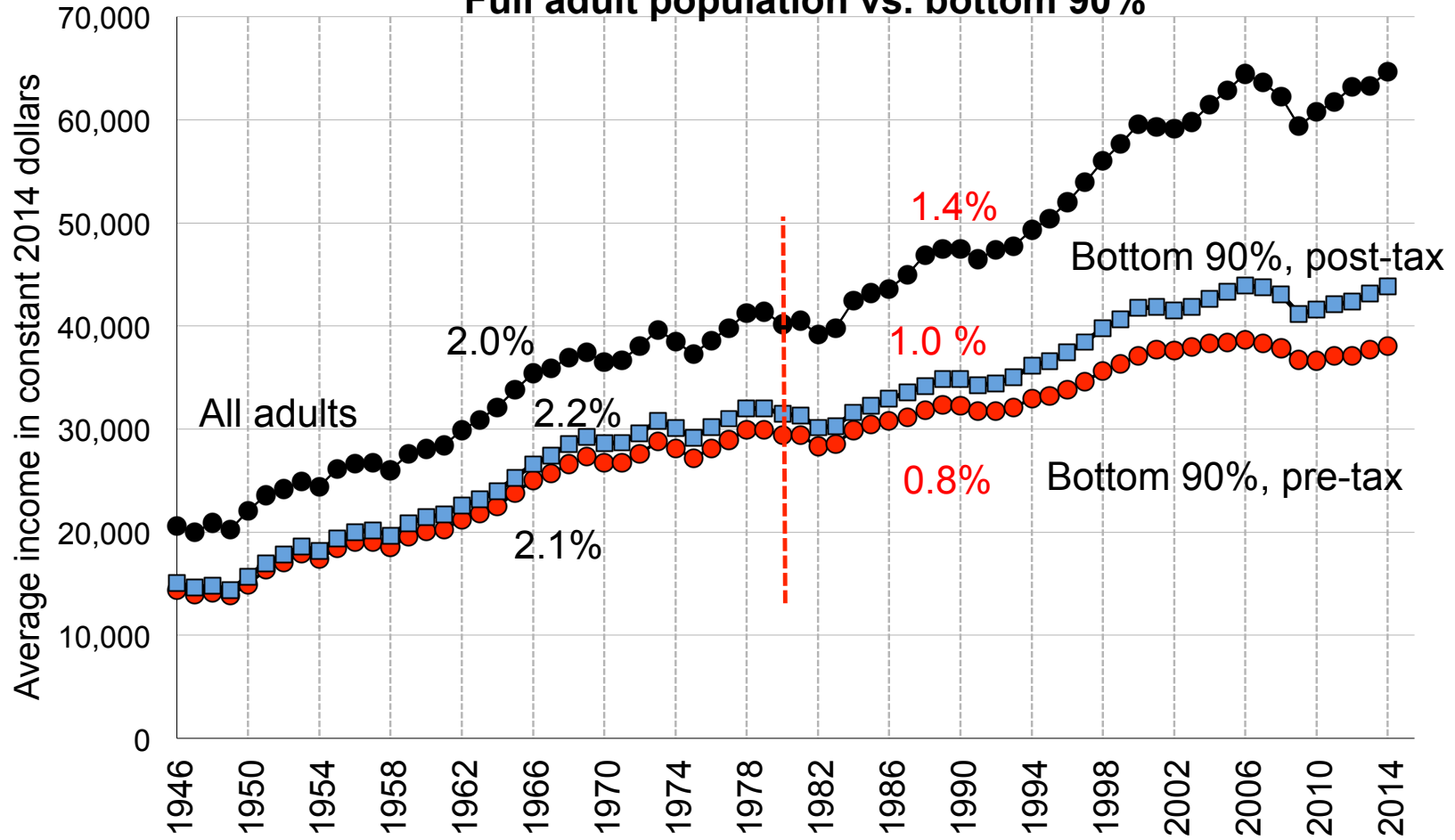
Source: Appendix Tables II-D1 and Piketty and Saez (2003, updated to 2014).

**Figure S.15b: Top 10% income share: tax units vs. equal-split adults**



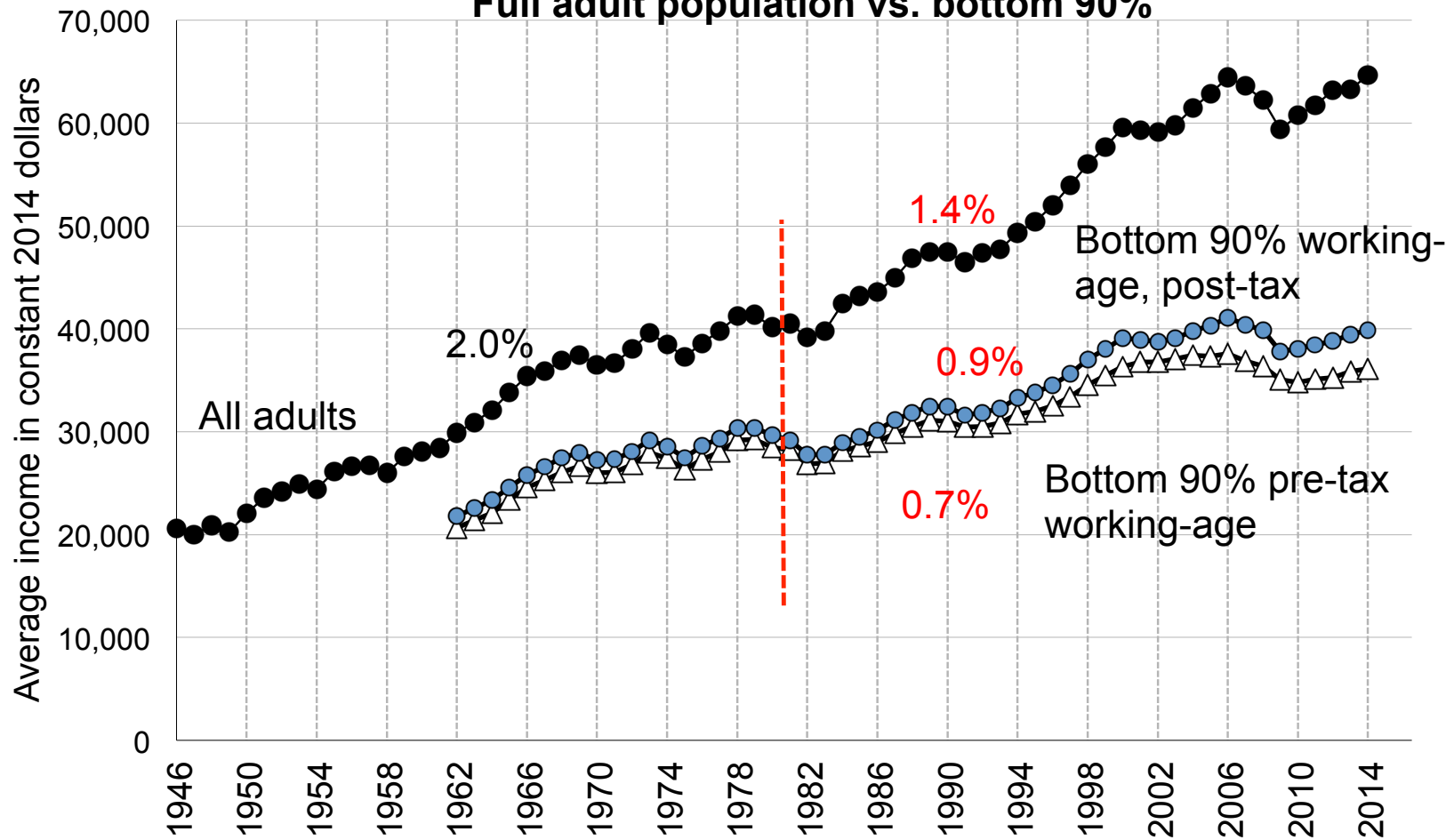
Source: Appendix TablesII- B1 and II-B9.

**Figure S.16: Real average national income:  
Full adult population vs. bottom 90%**



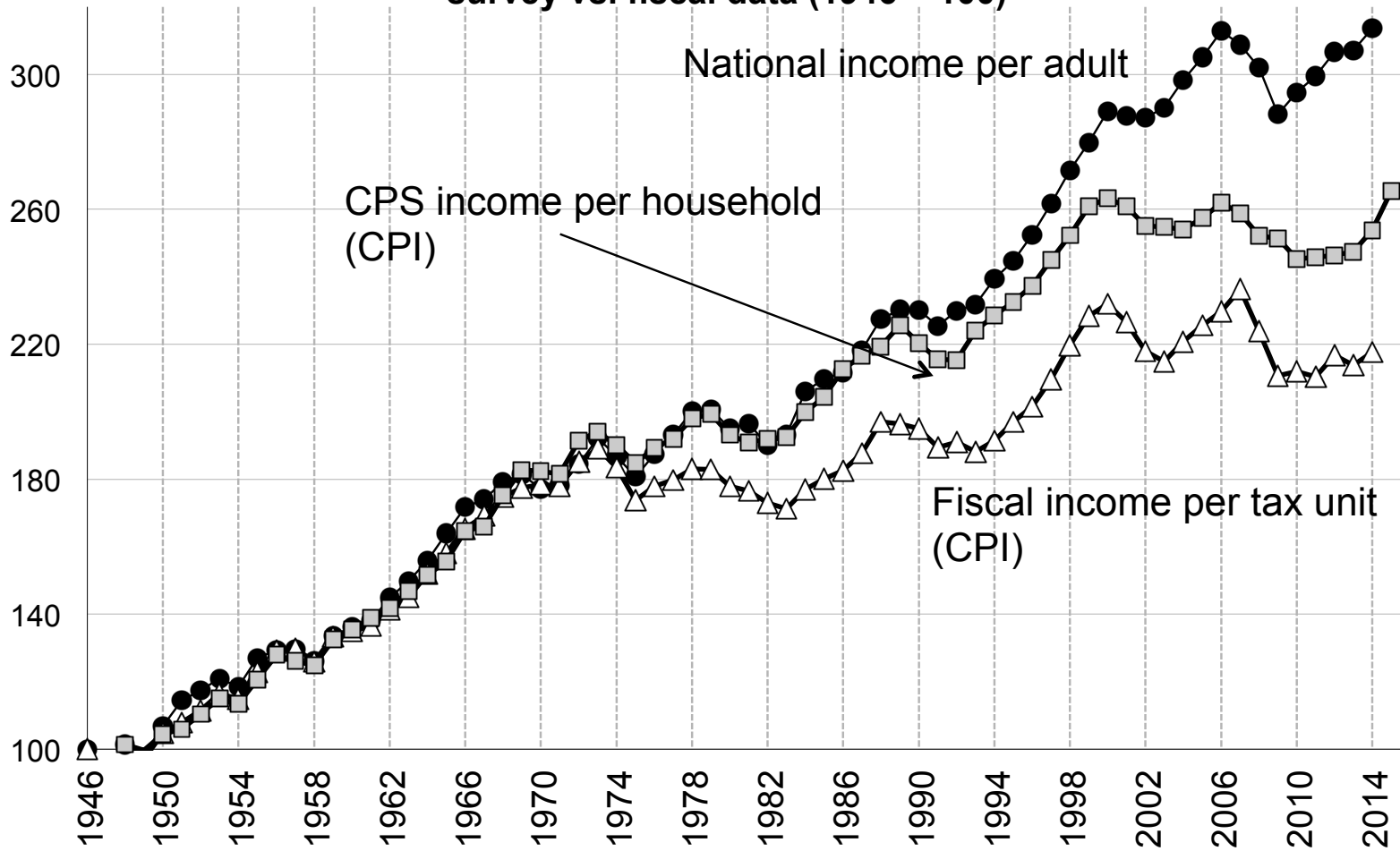
Source: Appendix Table II-B3 and II-C3

**Figure S.17: Real average national income:  
Full adult population vs. bottom 90%**



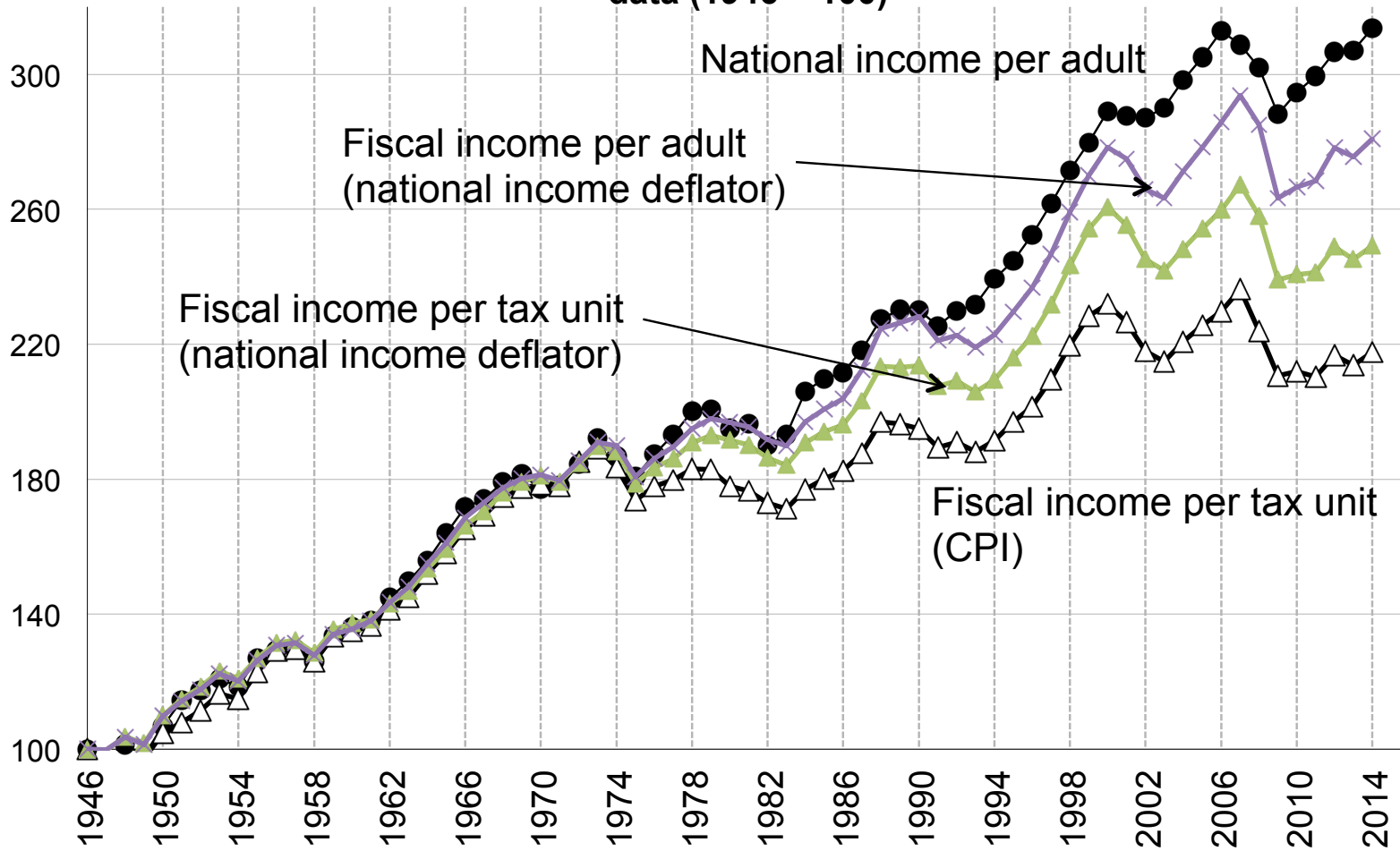
Source: Tables II-B6 and II-C6.

