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

Thomas Piketty, Emmanuel Saez
and Gabriel Zucman

December 2016

WORLD
WEALTH & INCOME
DATABASE
THE SOURCE FOR
GLOBAL INEQUALITY DATA

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)

December 15, 2016

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.

Contents

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 proving 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".[1] Both the NIPAs and Financial Accounts are frequently updated; we regularly

---

[1]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.[2]

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.[3] 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[4]); 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 no exist in the NIPAs (such as the decomposition of national income by sector[5]). 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).

[2]Downloading the raw national accounts series and integrating them into Appendix Tables I (Macro) is done by the program scrap_macro.do.

[3]The same historical income and wealth series were used by Saez and Zucman (2016).

[4]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).

[5]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.[6]

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.[7] 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.[8] The most recent file available is for 2014. Typically, the file for year t becomes available around July of

---

[6]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).

[7]There are no files for 1963 and 1965. A file for 1960 exists but with much fewer variables.

[8]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 renames 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.[9] 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

---

[9]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 programuse_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 then 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 sharefbuild.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).[10] 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).

---

[10]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 in 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 dweght) 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

- **dweght**: Population weight × 100,000

- **dweghttaxu**: 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 ¡20)

- **filer**: Tax filer dummy

2. Core income and wealth series

- **fiinc**: 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 = flinc + fkinc

- **flinc**: 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 = flinc + 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

- **fikgi**: 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

- **flmil**: 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

- **fknmo**: 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. (alloced 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 (allocted prop.

- **prisupgov**: Government primary surplus (= taxes - benefits) (alloced 50% prop. to taxes, 50

- **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

16

- **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 assitance 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.[11] 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.[12]

---

[11]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.

[12]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).[13] 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.

---

[13]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.[14] Our imputation is designed to respect the correlation of benefits recipiency with income, marital, and elderly status.

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

---

[14]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.[15]

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

---

[15]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 × marital status × 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 × marital status × number of kids × 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.[16] 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.

---

[16]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 around 25% × 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.[17]

CBO assumes that corporate taxes fall 75% on capital and 25% on labor (CBO earlier assumed that corporate taxes fell 100% on capital).[18] 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

---

[17]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.

[18]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 that 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.[19]

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.

---

[19]A more sophisticated possibility would be to allocate the deficit based on age × taxes paid × transfers received (rather than taxes paid × 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 → 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

29

income reinvested in pension plans.

An economy with funded pensions naturally tends to have contributions greater than distributions (see Alvaredo 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 β_P is the steady-state pension wealth-factor income ratio), one can immediately see that contributions (including accrued investment income) exceed pension distributions by $g \cdot W_{Pt}$.[20] 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.[21] Second, we allocate the (very small) primary surplus of the private pensions system proportionally to wages. Last, we re-attribute the capital

---

[20]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.

[21]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.[22] 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 (we discuss below the issue of transfers and post-tax income).

Second, top income shares studies so far were computed using the tax unit (the family in the United States, i.e. married couples or single individuals), with no adjustment at all for size. This can create various biases in levels and evolutions, e.g. an increase in the fraction of single individuals can artificially lead to an increase in inequality. So we prefer in our benchmark *equal-split* series to divide by two the income of married couples.

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 (which we will compare later in the paper, so that one can very easily assess the effect of the unit of observation on inequality trends). We prefer however not to use explicit non-linear equivalence scales, since they are somewhat arbitrary and introduce non-linearities in our growth decomposition exercises that are difficult to justify.

Third, the family tax unit (as used, e.g., in Piketty and Saez 2003, and Saez and Zucman

---

[22]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).

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.[23] Similarly, the choice of the household as unit which is done in official Census based measures derived from the Current Population Survey (see US Census 2015) is problematic as a trend toward larger households (e.g. children staying with their parents longer) increases income growth per household. Against this backdrop, the most sensible choice for the study seems to use the adult as our unit.

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.[24] 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.

---

[23]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).

[24]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.

Appendix References

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.

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.

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.
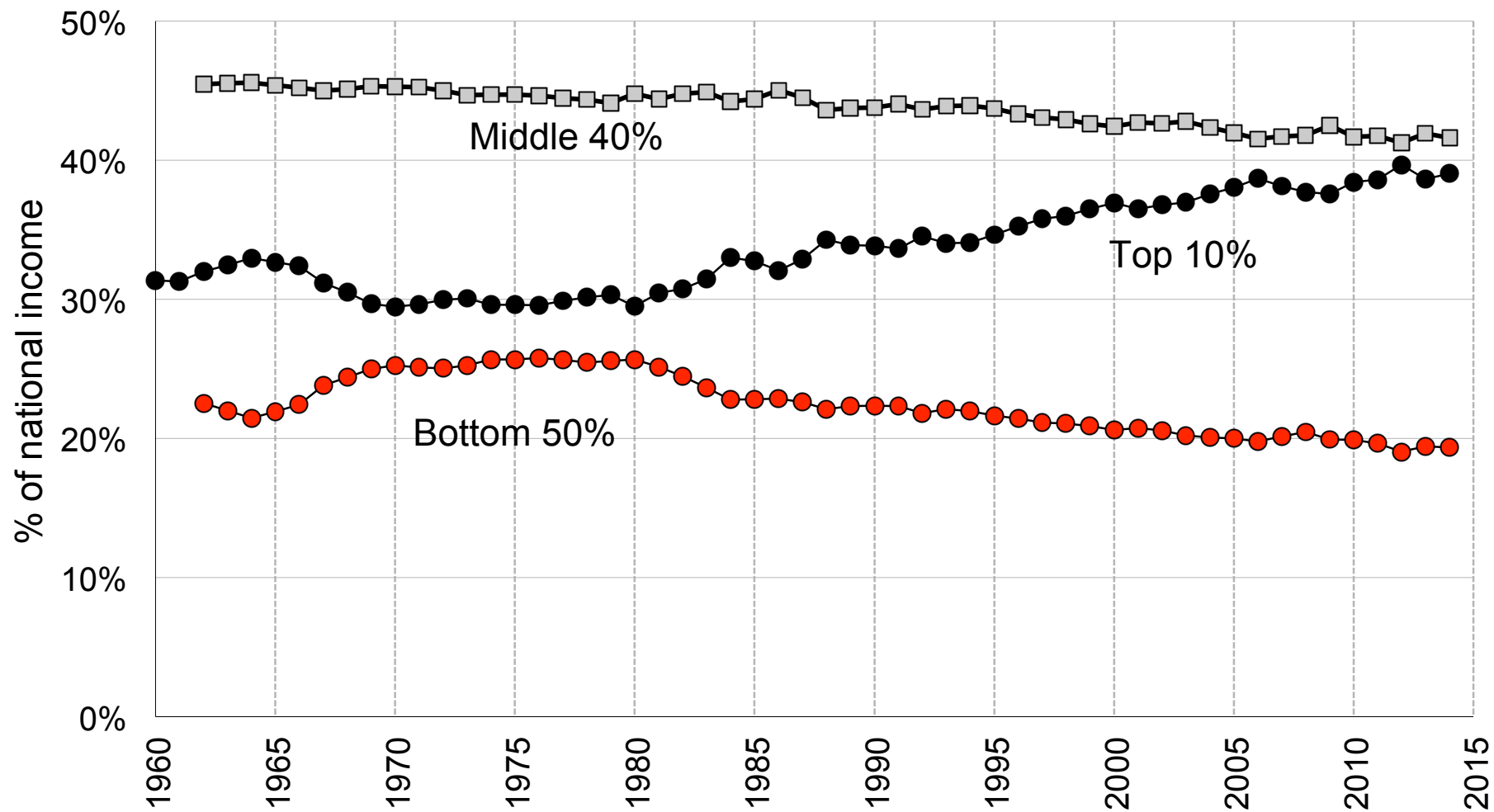
Saez, Emmanuel and Gabriel Zucman. 2016. "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data," *Quarterly Journal of Economics* 131(2), 519-578.