**Figure S.18: Average real income growth: national accounts vs. survey vs. fiscal data (1946 = 100)**



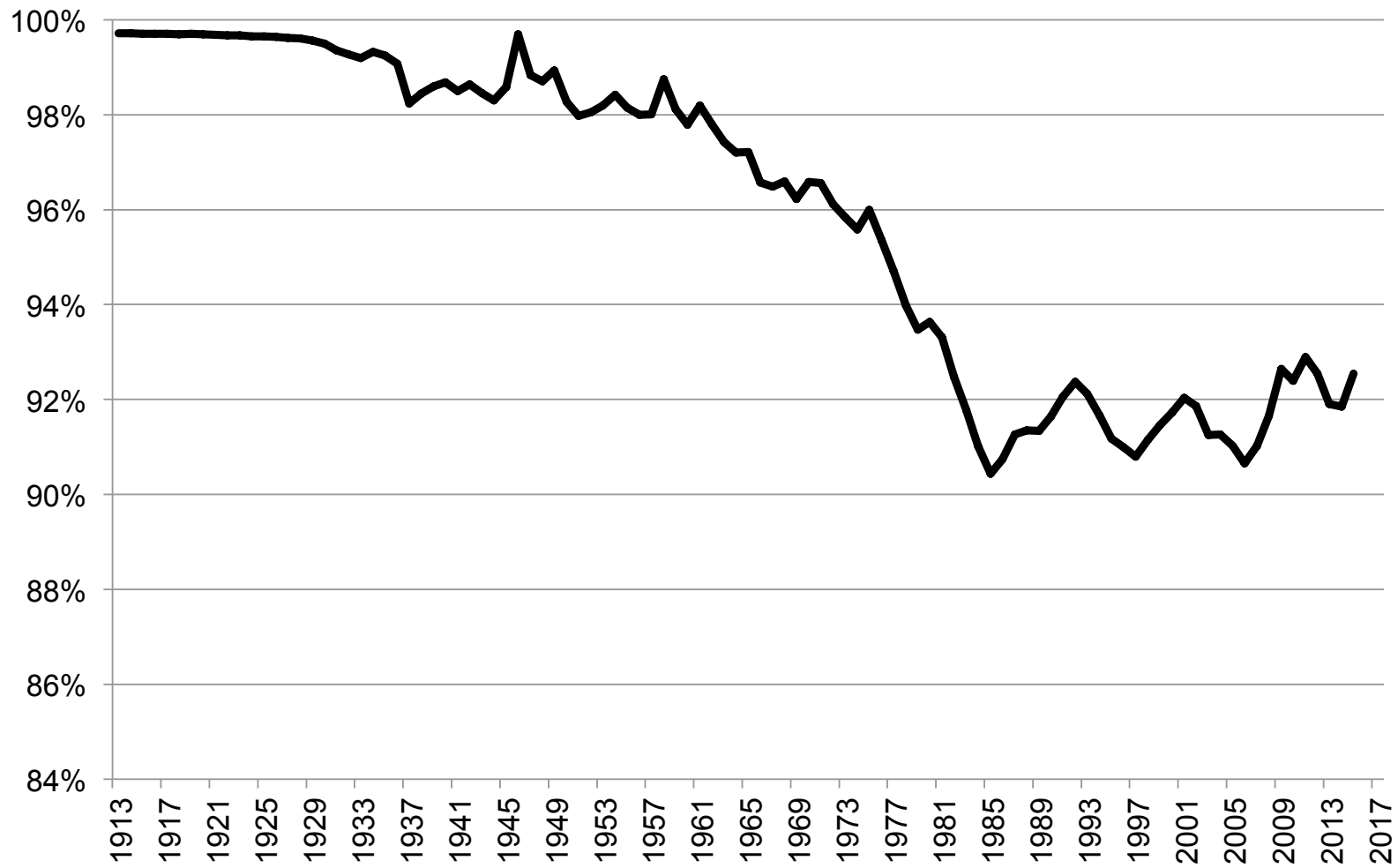
Source: Appendix Table I-A0 and Census Bureau (2016).

**Figure S.19: Average real income growth: national accounts vs. fiscal data (1946 = 100)**



Source: Appendix Table I-A0.

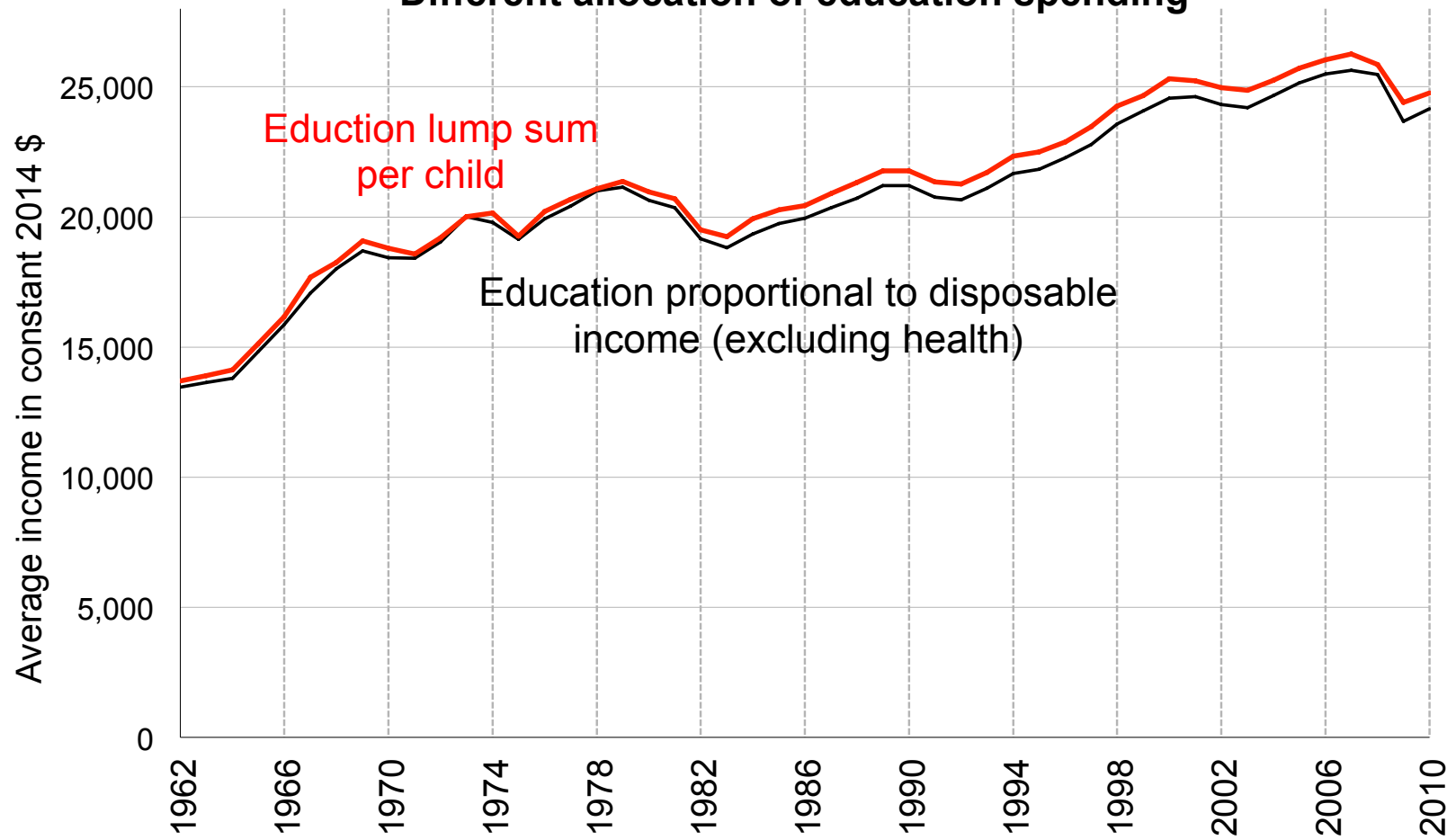
**Figure S.20: Ratio of personal pre-tax income to personal factor income**



Source: Appendix Table I-A7.

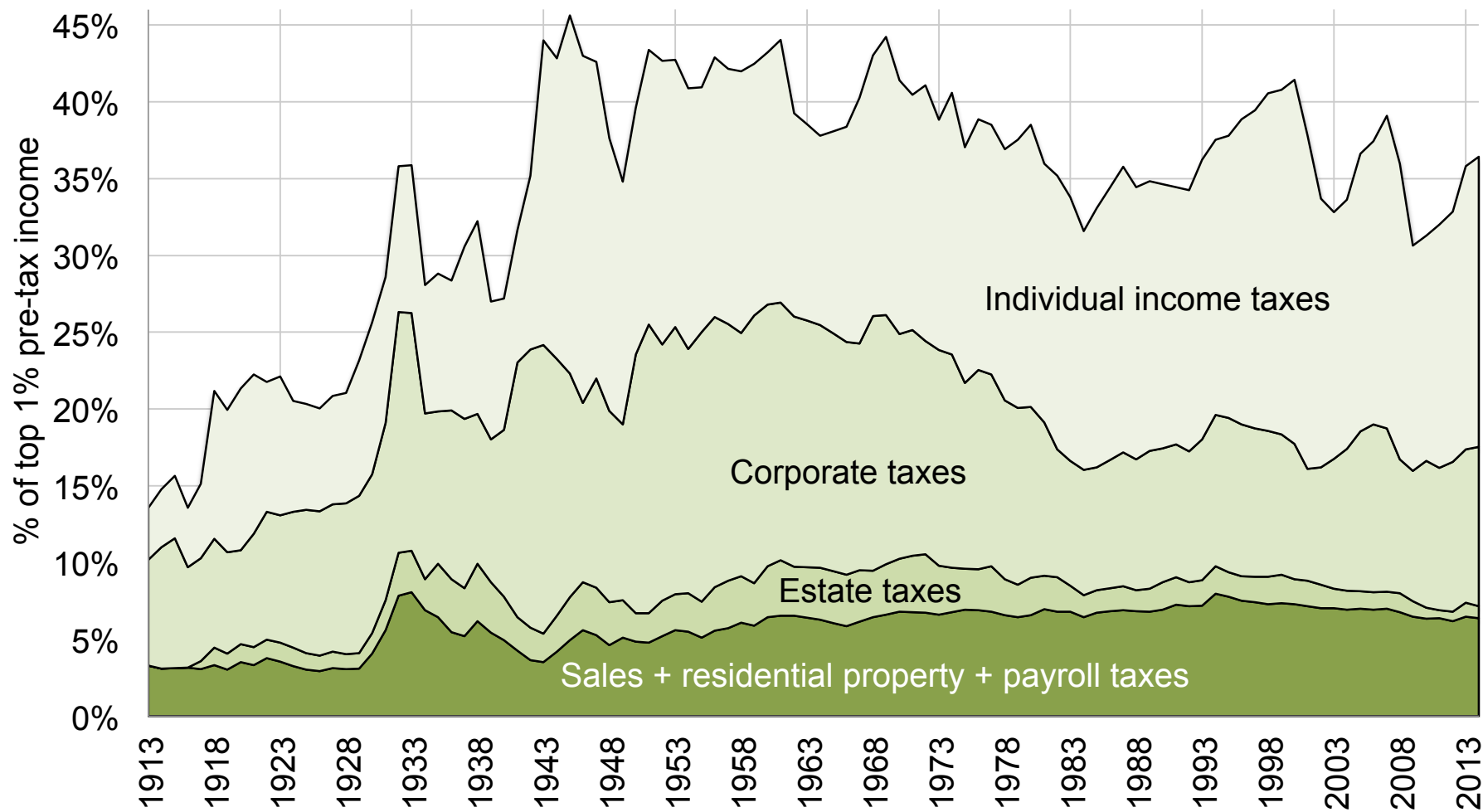


**Figure S.21: Real post-tax income of bottom 50%:  
Different allocation of education spending**



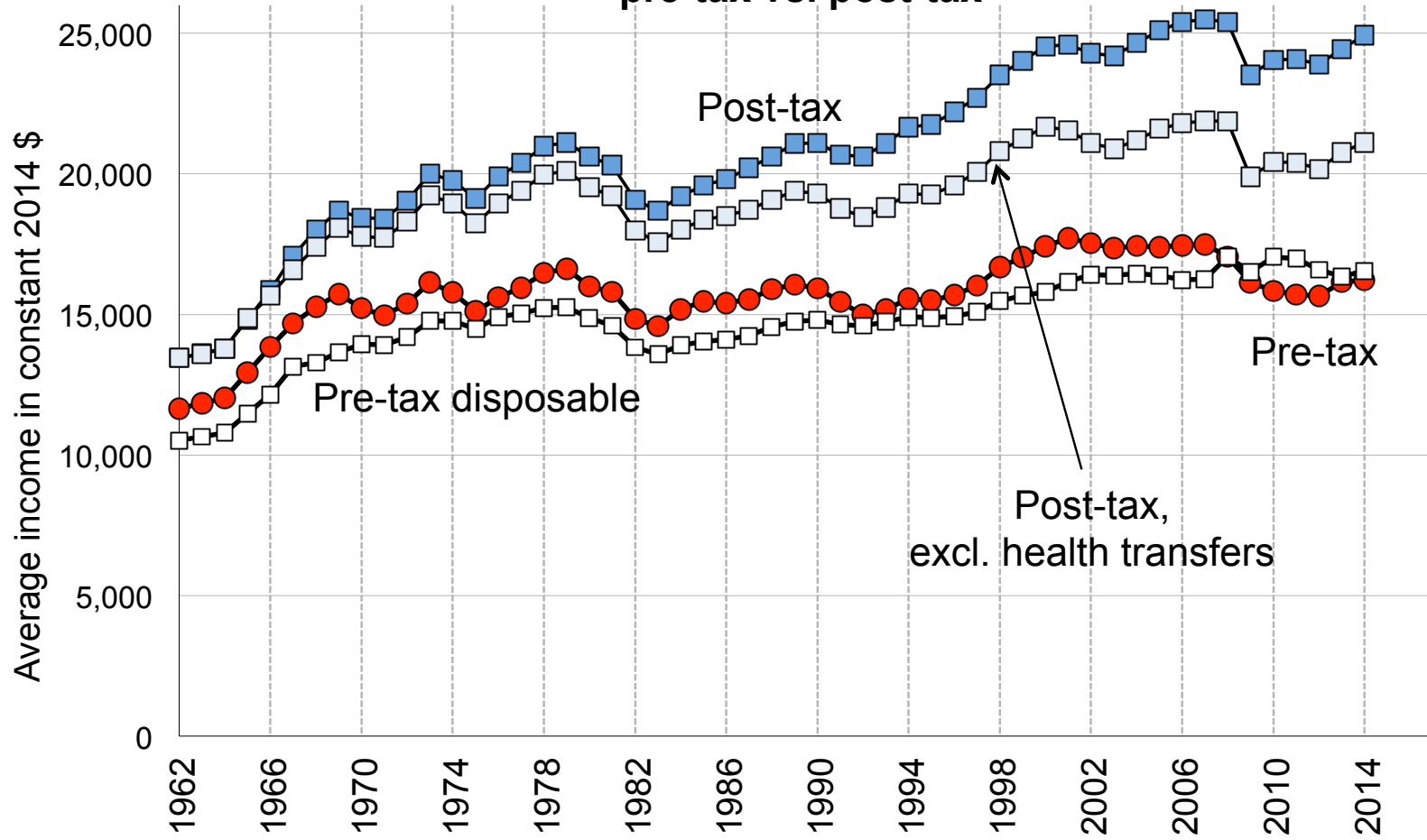
Source: Appendix Tables II-C3d.