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)

Top 10%

Middle 40%

Bottom 50%

% of national income
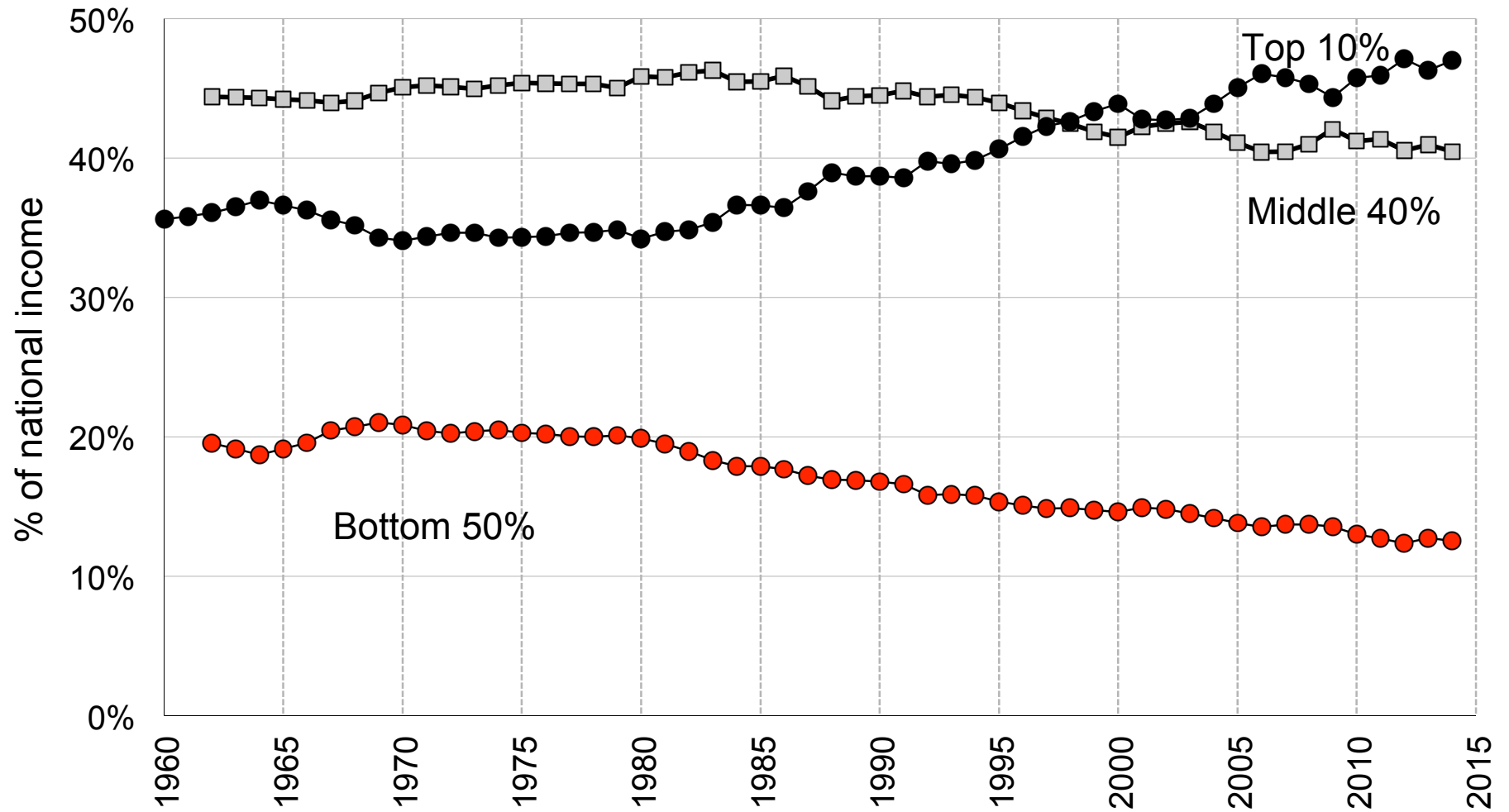
Source: Table II-B1

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

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

Transfers (in cash + in-kind + collective expenditure)

Post-tax income excluding transfers

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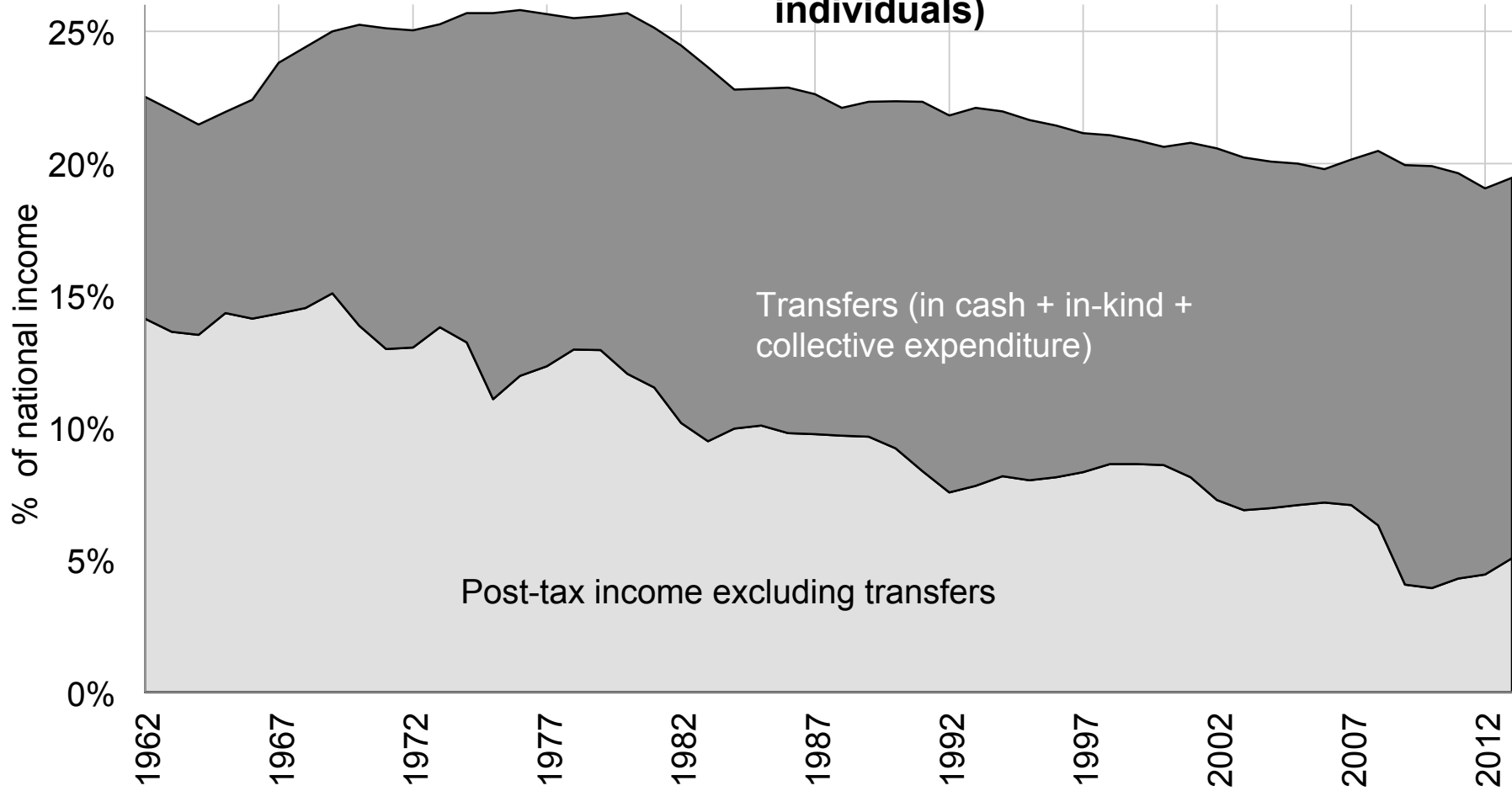
Figure S.4: Real post-tax income of the bottom 50%: the role of transfers

Post-tax income

Transfers

Post-tax income excluding transfers

Average income in constant 2014 dollars

25,000

20,000

15,000

10,000

5,000

0

1962 1966 1970 1974 1978 1982 1986 1990 1994 1998 2002 2006 2010 2014
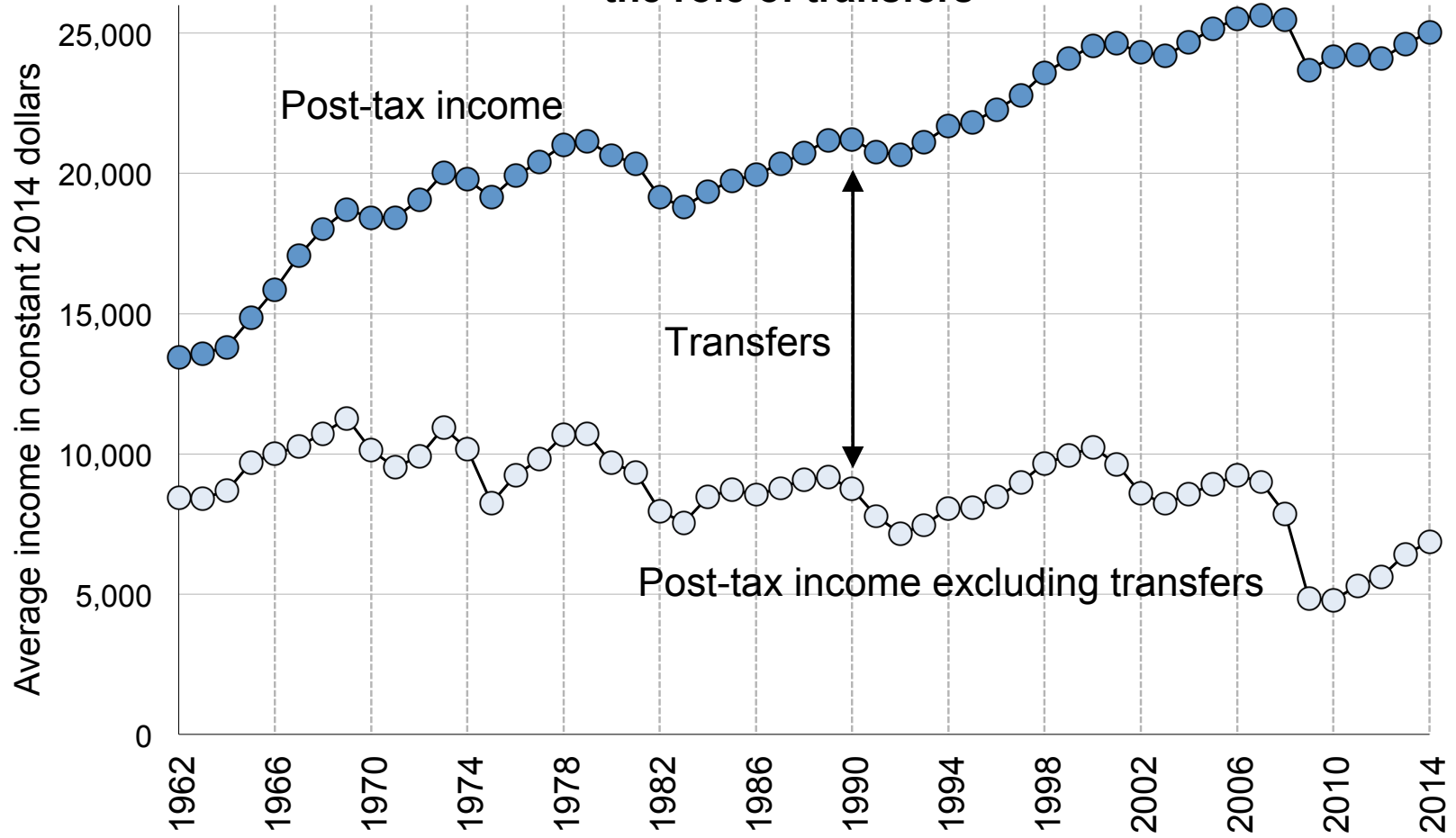
Source: Table II-C3b.

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

Post-tax income

Medicare + Medicaid

Post-tax income excluding health benefits

Average income in constant 2014 dollars
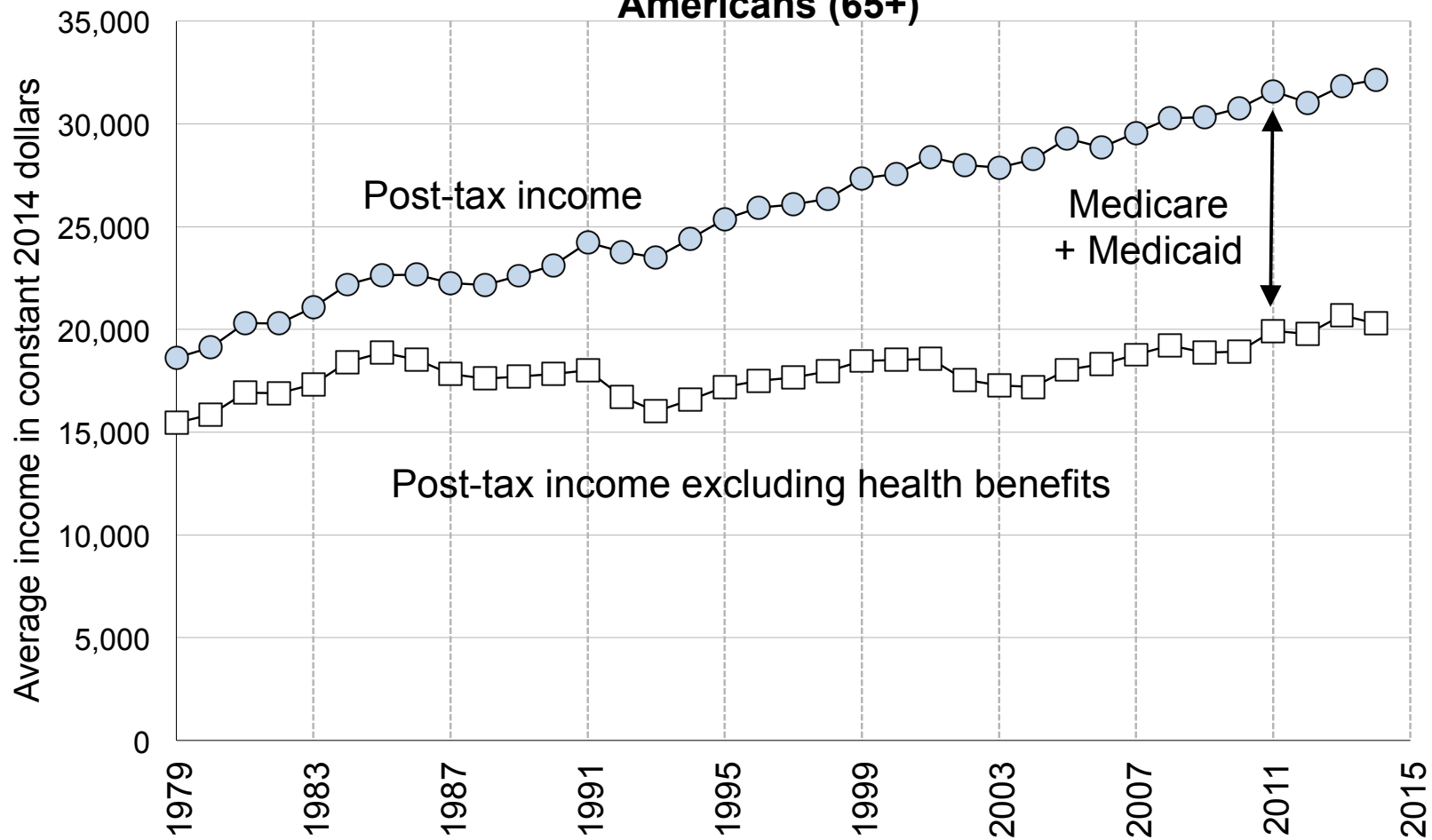
Figure S.6: Real average pre-tax income of bottom 90% and top 1% adults

Top 1%
(left axis)

Bottom 90%
(right axis)

Top 1% real average pre-tax income (2014$)

Bottom 90% real average pre-tax income (2014$)
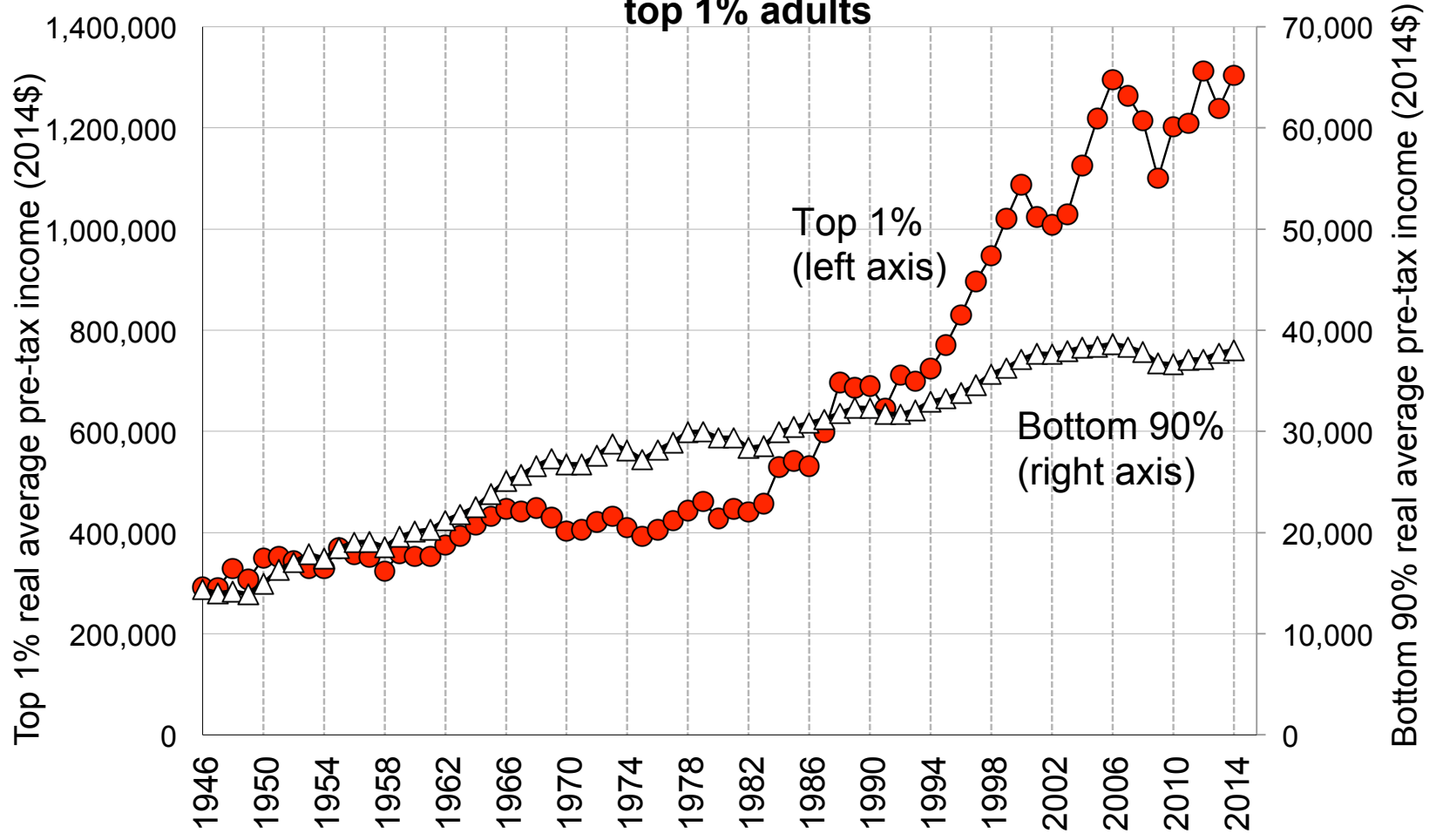
Source: Appendix Table II-B3