**Figure S.22: Taxes paid by the top 1%**



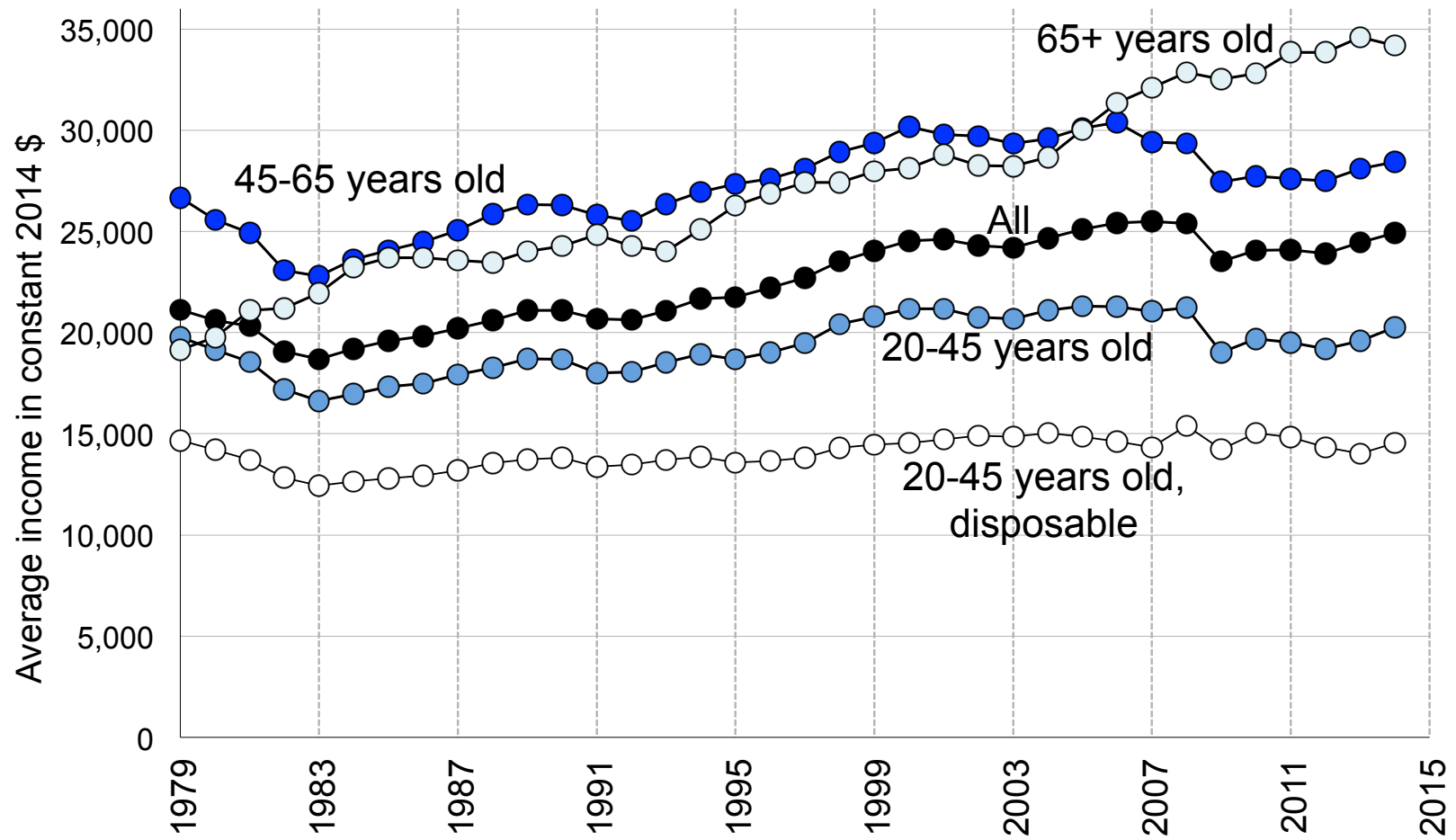
Source: Appendix Table II-G2

**Figure S.23: Real income of bottom 50%:  
pre-tax vs. post-tax**



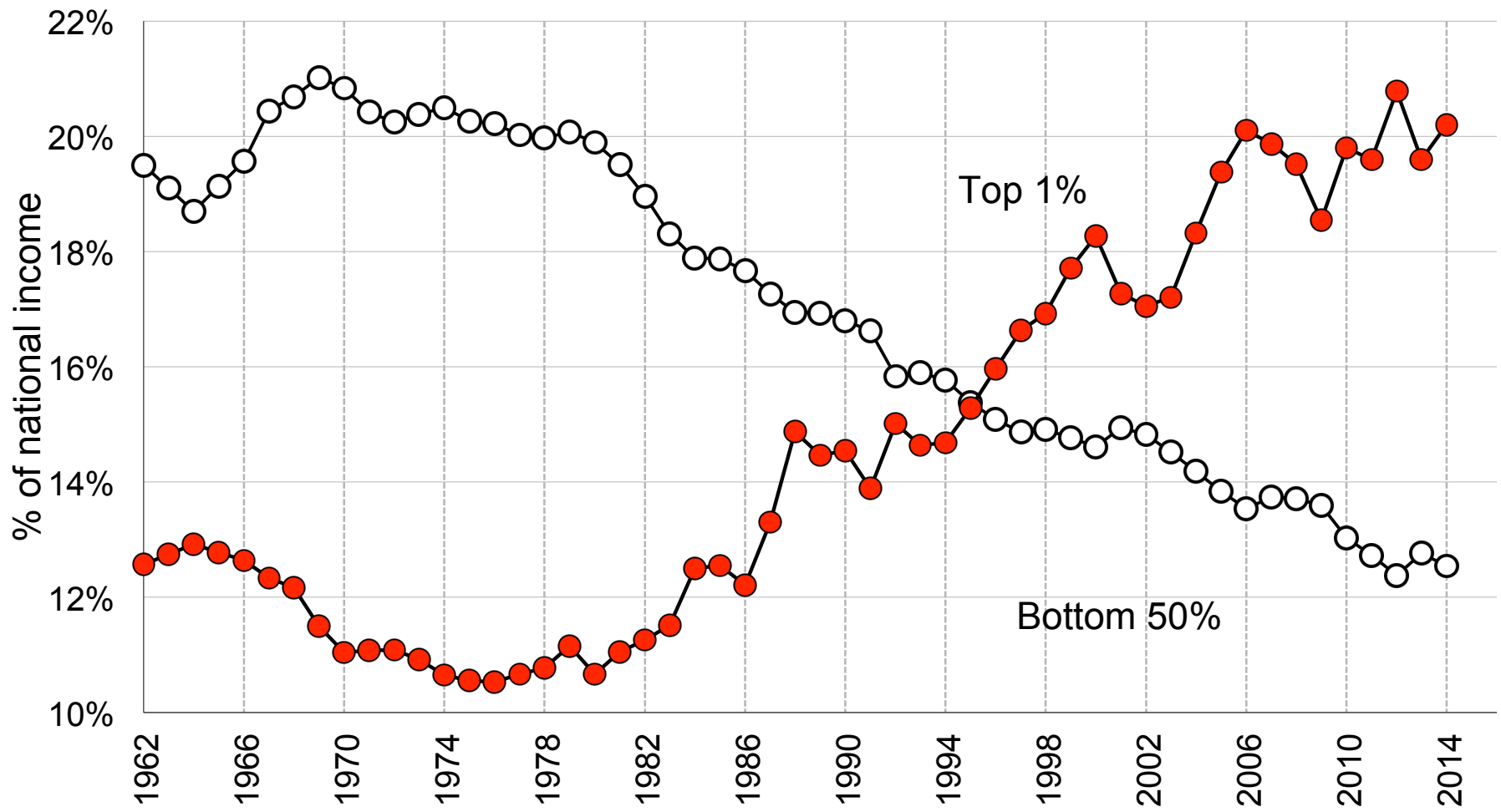
Source: Appendix Tables II-B7, II-C7 and II-C3c.

**Figure S.24: Real post-tax income of bottom 50%, by age**



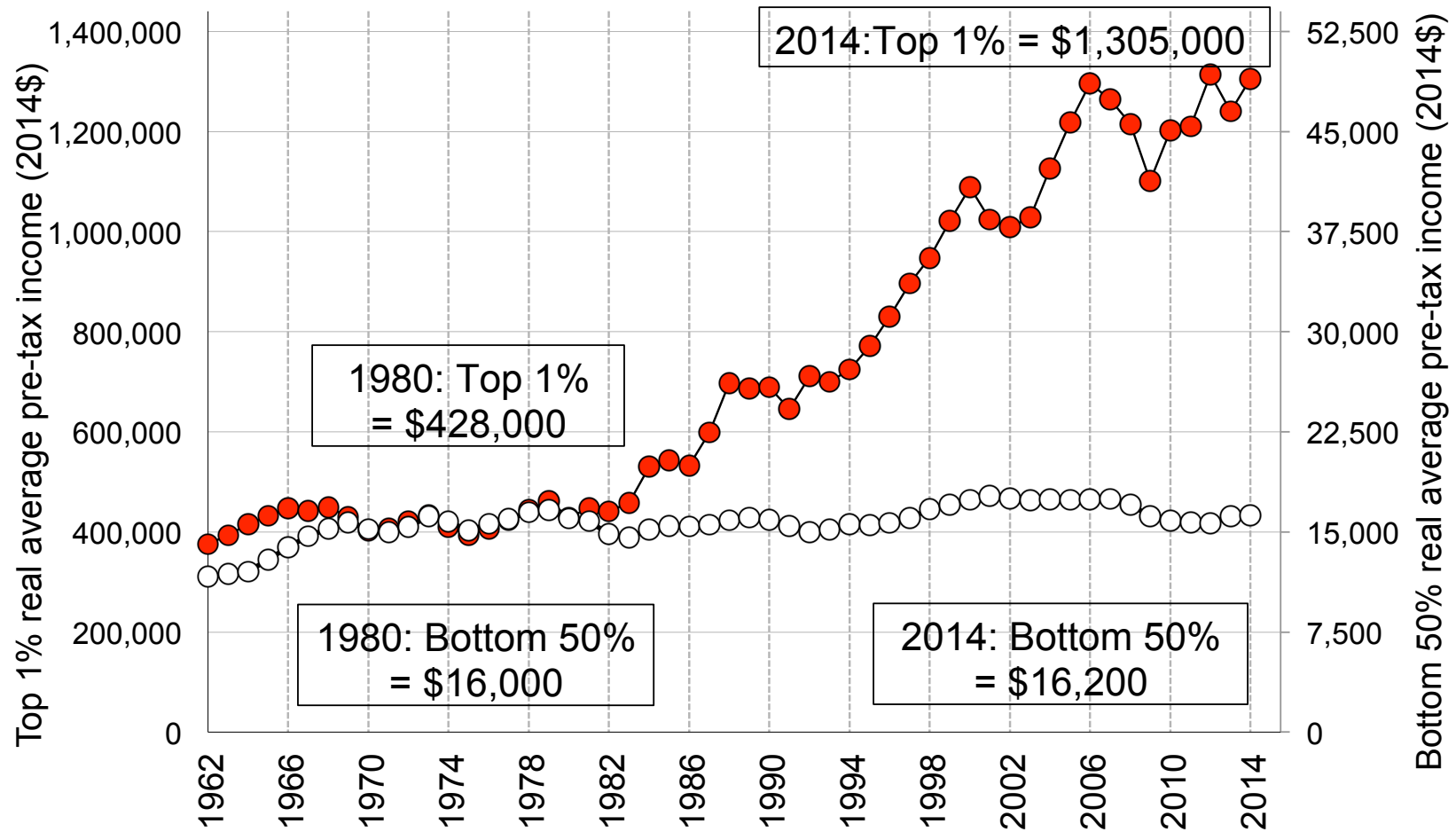
Source: Appendix Tables II-C7, II-C7b and II-C7d.

**Figure S.25: Pre-tax national income share: top 1% vs. bottom 50%**



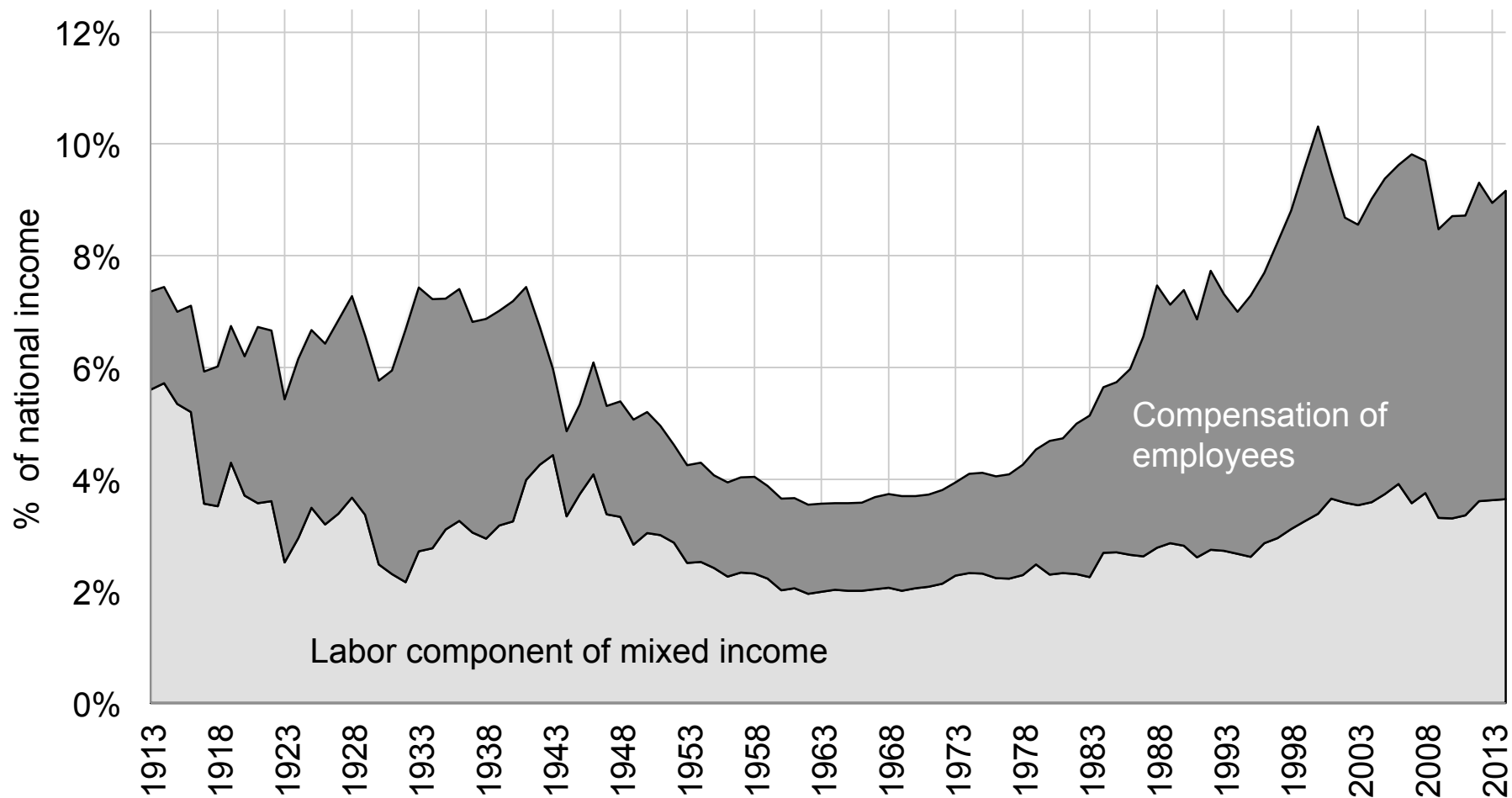
Source: Appendix Table II-B1

**Figure S.25b: Real pre-tax income of bottom 50% vs. top 1%**



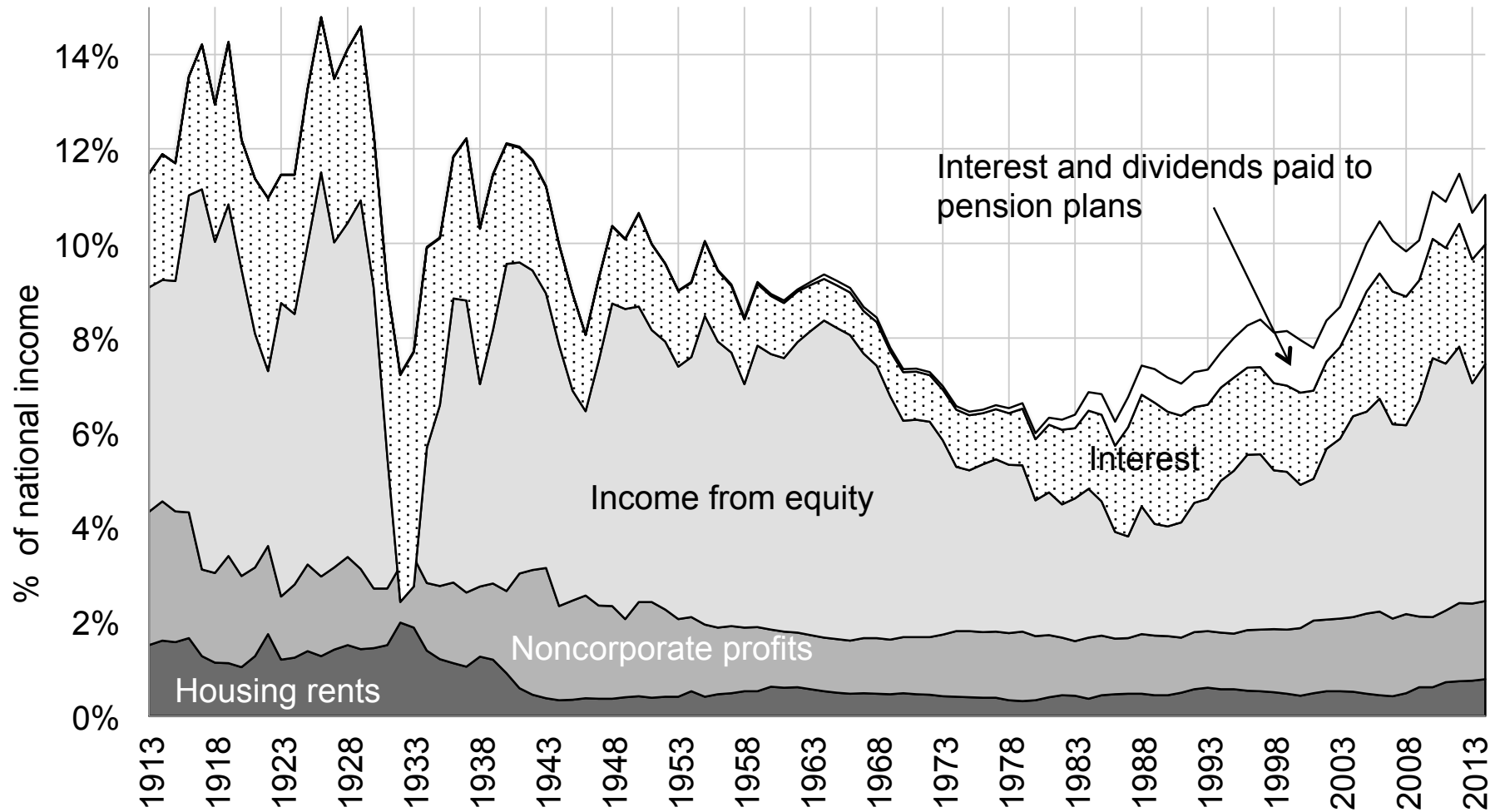
Source: Appendix Tables II-B7 and II-B10

**Figure S.26: Pre-tax labor income of top 1% adult income earners**



Source: Appendix Table II-B2b.