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

All

45 to 64

20 to 44
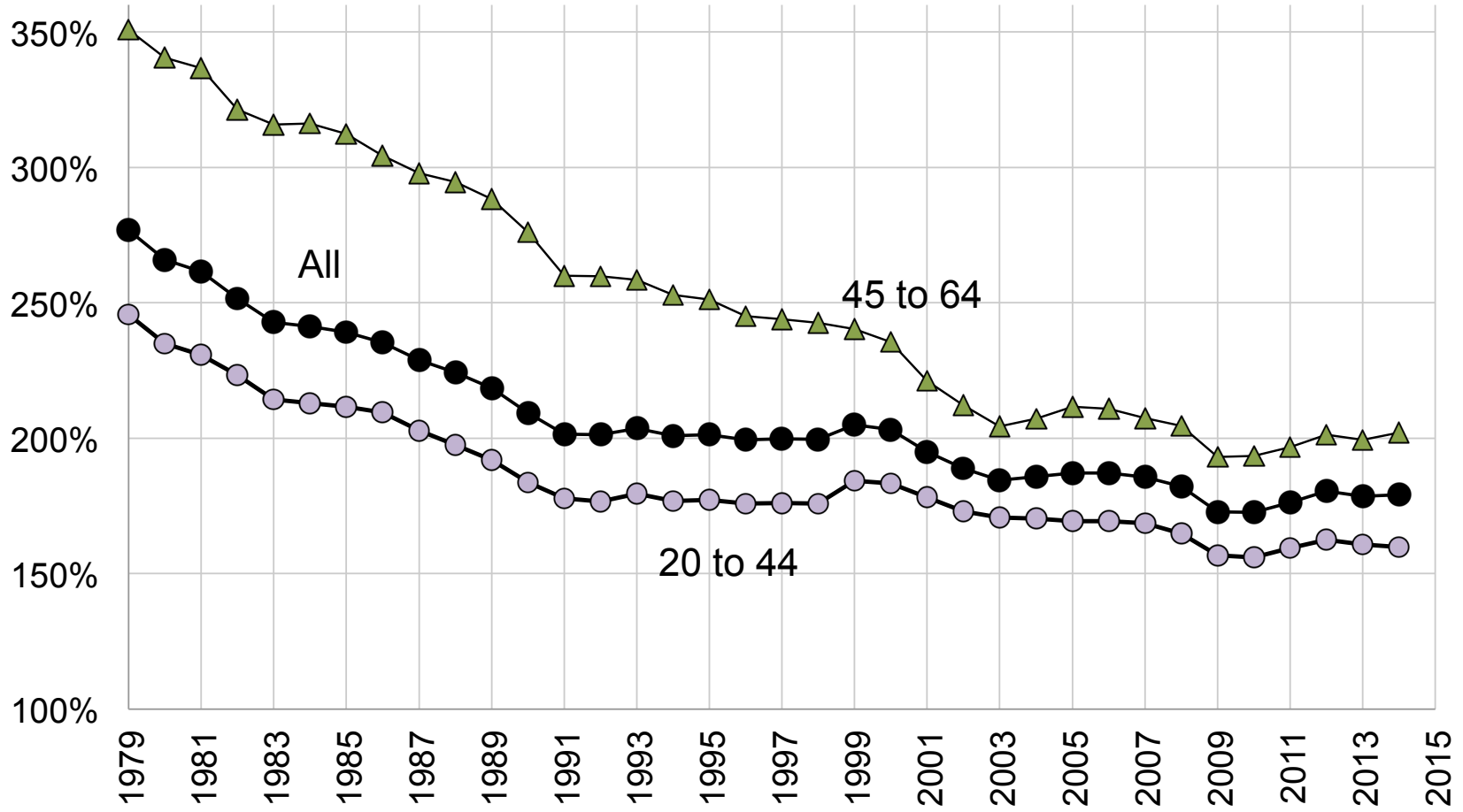
Source: Appendix Table F1.

Figure S.8: Median pre-tax vs. post-tax income of men
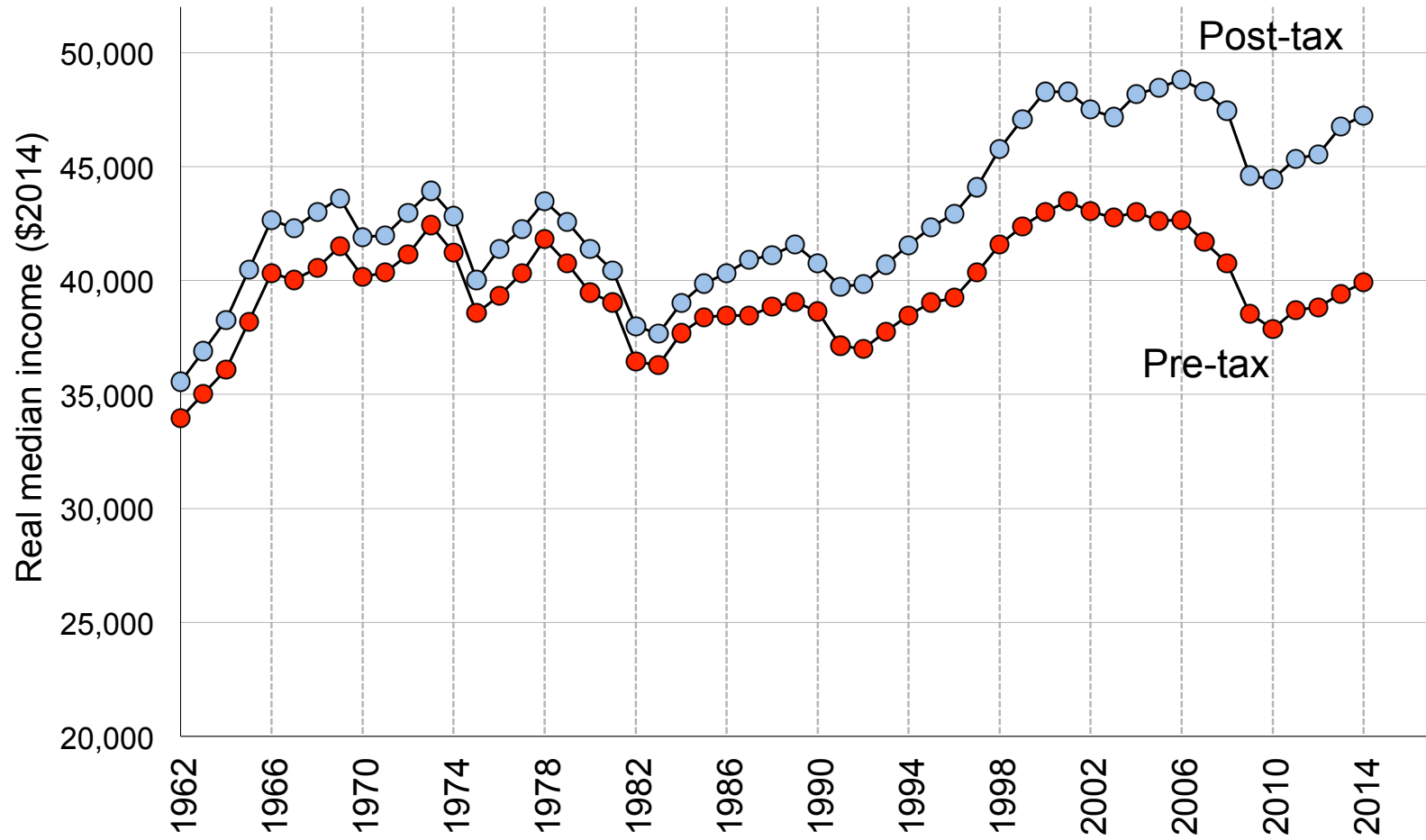
Source: Appendix Table II-B13 and II-C13.

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

Average income in constant 2014 $

Equal split

Individualized
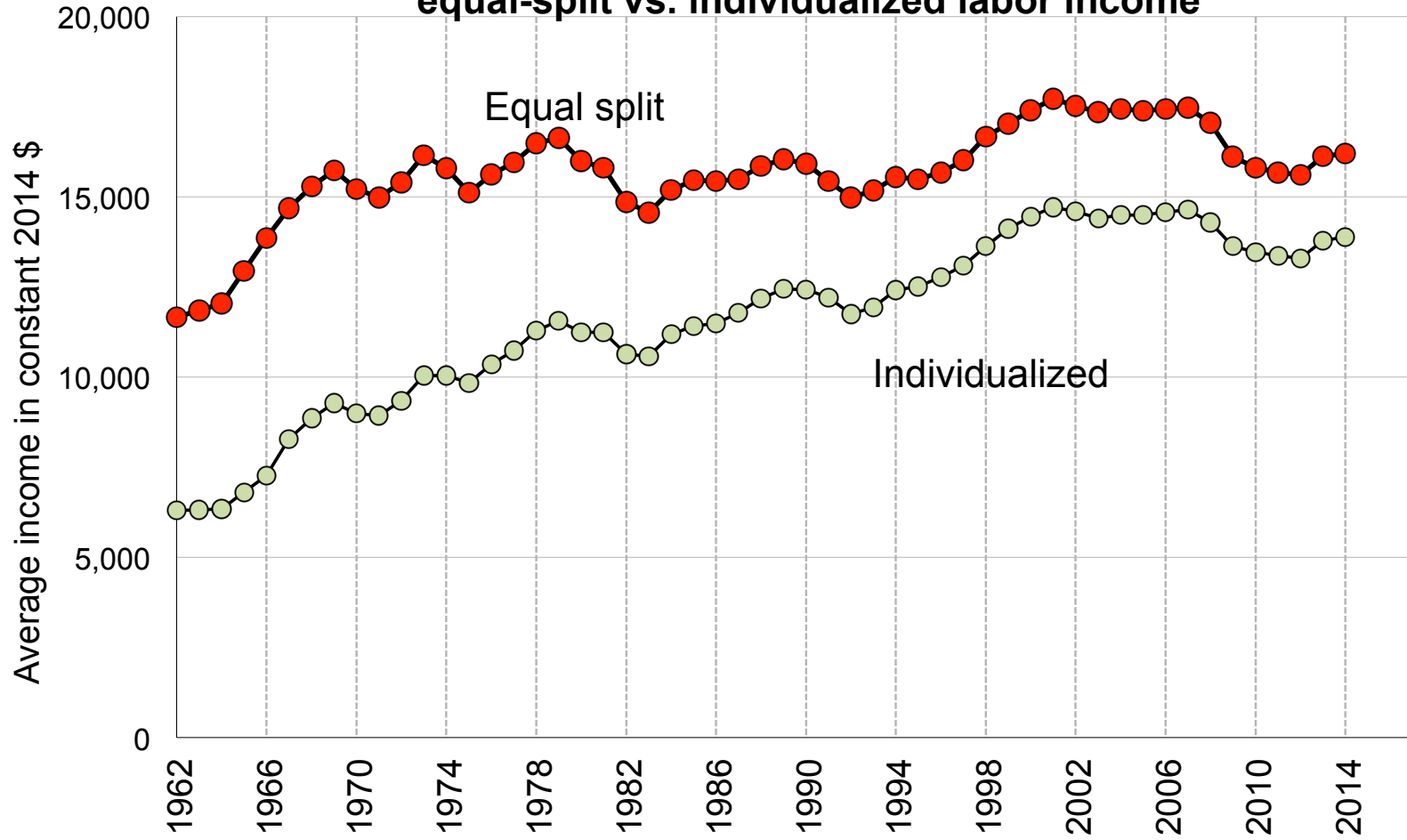
Source: Appendix Tables II-B7.

Figure S.10: The share of capital in pre-tax income, assuming constant 5% return on capital
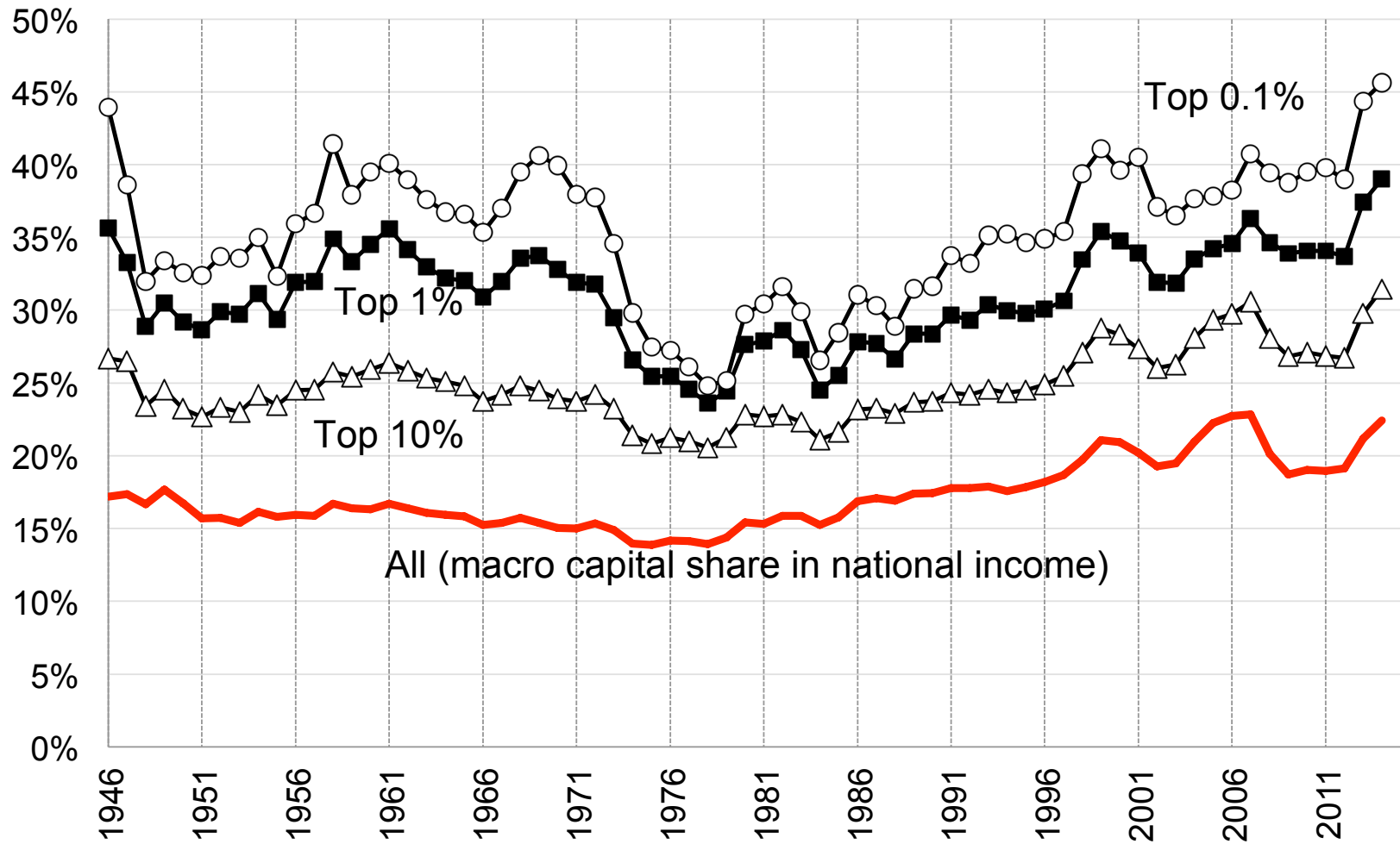
Figure S.11: Top 0.1% pre-tax national income share by age group
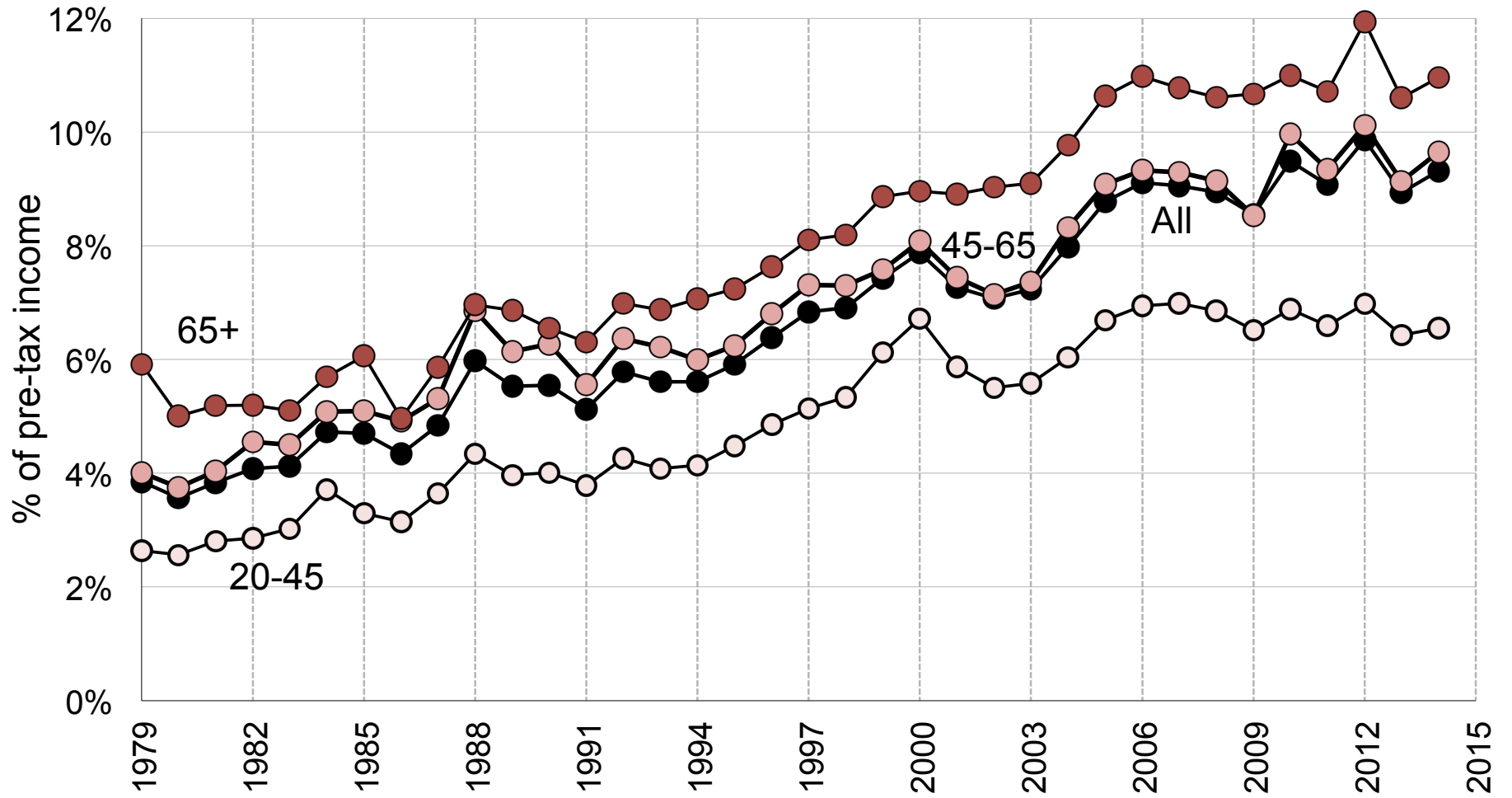
Figure S.12: Average transfers: individualized vs. collective consumption expenditure