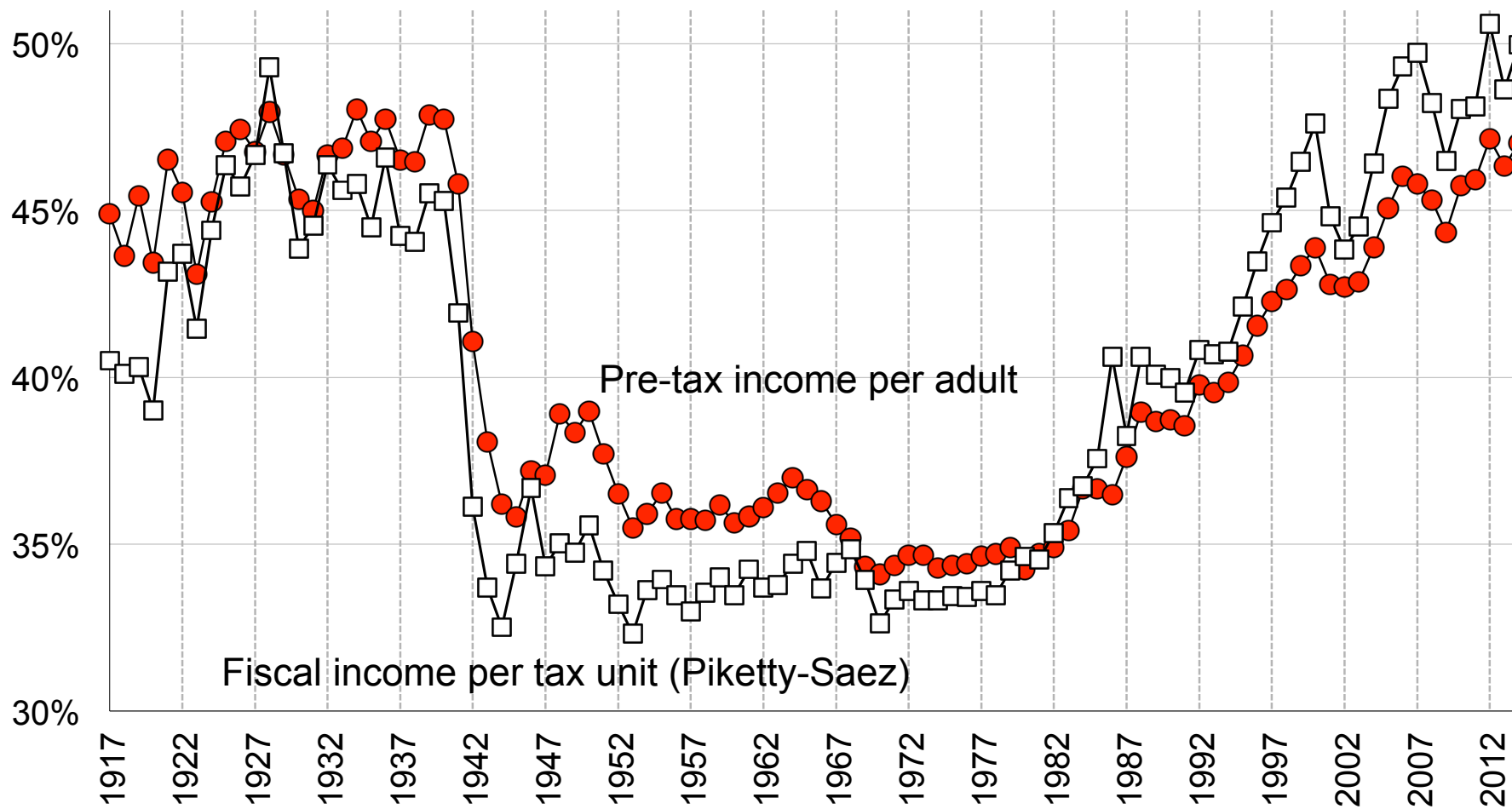
**Figure S.27: Pre-tax capital income of top 1% adult income earners**



Source: Appendix Table II-B2b

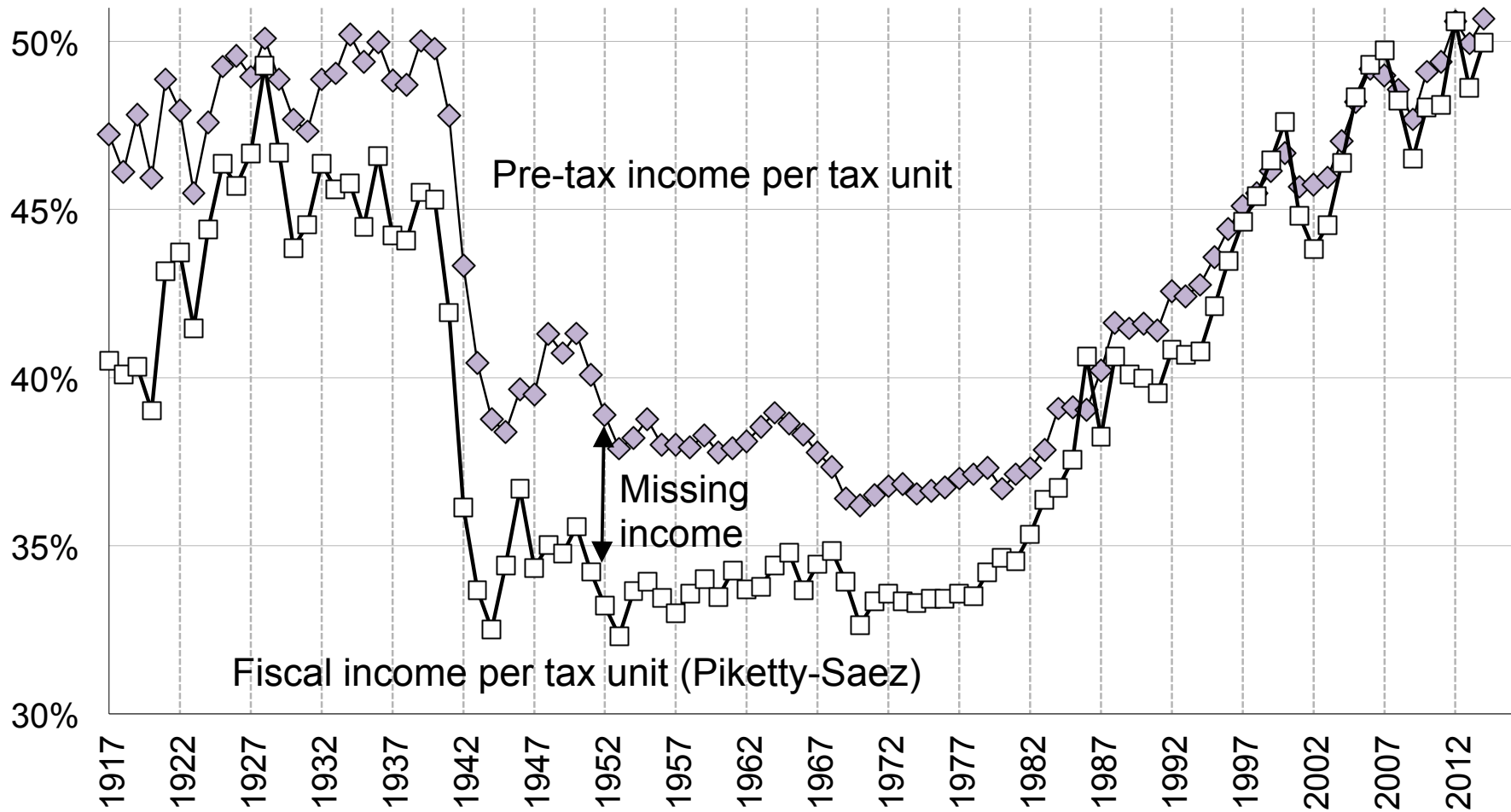


**Figure S.28: Top 10% income share: comparison of estimates**



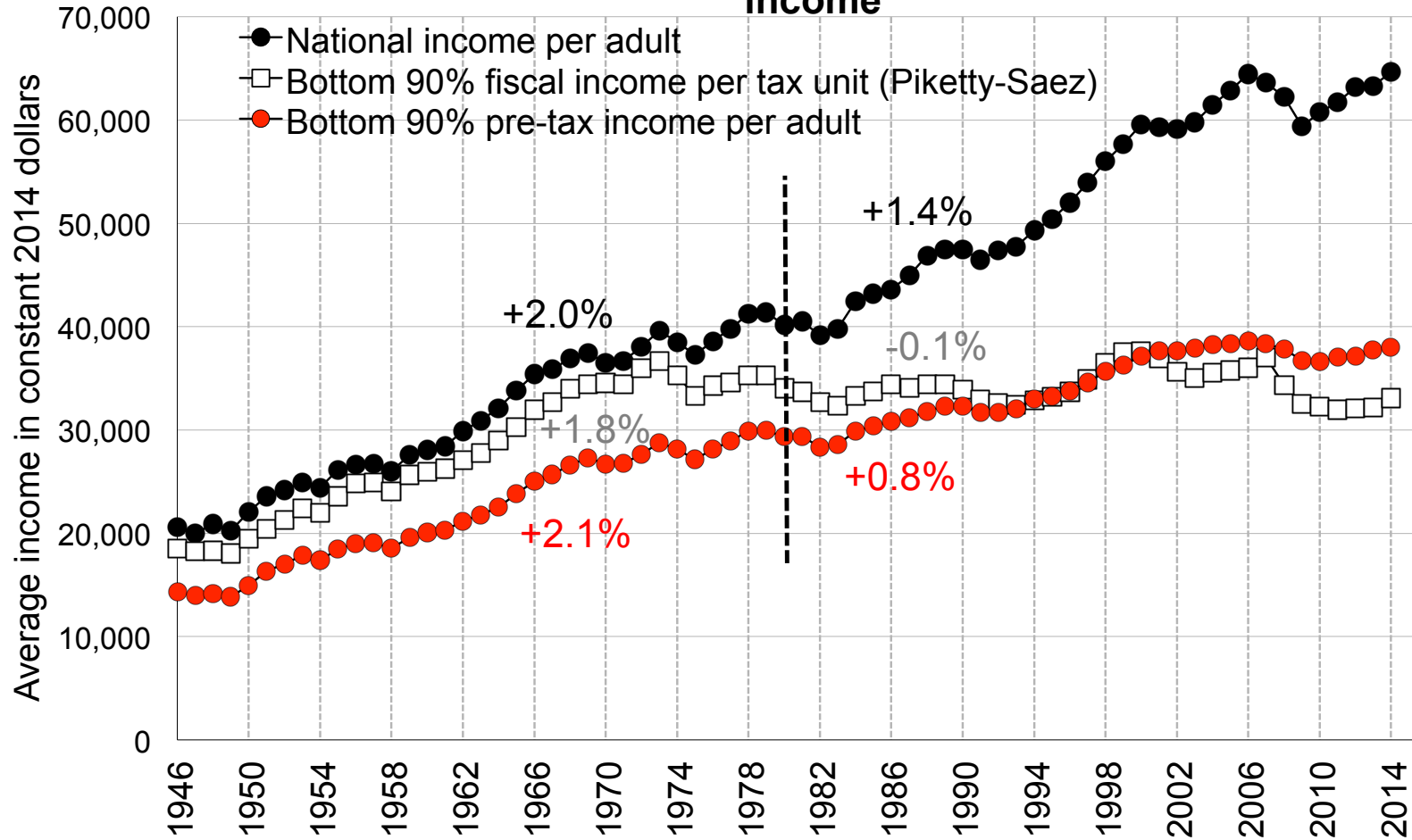
Source: Appendix Table II-B1 and Piketty and Saez (2003, updated to 2014).

**Figure S.29: Top 10% income share: fiscal vs. pre-tax income**



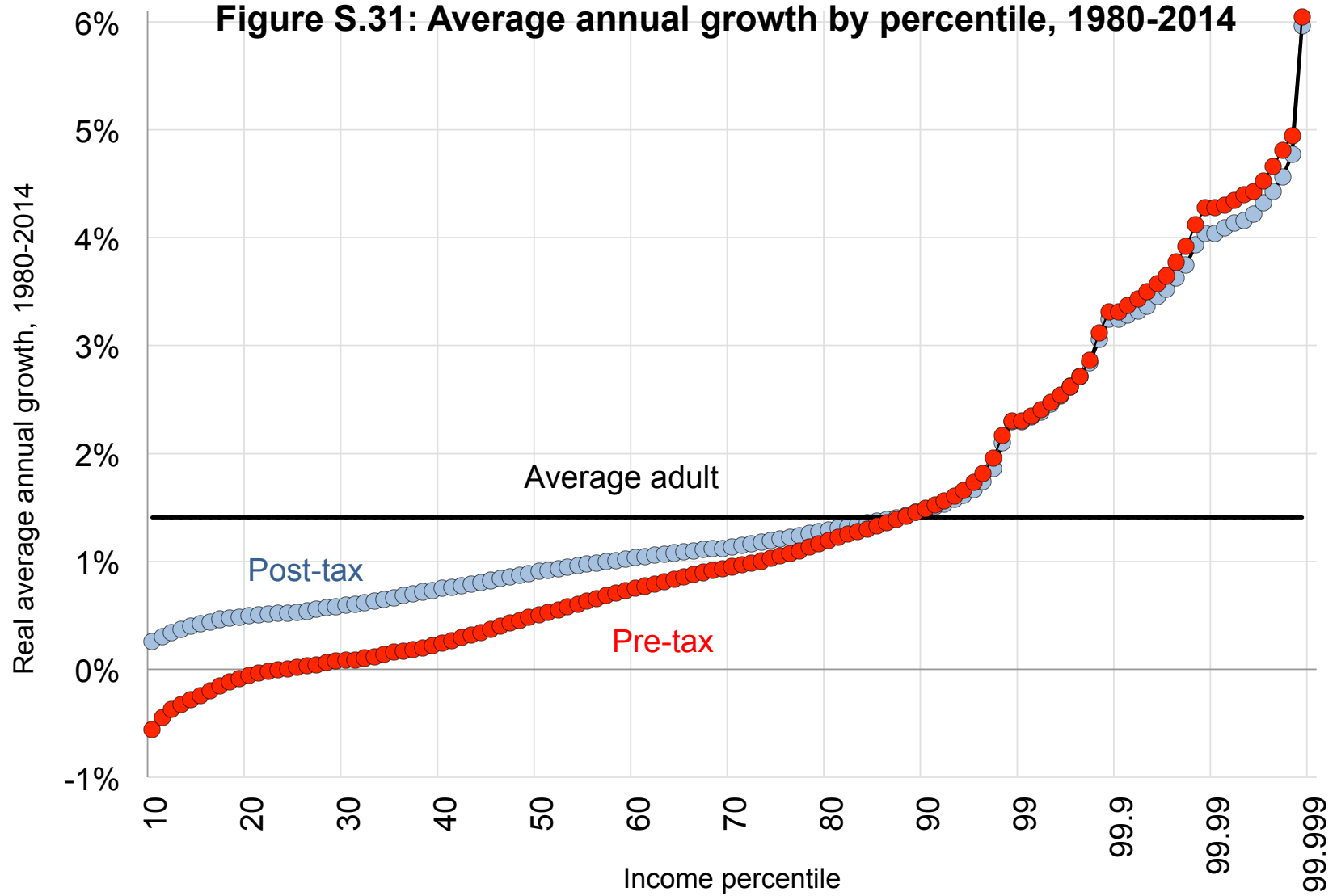
Source: Appendix Tables II-B9 and Piketty and Saez (2003, updated to 2014)