Collective consumption expenditure

Individualized transfers (cash + in-kind)
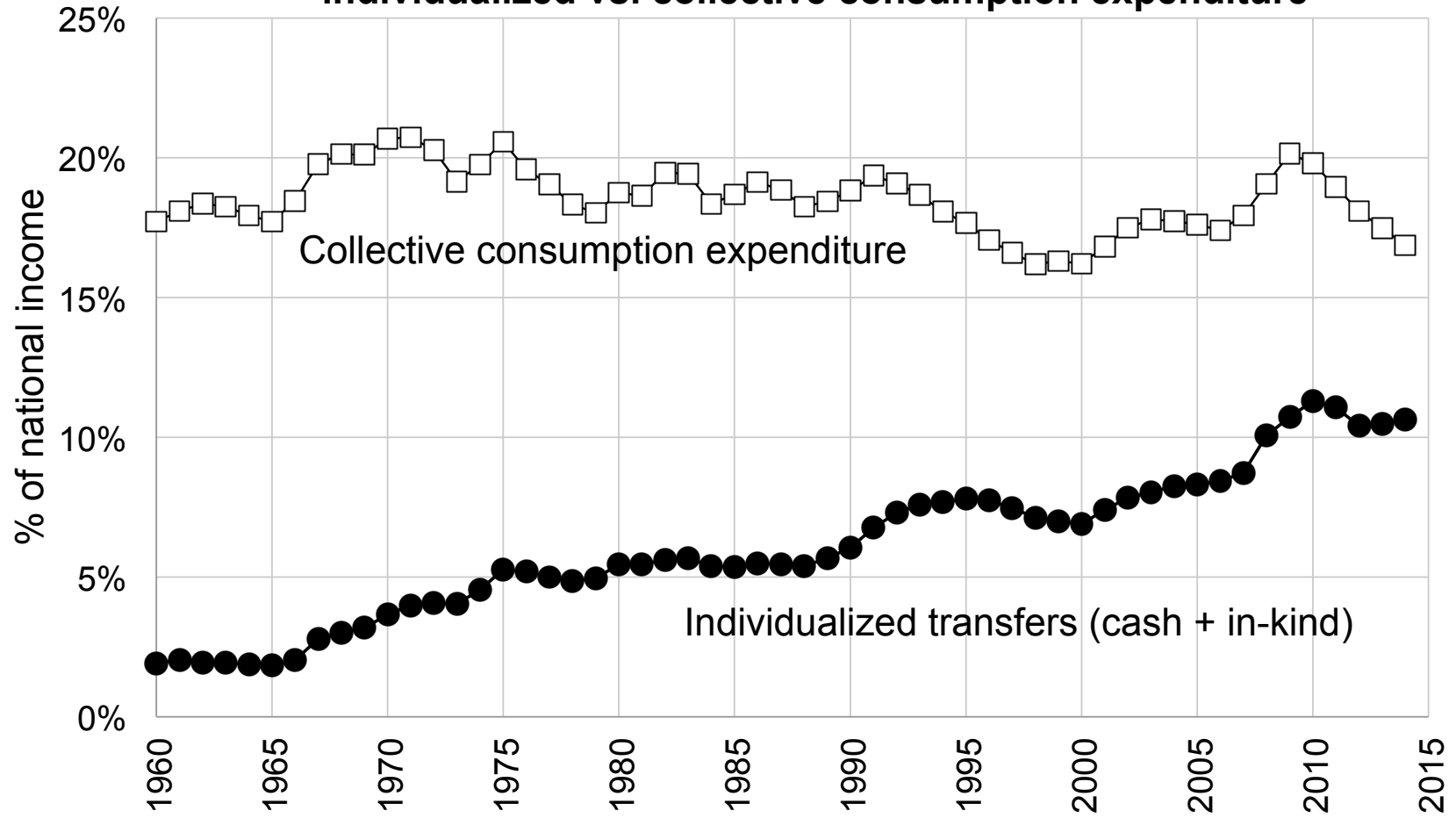
Source: Appendix Table G4

Figure S.13: Average individualized transfer by post-tax income group (including Social Security)
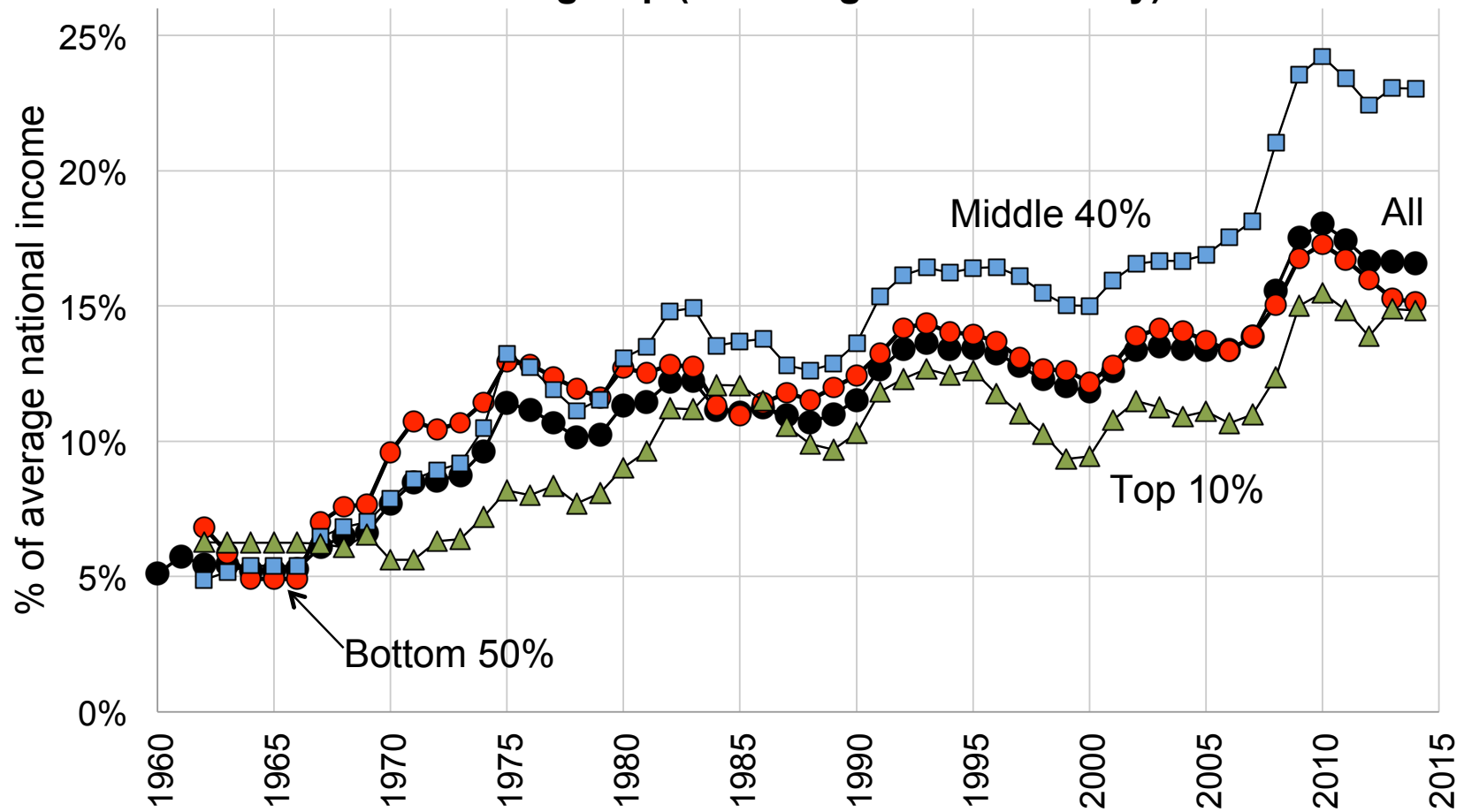
Source: Appendix Table II-G4b.

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

Middle 40%

All

Top 10%

Bottom 50%

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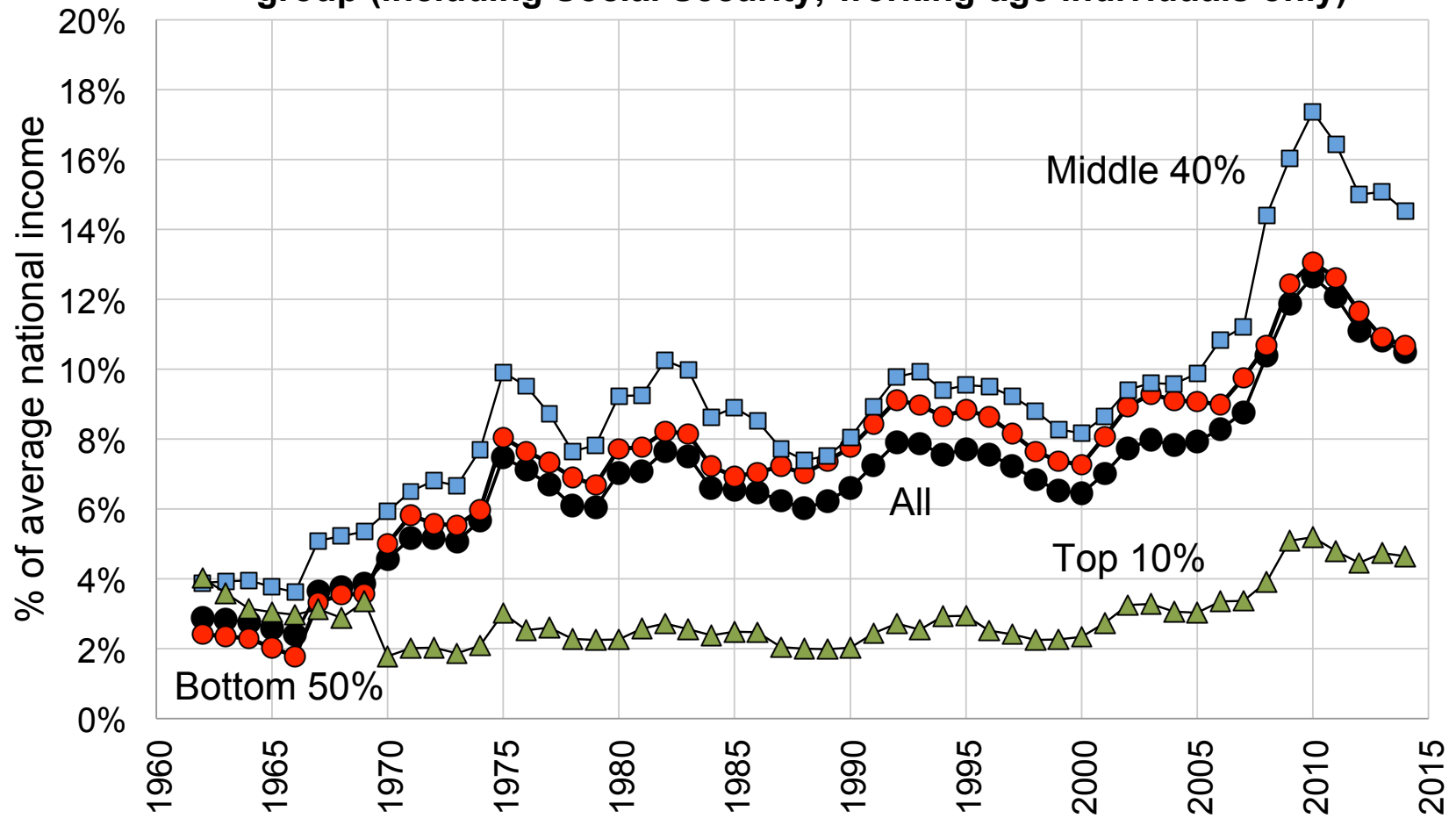
Figure S.15a: Top 10% income share: comparison of estimates

Fiscal income per tax unit
(Piketty-Saez-Zucman)

Fiscal income per tax unit
(Piketty-Saez)
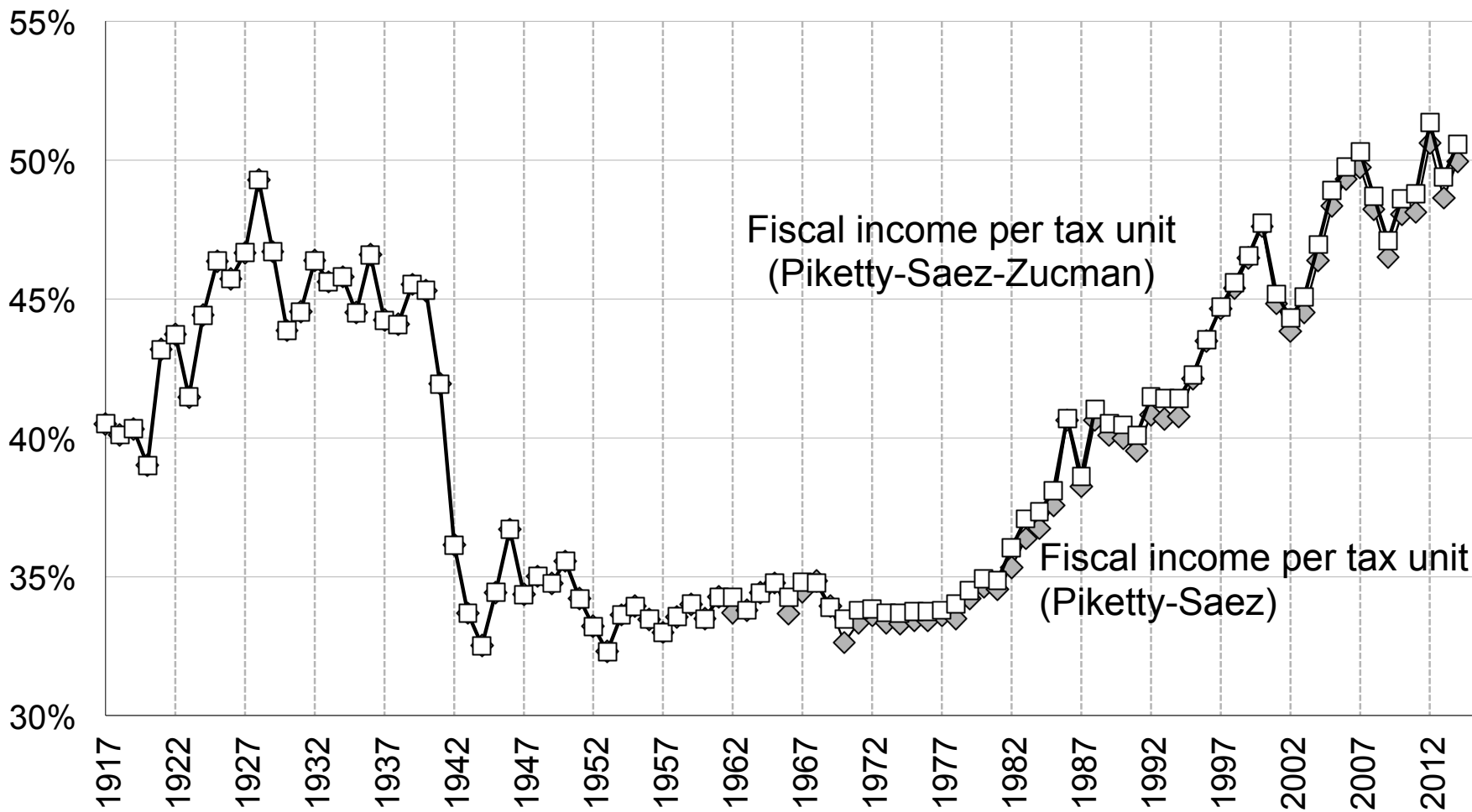
Source: Appendix Tables II-D1 and Piketty and Saez (2003, updated to 2014).