**Figure S.30: Bottom 90% income growth: Pre-tax vs. fiscal income**

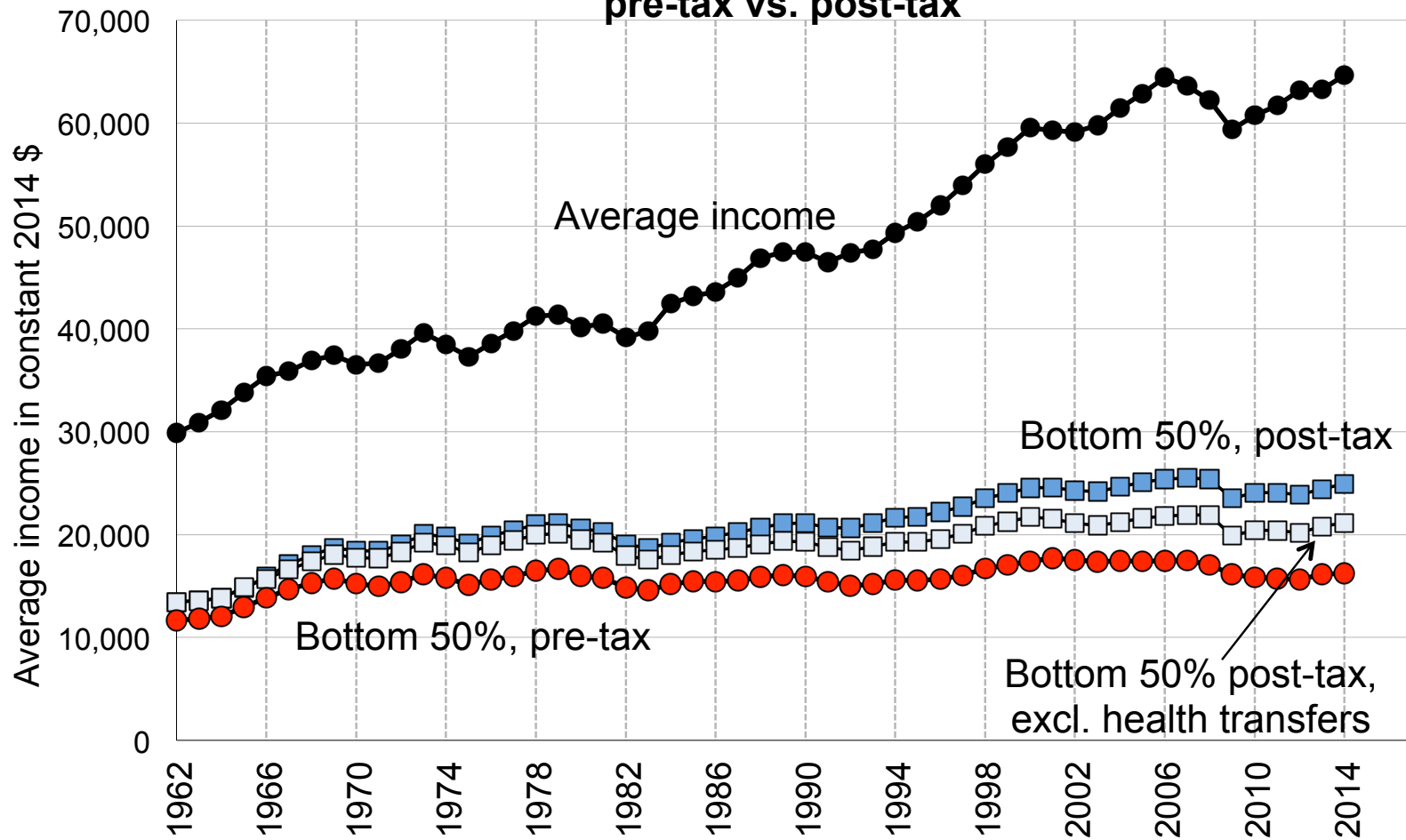


Source: Appendix Table II-B3 and Piketty and Saez (2003, updated to 2014)

**Figure S.31: Average annual growth by percentile, 1980-2014**

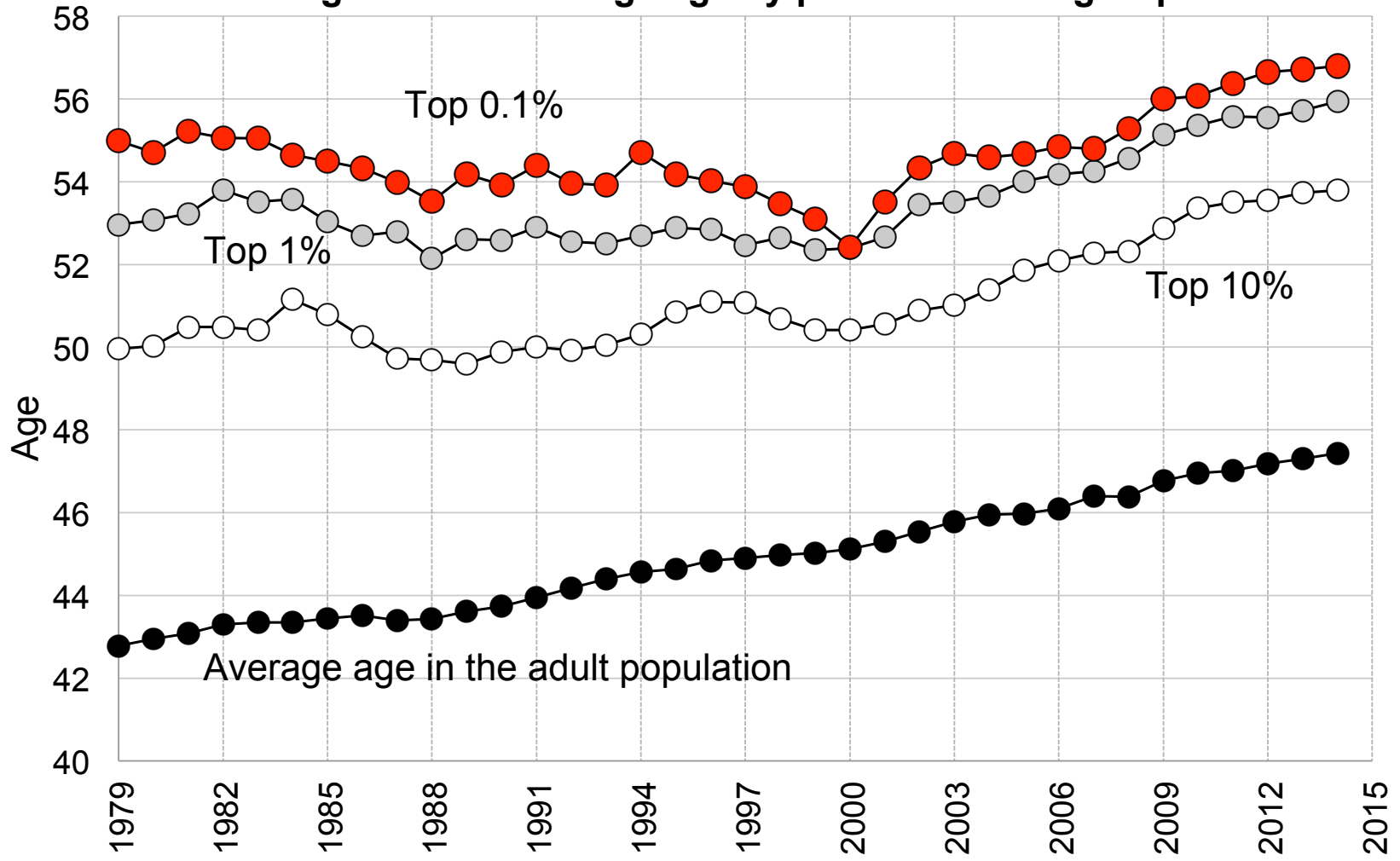


**Figure S.32: Real income of bottom 50%:  
pre-tax vs. post-tax**



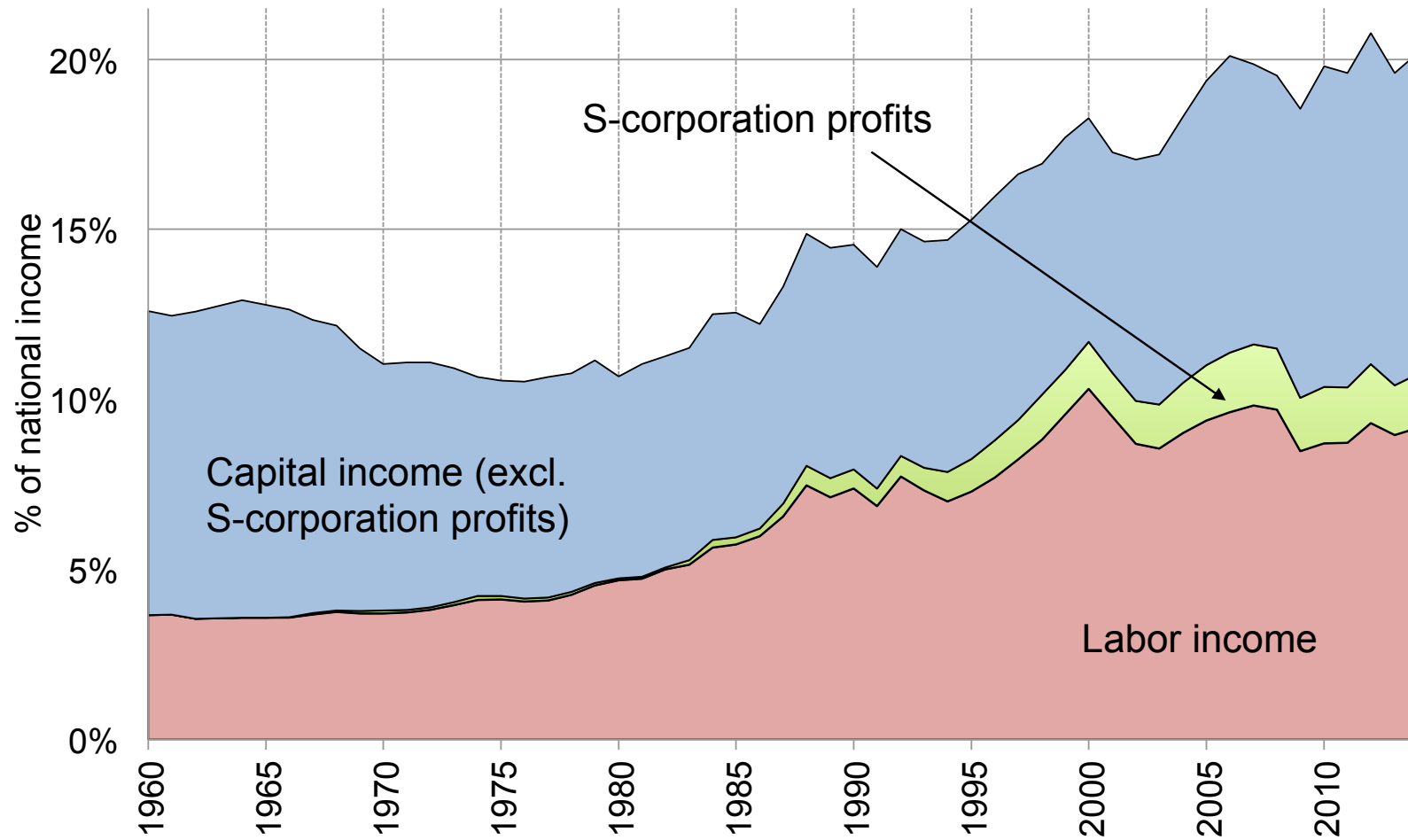
Source: Appendix Tables II-B7, II-C7 and II-C3c.

**Figure S.33: Average age by pre-tax income group**



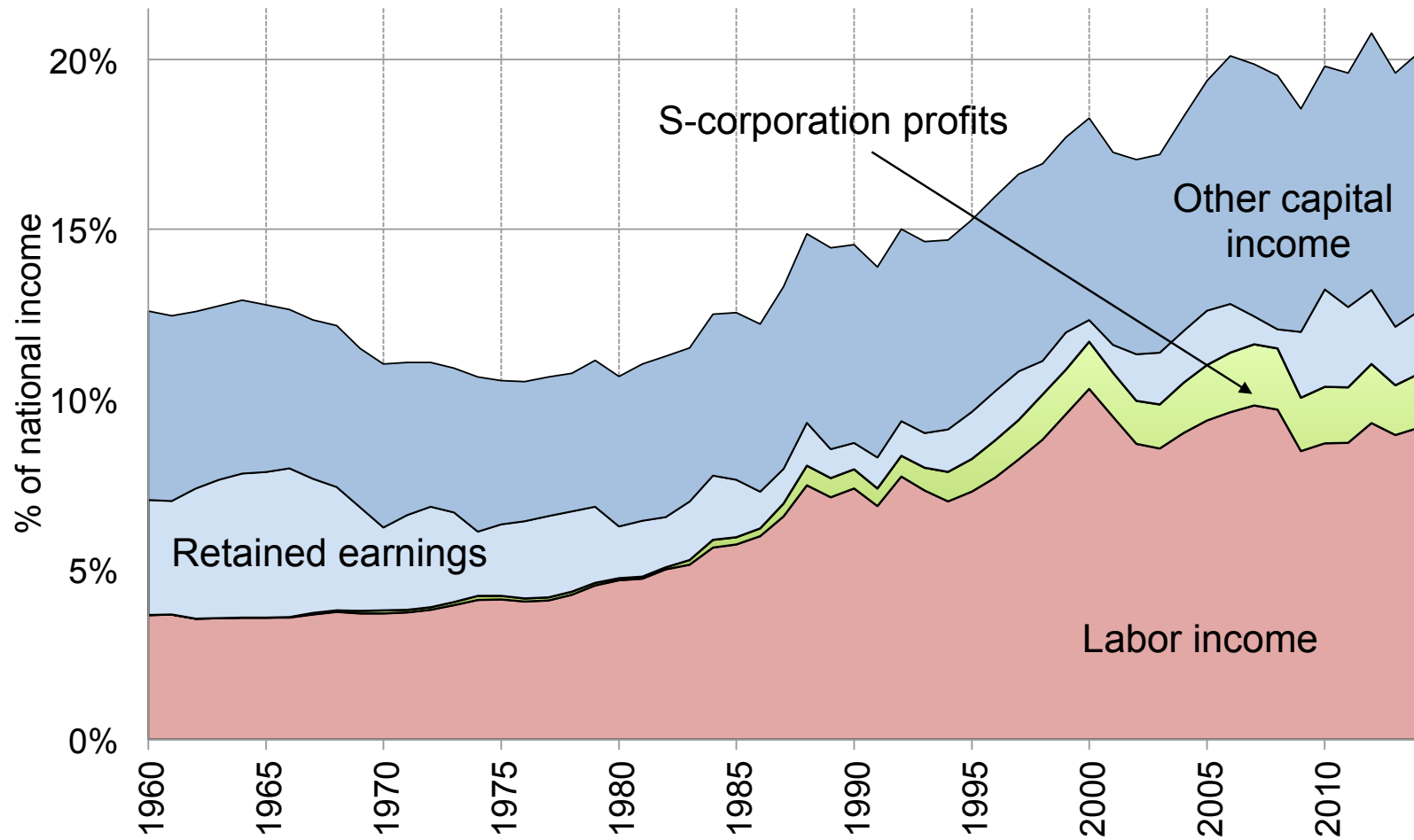
Source: Appendix Table II-F2.

**Figure S.34: The role of S-corporation profits for the top 1%**



Source: Appendix Table II-B2g.

**Figure S.35: The role of retained earnings for the top 1%**

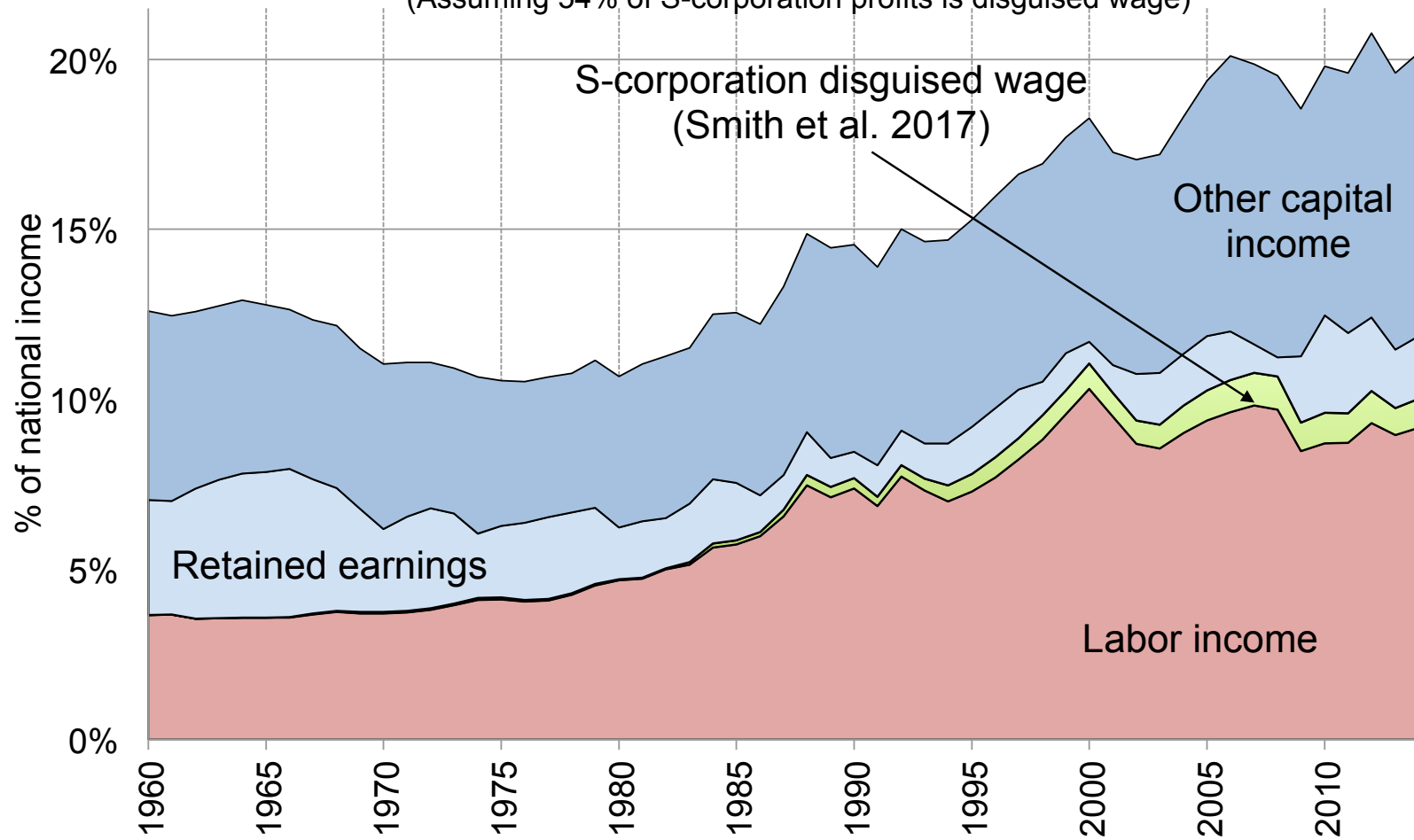


Source: Appendix Table II-B2g.



**Figure S.36: S-corporation disguised wage in the top 1 %**

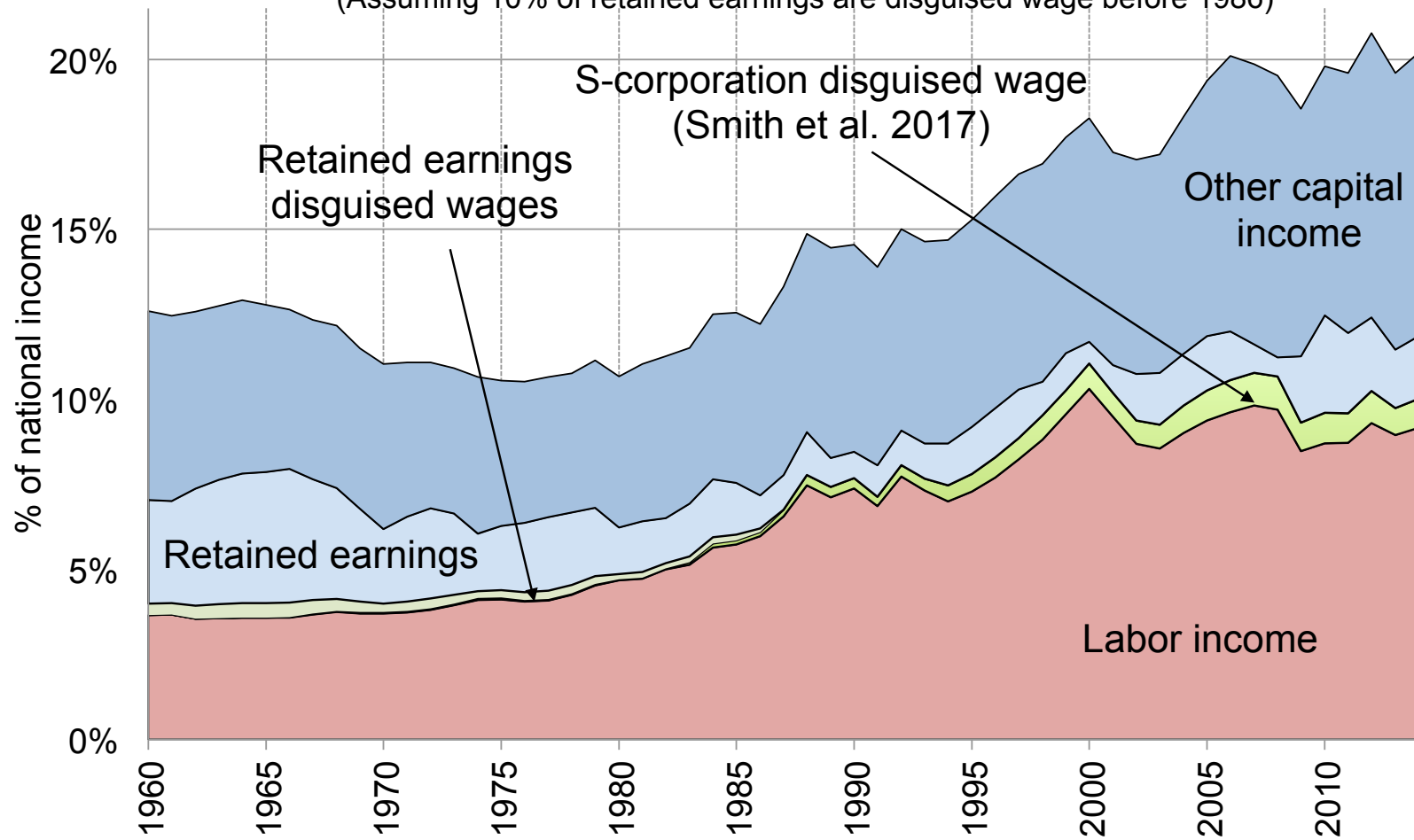
(Assuming 54% of S-corporation profits is disguised wage)



Source: Appendix Table II-B2g.

**Figure S.37: Retained earnings disguised wage in the top 1 %**

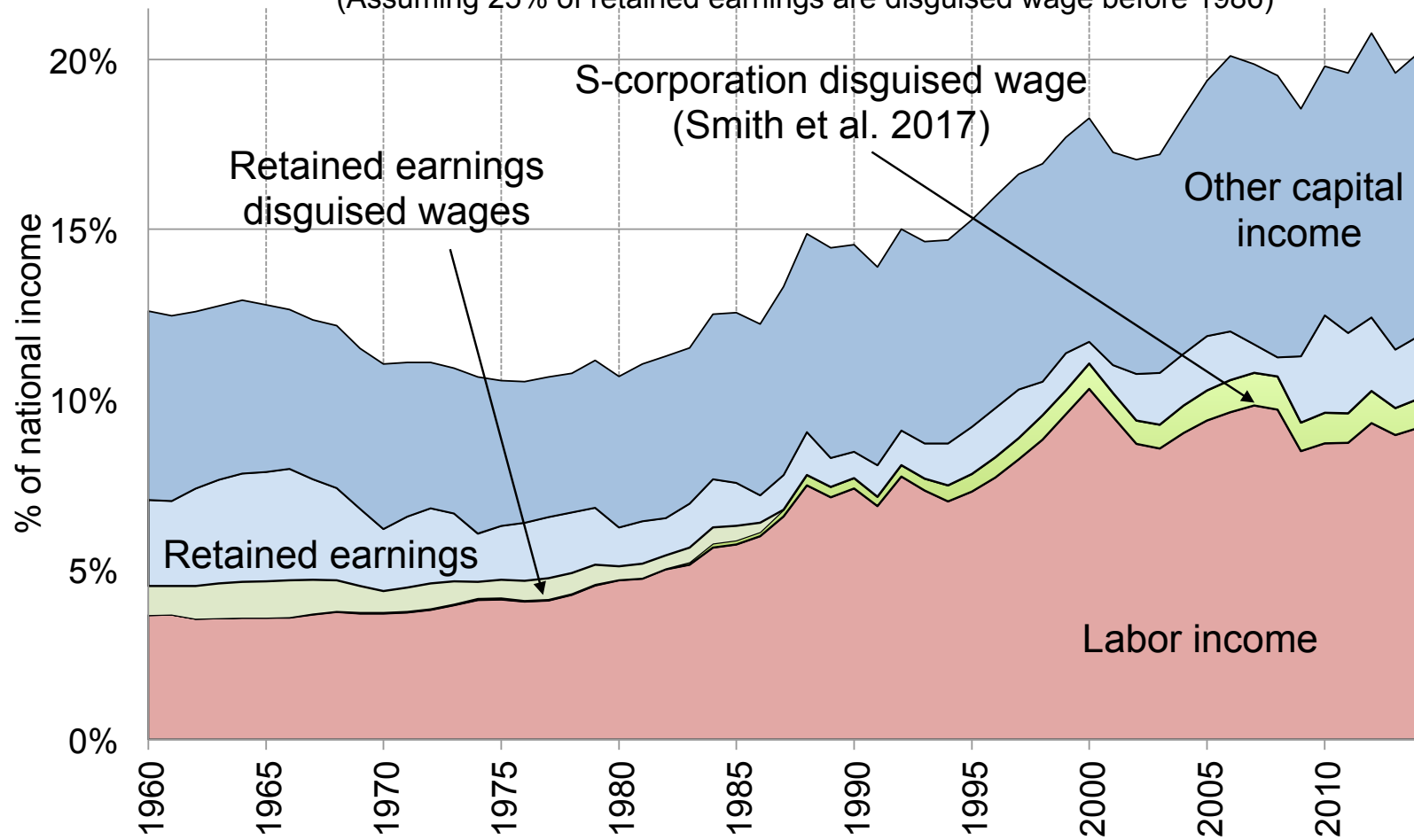
(Assuming 10% of retained earnings are disguised wage before 1986)



Source: Appendix Table II-B2g.

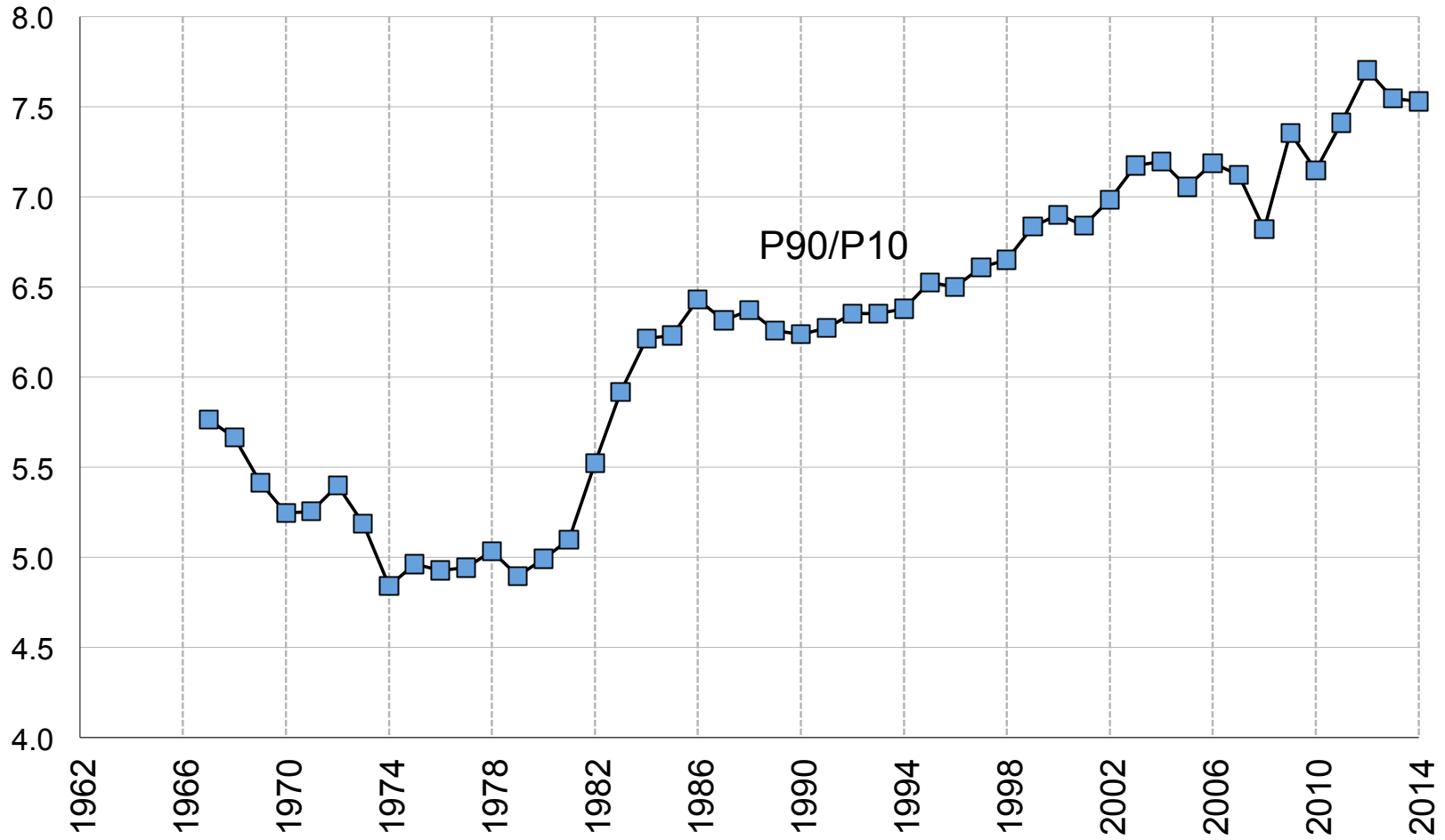
**Figure S.38: Retained earnings disguised wage in the top 1 %**

(Assuming 25% of retained earnings are disguised wage before 1986)



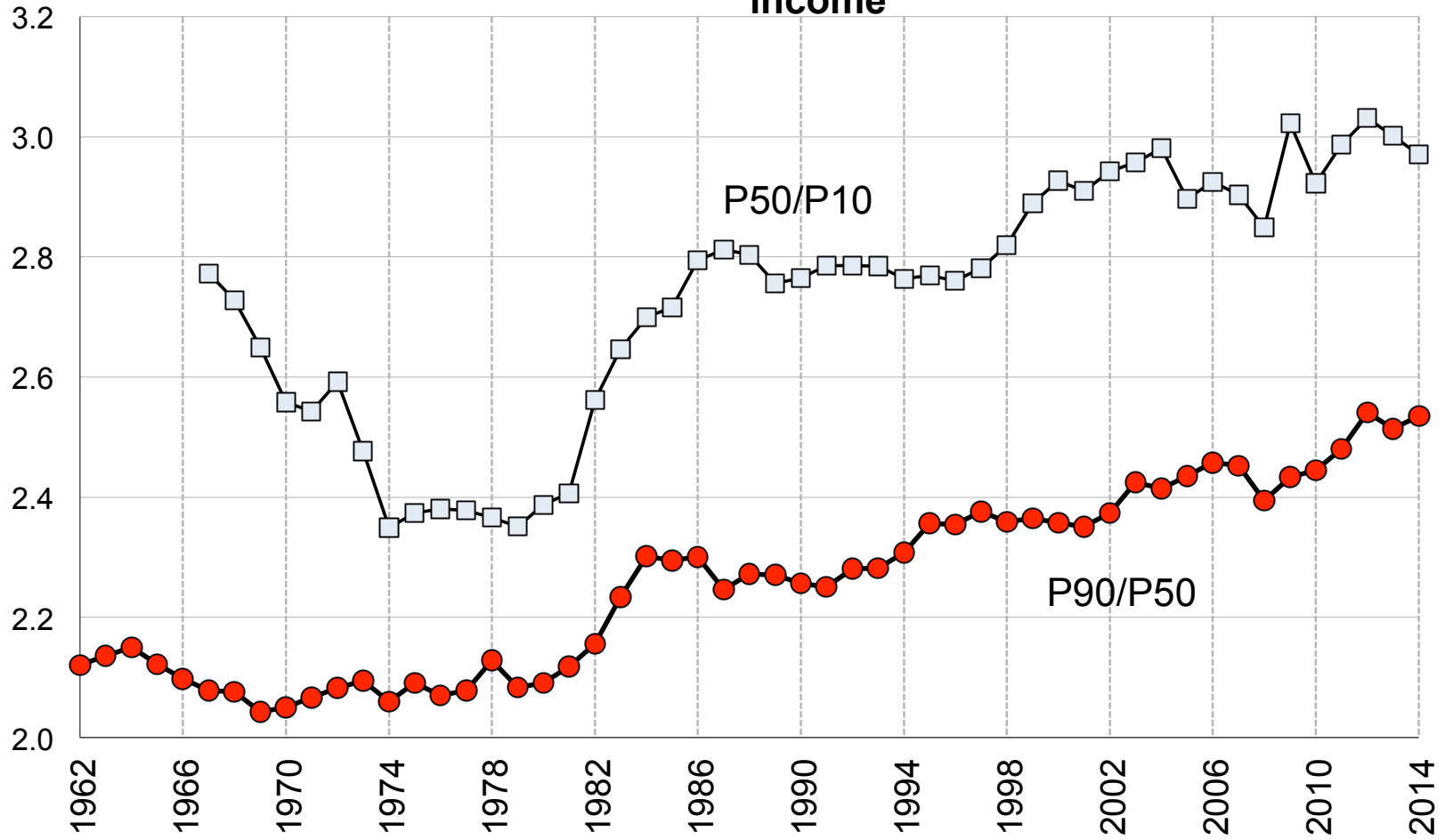
Source: Appendix Table II-g.

**Appendix Figure S.39: P90/P10 ratio, post-tax income**



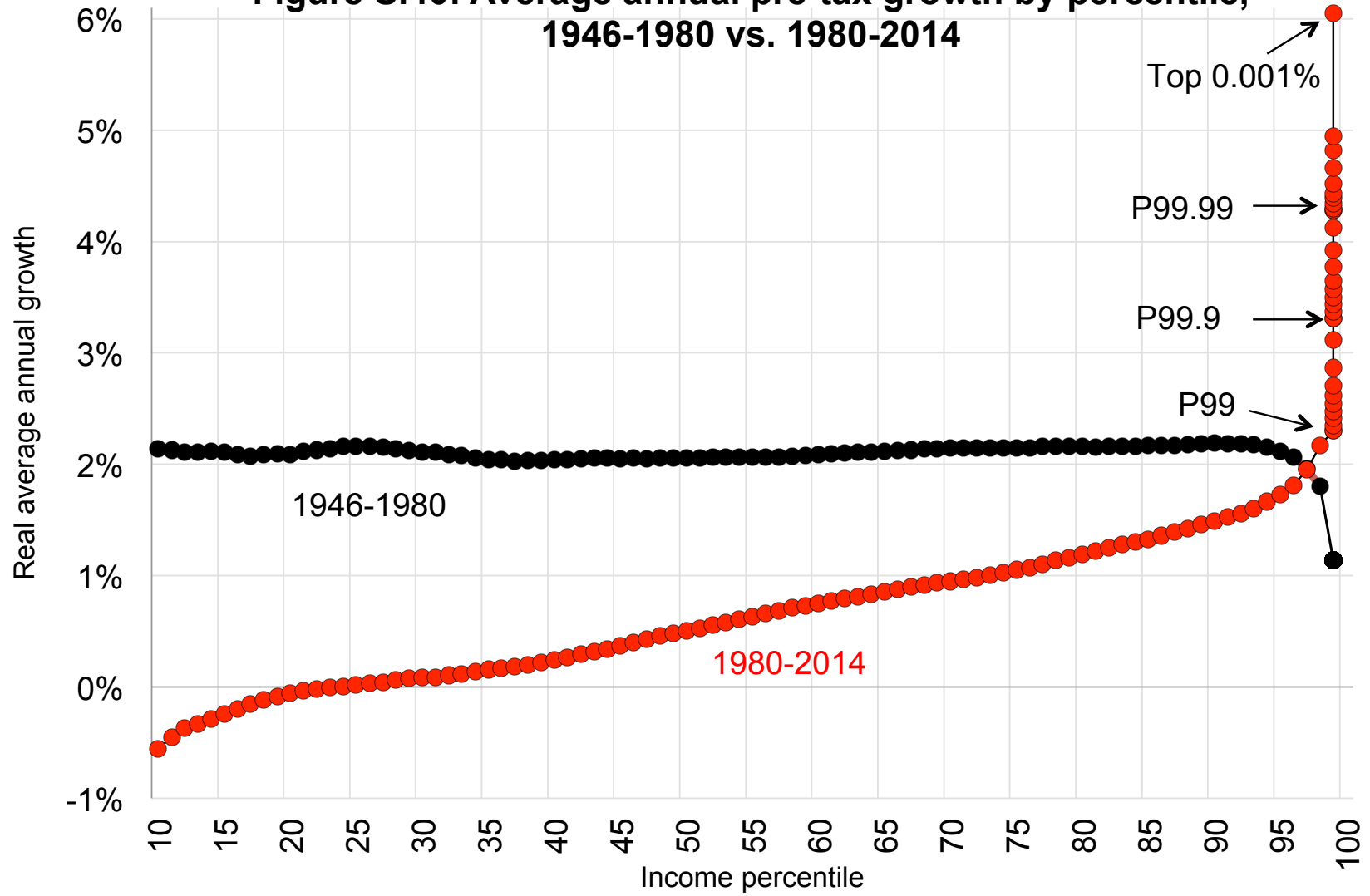
Source: Appendix Tables II-C4.

**Appendix Figure S.39b: P50/P10 and P90/P50 ratios, post-tax income**

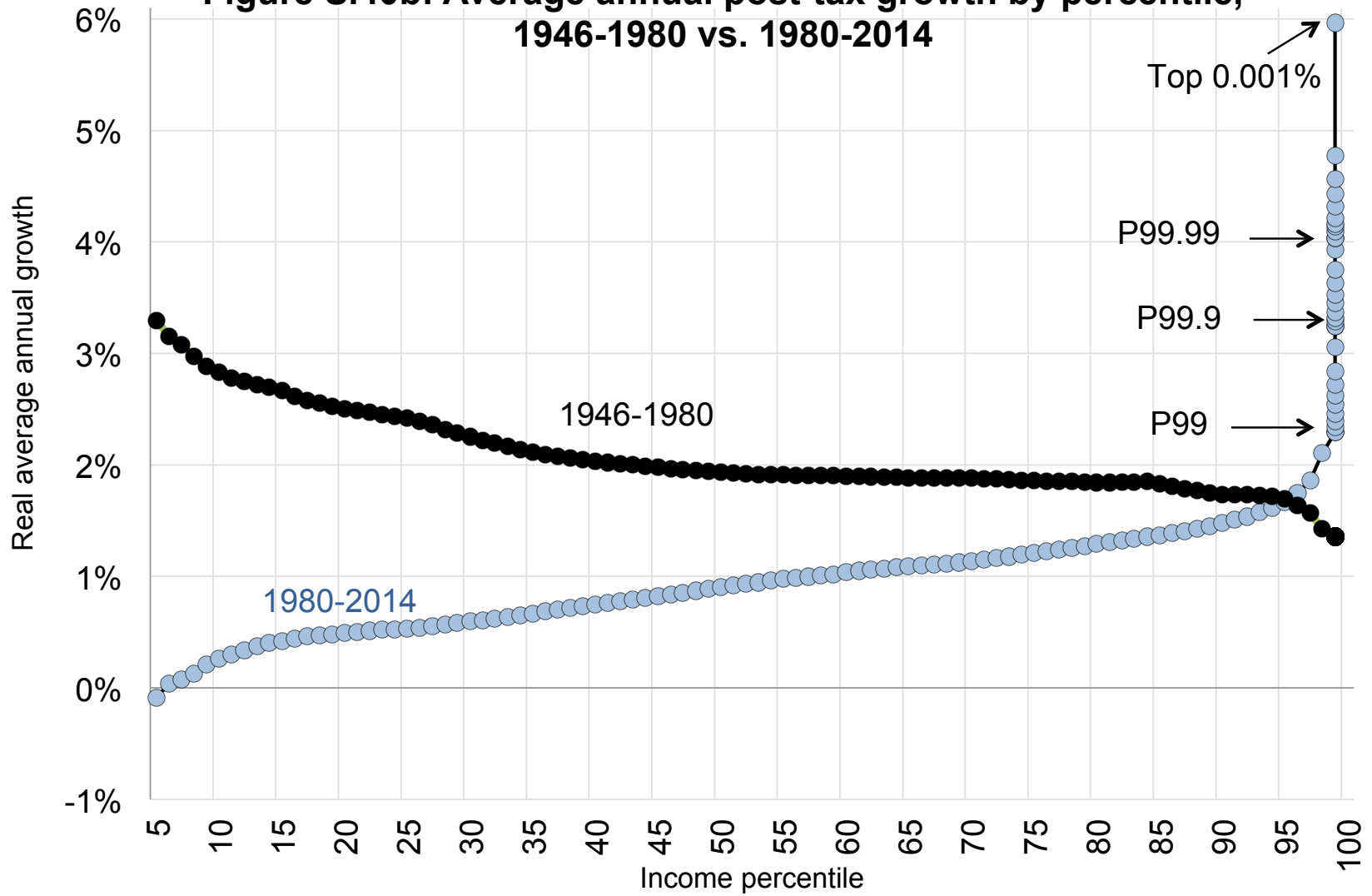


Source: Appendix Tables II-C4.

**Figure S.40: Average annual pre-tax growth by percentile, 1946-1980 vs. 1980-2014**



**Figure S.40b: Average annual post-tax growth by percentile, 1946-1980 vs. 1980-2014**



**Table S.1: Comparison of real growth rates: fiscal vs national income**

	[1]	[2]	[3]	[4]
	<b>Fiscal income per tax unit</b> (consumer price index)	Fiscal income per tax unit (national income deflator)	Fiscal income per adult (national income deflator)	<b>National income per adult</b> (national income deflator) (Our DINA series)
<b>1913-2014</b>	<b>1.2%</b>	1.3%	1.4%	<b>1.7%</b>
<b>1913-1946</b>	<b>1.4%</b>	1.3%	1.1%	<b>1.7%</b>
<b>1946-2014</b>	<b>1.1%</b>	1.4%	1.5%	<b>1.7%</b>
1946-1980	<b>1.7%</b>	1.9%	2.0%	<b>2.0%</b>
1980-2014	<b>0.6%</b>	0.8%	1.1%	<b>1.4%</b>



**Table S.2: Comparison of real growth rates: labor vs capital income**

	[1]	[2]	[3]	[4]	[5]
	<b>National income per adult</b>	Labor income per adult	Capital income per adult	Share of aggregate per-adult income growth attributed to income from...	
				Labor	Capital
<b>1980-2014</b>	<b>1.4%</b>	1.2%	2.0%	60%	40%
1980-1990	<b>1.7%</b>	1.5%	2.1%	67%	33%
1990-2000	<b>2.3%</b>	2.5%	1.8%	79%	21%
2000-2014	<b>0.6%</b>	0.0%	2.0%	6%	94%

**Table S.3: Decomposition of real growth rates of pre-tax national income**

	All	Bottom 90%	Bottom 50%	Next 40%	Top 10%	Top 5%	Top 1%	Top 0.5%	Top 0.1%	Top 0.01%
Average yearly growth rates of pre-tax income per adult (equal-split among spouses)										
<b>1913-2014</b>	<b>1.7%</b>	<b>1.6%</b>			<b>1.8%</b>	1.8%	1.8%	1.7%	1.7%	2.1%
<b>1913-1946</b>	<b>1.7%</b>	<b>2.0%</b>			<b>1.3%</b>	1.4%	0.9%	0.4%	-0.5%	-0.5%
1913-1929	<b>1.3%</b>	<b>0.8%</b>			<b>1.9%</b>	2.2%	2.0%	1.6%	1.3%	3.3%
1929-1946	<b>2.2%</b>	<b>3.2%</b>			<b>0.8%</b>	0.6%	-0.2%	-0.7%	-2.1%	-3.9%
<b>1946-2014</b>	<b>1.7%</b>	<b>1.4%</b>			<b>2.0%</b>	2.1%	2.2%	2.4%	2.8%	3.4%
1946-1980	<b>2.0%</b>	<b>2.1%</b>			<b>1.7%</b>	1.5%	1.1%	1.1%	1.3%	1.7%
1980-2014	<b>1.4%</b>	<b>0.8%</b>	0.0%	1.0%	<b>2.4%</b>	2.6%	3.3%	3.6%	4.3%	5.2%
2009-2014	<b>1.7%</b>	<b>0.7%</b>	0.1%	0.9%	<b>2.9%</b>	3.2%	3.5%	3.5%	3.5%	2.6%
Fraction of pre-tax income growth accruing to each group										
<b>1913-2014</b>	100%						20%	16%	9%	5%
<b>1913-1946</b>	100%						8%	3%	-2%	-1%
1913-1929	100%						32%	20%	10%	9%
1929-1946	100%	<b>84%</b>			<b>16%</b>	9%	-2%	-4%	-7%	-4%
<b>1946-2014</b>	100%	<b>49%</b>			<b>52%</b>	40%	23%	19%	12%	6%
1946-1980	100%	<b>68%</b>			<b>31%</b>	19%	7%	5%	3%	1%
1980-2014	100%	<b>34%</b>	0%	32%	<b>68%</b>	56%	36%	30%	19%	9%
2009-2014	100%	<b>37%</b>	1%	22%	<b>77%</b>	65%	39%	31%	18%	6%

**Table S.3b: Decomposition of real growth rates of pre-tax national income**

	<b>All</b>	<b>Bottom 90%</b>	Bottom 50%	Next 40%	<b>Top 10%</b>	Top 5%	Top 1%	Top 0.5%	Top 0.1%	Top 0.01%
Cumulative growth rates of pre-tax income per adult (equal split among spouses)										
<b>1913-2014</b>	455%	409%			516%		495%	457%	453%	719%
<b>1913-1946</b>	77%	93%			56%		33%	14%	-15%	-16%
1913-1929	22%	13%			35%		37%	28%	23%	67%
1929-1946	<b>45%</b>	<b>70%</b>			<b>15%</b>	11%	-3%	-11%	-2%	-50%
<b>1946-2014</b>	<b>214%</b>	<b>165%</b>			<b>296%</b>	302%	347%	389%	549%	871%
1946-1980	<b>95%</b>	<b>104%</b>			<b>79%</b>	65%	47%	45%	54%	75%
1980-2014	<b>61%</b>	<b>30%</b>	1%	42%	<b>121%</b>	143%	205%	237%	321%	454%
2009-2014	<b>9%</b>	<b>4%</b>	1%	5%	<b>15%</b>	17%	18%	19%	19%	14%
Fraction of pre-tax income growth accruing to each group										
<b>1913-2014</b>	100%						20%	16%	9%	5%
<b>1913-1946</b>	100%						8%	3%	-2%	-1%
1913-1929	100%						32%	20%	10%	9%
1929-1946	100%	<b>84%</b>			<b>16%</b>	9%	-2%	-4%	-7%	-4%
<b>1946-2014</b>	100%	<b>48%</b>			<b>52%</b>	40%	23%	19%	12%	6%
1946-1980	100%	<b>68%</b>			<b>31%</b>	19%	7%	5%	3%	1%
1980-2014	100%	<b>32%</b>	0%	32%	<b>68%</b>	56%	36%	30%	19%	9%
2009-2014	100%	<b>28%</b>	1%	22%	<b>77%</b>	65%	39%	31%	18%	6%

**Table S.4: Decomposition of real growth rates of post-tax national income**

	All	Bottom 90%	Bottom 50%	Next 40%	Top 10%	Top 5%	Top 1%	Top 0.5%	Top 0.1%	Top 0.01%
Average yearly growth rates of post-tax income per adult (equal split among spouses)										
<b>1913-2014</b>	1.7%						1.6%	1.5%	1.5%	2.0%
<b>1913-1946</b>	1.7%						0.2%	-0.5%	-1.7%	-2.2%
1913-1929	1.3%						1.7%	1.1%	0.7%	3.1%
1929-1946	2.2%	<b>3.2%</b>			<b>0.7%</b>	0.0%	-1.2%	-1.9%	-3.9%	-7.0%
<b>1946-2014</b>	1.7%	<b>1.6%</b>			<b>1.9%</b>	2.0%	2.3%	2.5%	3.1%	4.1%
1946-1980	2.0%	<b>2.2%</b>			<b>1.5%</b>	1.5%	1.4%	1.5%	2.1%	3.3%
1980-2014	1.4%	<b>1.0%</b>	0.6%	1.2%	<b>2.2%</b>	2.5%	3.2%	3.5%	4.2%	5.0%
2009-2014	1.7%	<b>1.2%</b>	1.1%	1.3%	<b>2.5%</b>	2.6%	2.4%	2.2%	1.7%	0.3%
Fraction of post-tax income growth accruing to each group										
<b>1913-2014</b>	100%						15%	12%	7%	3%
<b>1913-1946</b>	100%						1%	-3%	-5%	-2%
1913-1929	100%						24%	13%	4%	7%
1929-1946	100%	<b>88%</b>			<b>12%</b>	0%	-8%	-9%	-8%	-5%
<b>1946-2014</b>	100%	<b>59%</b>			<b>41%</b>	32%	18%	14%	9%	4%
1946-1980	100%	<b>75%</b>			<b>25%</b>	16%	6%	5%	3%	1%
1980-2014	100%	<b>45%</b>	9%	36%	<b>55%</b>	44%	27%	22%	14%	7%
2009-2014	100%	<b>44%</b>	13%	31%	<b>56%</b>	43%	22%	16%	7%	1%

**Table S.5a: Distribution of pre-tax vs. post-tax national income in 1962 vs. 2014**

<b>Adult individuals</b>						
	<b>1962</b>			<b>2014</b>		
	Share pre-tax income	Share post-tax income	Implicit net tax rate	Share pre-tax income	Share post-tax income	Implicit net tax rate
P0-50	10.5%	12.3%	-16.8%	10.8%	17.0%	-58.1%
P50-P90	48.7%	50.4%	-3.3%	40.2%	41.8%	-4.1%
Top 10%	40.7%	37.4%	8.3%	49.1%	41.1%	16.1%
Top 1%	13.5%	11.2%	17.3%	21.2%	16.5%	22.1%
Top 0.1%	4.6%	3.5%	24.2%	9.7%	7.2%	26.2%
Top 0.01%	1.7%	1.3%	25.5%	4.5%	3.3%	27.4%

<b>Adults (equal split among spouses)</b>						
	<b>1962</b>			<b>2014</b>		
	Share pre-tax income	Share post-tax income	Implicit net tax rate	Share pre-tax income	Share post-tax income	Implicit net tax rate
P0-50	19.5%	22.5%	-15.4%	12.5%	19.3%	-53.7%
P50-P90	44.4%	45.5%	-2.4%	40.4%	41.6%	-2.9%
Top 10%	36.1%	32.0%	11.3%	47.0%	39.1%	16.8%
Top 1%	12.6%	10.1%	19.9%	20.2%	15.7%	22.4%
Top 0.1%	4.4%	3.3%	25.0%	9.3%	6.8%	26.5%
Top 0.01%	1.7%	1.3%	25.6%	4.4%	3.1%	27.7%

<b>Tax units</b>						
	<b>1962</b>			<b>2014</b>		
	Share pre-tax income	Share post-tax income	Implicit net tax rate	Share pre-tax income	Share post-tax income	Implicit net tax rate
P0-50	16.1%	18.7%	-15.5%	10.4%	16.0%	-53.3%
P50-P90	45.8%	47.1%	-3.0%	38.9%	41.2%	-5.9%
Top 10%	38.1%	34.2%	10.2%	50.7%	42.8%	15.4%
Top 1%	13.1%	10.7%	18.9%	21.8%	17.1%	21.7%
Top 0.1%	4.6%	3.5%	24.4%	10.0%	7.4%	25.9%
Top 0.01%	1.7%	1.3%	25.1%	4.6%	3.4%	27.3%

**Table S.5b: Distribution of factor vs. post-tax national income in 1962 vs. 2014**

<b>Adult individuals</b>						
	<b>1962</b>			<b>2014</b>		
	Share factor income	Share post-tax income	Implicit net tax rate	Share factor income	Share post-tax income	Implicit net tax rate
P0-50	7.9%	12.3%	-56.0%	8.2%	17.0%	-106.9%
P50-P90	50.8%	50.4%	0.8%	41.6%	41.8%	-0.6%
Top 10%	41.4%	37.4%	9.7%	50.2%	41.1%	18.0%
Top 1%	13.5%	11.2%	17.4%	20.8%	16.5%	20.7%
Top 0.1%	4.6%	3.5%	24.2%	9.7%	7.2%	25.7%
Top 0.01%	1.7%	1.3%	25.6%	4.5%	3.3%	27.7%

<b>Adults (equal split among spouses)</b>						
	<b>1962</b>			<b>2014</b>		
	Share factor income	Share post-tax income	Implicit net tax rate	Share factor income	Share post-tax income	Implicit net tax rate
P0-50	17.4%	22.5%	-29.1%	10.3%	19.3%	-87.0%
P50-P90	46.3%	45.5%	1.7%	42.4%	41.6%	2.0%
Top 10%	36.3%	32.0%	11.8%	47.2%	39.1%	17.2%
Top 1%	12.5%	10.1%	19.7%	19.6%	15.7%	20.0%
Top 0.1%	4.4%	3.3%	24.9%	9.2%	6.8%	25.9%
Top 0.01%	1.7%	1.3%	25.7%	4.4%	3.1%	27.8%

<b>Tax units</b>						
	<b>1962</b>			<b>2014</b>		
	Share factor income	Share post-tax income	Implicit net tax rate	Share factor income	Share post-tax income	Implicit net tax rate
P0-50	13.9%	18.7%	-33.8%	8.9%	16.0%	-79.7%
P50-P90	47.6%	47.1%	1.1%	39.9%	41.2%	-3.2%
Top 10%	38.4%	34.2%	11.0%	51.2%	42.8%	16.3%
Top 1%	13.1%	10.7%	18.8%	21.1%	17.1%	19.3%
Top 0.1%	4.6%	3.5%	24.4%	9.8%	7.4%	25.0%
Top 0.01%	1.7%	1.3%	25.3%	4.6%	3.4%	27.4%

**Table S.6: Taxes paid by pre-tax income group**

<b>Taxes paid as a fraction of pre-tax income, 1962</b>							
	All taxes	Sales tax	Residential property tax	Payroll tax	Income tax	Corporate (& business property) tax	Estate tax
All	27.2%	0.9%	1.6%	3.6%	9.6%	6.0%	0.5%
P0-50	19.2%	5.7%	1.5%	5.1%	5.1%	1.8%	0.0%
P50-P90	26.0%	6.5%	1.7%	4.7%	10.2%	2.9%	0.0%
Top 10%	33.2%	5.4%	1.6%	1.4%	11.4%	12.0%	1.3%
Top 1%	39.3%	5.0%	1.3%	0.3%	13.2%	16.3%	3.2%
Top 0.1%	43.6%	4.9%	0.8%	0.1%	13.6%	19.0%	5.1%
Top 0.01%	44.1%	5.0%	0.5%	0.0%	12.9%	20.0%	5.7%
<b>Taxes paid as a fraction of pre-tax income, 2010</b>							
	All taxes	Sales tax	Residential property tax	Payroll tax	Income tax	Corporate (& business property) tax	Estate tax
All	28.2%	2.3%	1.2%	7.7%	9.7%	4.6%	0.2%
P0-50	23.6%	4.7%	1.4%	11.3%	5.7%	0.5%	0.0%
P50-P90	26.9%	5.1%	1.4%	10.2%	7.1%	3.1%	0.0%
Top 10%	30.7%	4.3%	1.1%	4.5%	13.2%	7.2%	0.3%
Top 1%	31.3%	4.0%	0.7%	1.7%	14.7%	9.6%	0.7%
Top 0.1%	32.4%	4.0%	0.4%	0.9%	14.8%	11.4%	1.0%
Top 0.01%	32.0%	4.0%	0.2%	0.5%	13.5%	12.9%	0.9%
<b>Taxes paid as a fraction of pre-tax income, 2014</b>							
	All taxes	Sales tax	Residential property tax	Payroll tax	Income tax	Corporate (& business property) tax	Estate tax
All	30.6%	2.5%	1.1%	7.6%	11.8%	5.0%	0.2%
P0-50	24.4%	5.0%	1.0%	11.4%	6.4%	0.6%	0.0%
P50-P90	28.6%	5.3%	1.2%	10.2%	8.6%	3.3%	0.0%
Top 10%	33.9%	4.4%	1.0%	4.4%	16.0%	7.7%	0.3%
Top 1%	36.4%	4.0%	0.6%	1.7%	18.9%	10.4%	0.8%
Top 0.1%	39.8%	4.0%	0.4%	0.9%	20.7%	12.5%	1.2%
Top 0.01%	40.8%	3.9%	0.3%	0.6%	20.3%	14.4%	1.3%

**Table S.7: The Distribution of National Income in the United States, 2014**

Income group	Number of adults	Factor income		Pre-tax income		Post-tax income	
		Average income	Income share	Average income	Income share	Average income	Income share
Full Population	234,400,000	\$64,600	100%	\$64,600	100%	\$64,600	100%
Bottom 50%	117,200,000	\$13,300	10.3%	\$16,200	12.5%	\$24,900	19.3%
Middle 40%	93,760,000	\$68,600	42.4%	\$65,300	40.4%	\$67,200	41.6%
Top 10%	23,440,000	\$305,000	47.2%	\$304,000	47.0%	\$253,000	39.1%
Top 1%	2,344,000	\$1,270,000	19.6%	\$1,310,000	20.2%	\$1,010,000	15.7%
Top 0.1%	234,400	\$6,000,000	9.2%	\$6,000,000	9.3%	\$4,400,000	6.8%
Top 0.01%	23,440	\$28,200,000	4.4%	\$28,100,000	4.4%	\$20,300,000	3.1%
Top 0.001%	2,344	\$122,300,000	1.9%	\$121,900,000	1.9%	\$88,700,000	1.4%

**Notes:** This table reports statistics on the income distribution in the United States in 2014. Factor, pre-tax and post-tax income match national income. The unit is the adult individual (aged 20 or above). Income is split equally among spouses. Fractiles are defined relative to the total number of adults in the population.



**Table S.7b: The Distribution of National Income in the United States Among the Working-Age Population, 2014**

Income group	Number of working-age adults	Factor income		Pre-tax income		Post-tax income	
		Average income	Income share	Average income	Income share	Average income	Income share
Full Population	188,500,000	\$68,700	100%	\$63,400	100%	\$61,800	100%
Bottom 50%	94,250,000	\$16,700	12.2%	\$16,100	12.7%	\$23,600	19.1%
Middle 40%	75,400,000	\$74,500	43.4%	\$65,100	41.1%	\$65,000	42.1%
Top 10%	18,850,000	\$305,000	44.4%	\$292,000	46.1%	\$240,000	38.8%
Top 1%	1,885,000	\$1,220,000	17.7%	\$1,230,000	19.5%	\$950,000	15.4%
Top 0.1%	188,500	\$5,600,000	8.1%	\$5,600,000	8.8%	\$4,100,000	6.6%
Top 0.01%	18,850	\$26,000,000	3.8%	\$25,900,000	4.1%	\$18,500,000	3.0%

**Notes:** This table reports statistics on the income distribution in the United States among the working-age population (aged 20 to 64) in 2014. The unit of observation is the adult individual (aged 20 to 64). Income is split equally among spouses. Fractiles are defined relative to the total number of working-age adults in the population.

**Table S.8: The Growth of National Income in the United States since World War II**

Income group	Factor income growth		Pre-tax income growth		Post-tax income growth	
	1980-2014	1946-1980	1980-2014	1946-1980	1980-2014	1946-1980
Full Population	61%	95%	61%	95%	61%	95%
Bottom 50%	-4%	95%	1%	102%	21%	129%
Middle 40%	41%	106%	42%	105%	49%	98%
Top 10%	120%	82%	121%	79%	113%	69%
Top 1%	196%	49%	204%	47%	194%	58%
Top 0.1%	317%	57%	320%	54%	298%	104%
Top 0.01%	454%	79%	453%	76%	423%	201%
Top 0.001%	636%	60%	636%	57%	616%	163%

**Notes:** The table displays the cumulative real growth rates of pre-tax and post-tax national income per adult over two 34 years period: 1980 to 2014 and 1946 to 1980. Pre-tax and post-tax income match national income. The unit is the adult individual (aged 20 or above). Fractiles are defined relative to the total number of adults in the population. Income is split equally among spouses.

**Table S.8b: The Growth of National Income in the U.S. Since 1980: Working-Age vs. All Adults**

Income group	Factor income growth		Pre-tax income growth		Post-tax income growth	
	Working-age	All adults	Working-age	All adults	Working-age	All adults
Full Population	61%	61%	57%	61%	55%	61%
Bottom 50%	-4%	-4%	-6%	1%	13%	21%
Middle 40%	47%	41%	40%	42%	45%	49%
Top 10%	124%	120%	122%	121%	110%	113%
Top 1%	200%	196%	206%	204%	187%	194%
Top 0.1%	325%	317%	326%	320%	289%	298%
Top 0.01%	464%	454%	462%	453%	413%	423%

**Notes:** The table displays the cumulative real growth rates of factor, pre-tax and post-tax national income per working-age adult over 1980 to 2014. The unit is the working-age adult individual (aged 20 to 64) or the adult individual (aged above 20). Fractiles are defined relative to the total number of working-age adults or total number of adults in the population. Income is split equally among spouses.

**Table S.9: The Distribution of National Income in the United States, 2014**

Income group	Number of adults	Pre-tax national income		Post-tax disposable income		Post-tax national income	
		Average income	Income share	Average income	Income share	Average income	Income share
Full Population	234,400,000	\$64,600	100%	\$46,500	100%	\$64,600	100%
Bottom 50%	117,200,000	\$16,200	12.5%	\$16,600	17.8%	\$24,900	19.3%
Middle 40%	93,760,000	\$65,300	40.4%	\$47,900	41.2%	\$67,200	41.6%
Top 10%	23,440,000	\$304,000	47.0%	\$190,000	40.9%	\$253,000	39.1%
Top 1%	2,344,000	\$1,310,000	20.2%	\$790,000	17.0%	\$1,010,000	15.7%
Top 0.1%	234,400	\$6,000,000	9.3%	\$3,700,000	8.0%	\$4,400,000	6.8%
Top 0.01%	23,440	\$28,100,000	4.4%	\$17,300,000	3.7%	\$20,300,000	3.1%
Top 0.001%	2,344	\$121,900,000	1.9%	\$75,400,000	1.6%	\$88,700,000	1.4%

**Notes:** This table reports statistics on the income distribution in the United States in 2014. Pre-tax and post-tax income match national income. Post-tax disposable income excludes in-kind transfers, public goods consumption, and the government deficit. The unit is the adult individual (aged 20 or above). Income is split equally among spouses. Fractiles are defined relative to the total number of adults in the population.