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

Pre-tax national income per tax unit

Pre-tax national income per adult (equal split)
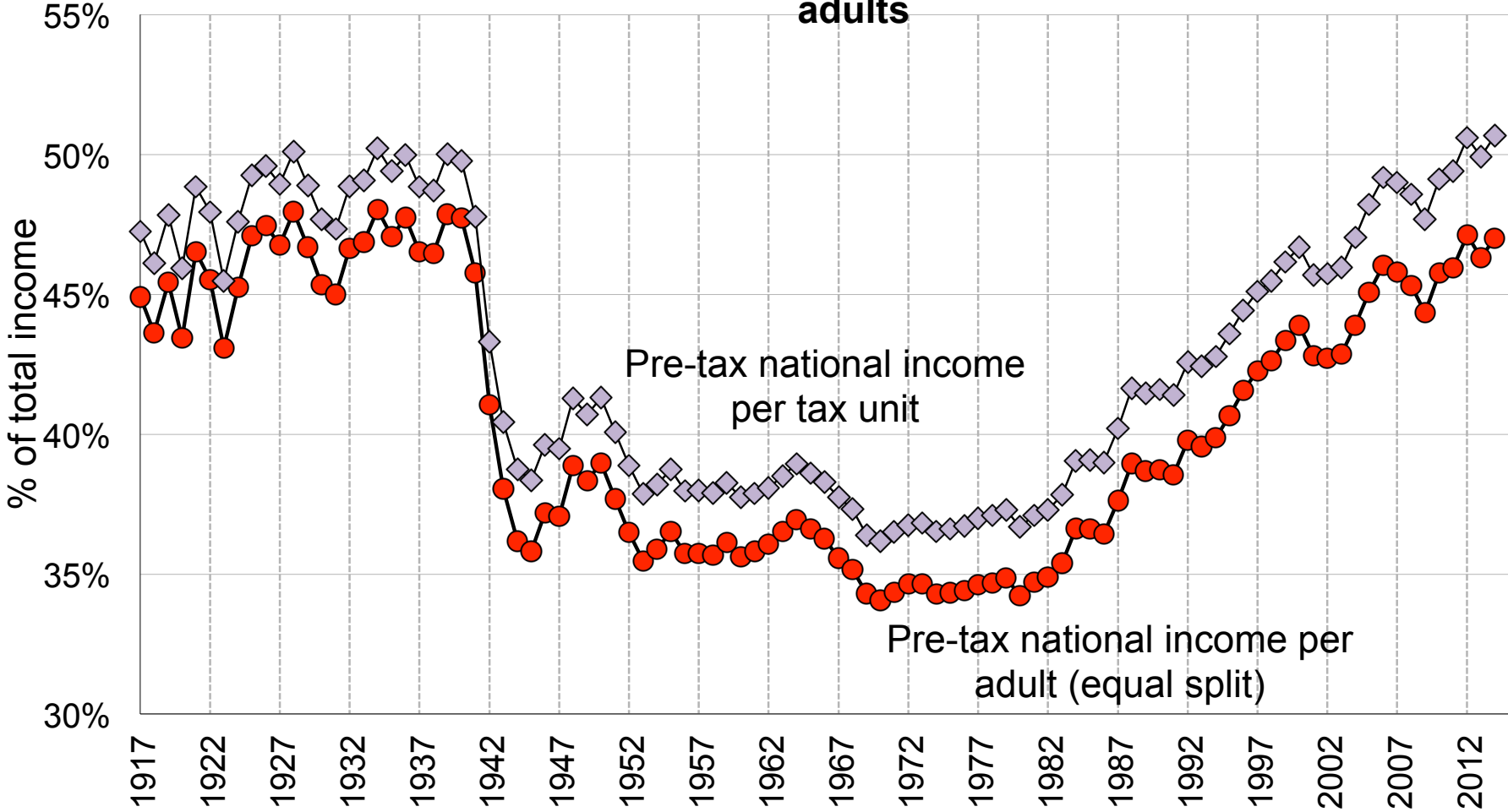
Source: Appendix TablesII- B1 and II-D1.

Figure S.16: Real average national income: Full adult population vs. bottom 90%
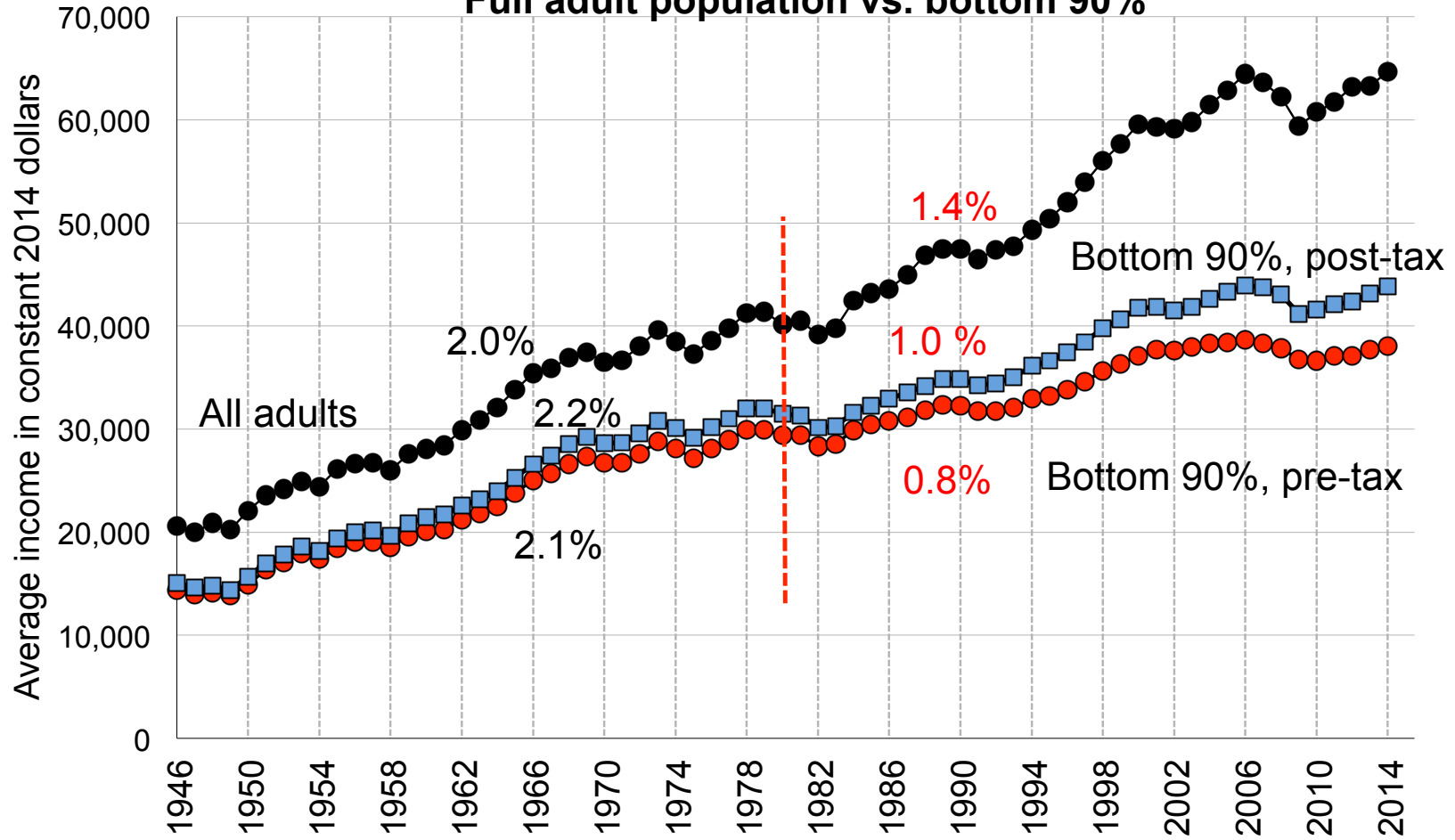
Source: Appendix Table II-B3 and II-C3

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

Average income in constant 2014 dollars

70,000

60,000

50,000

40,000

30,000

20,000

10,000

0

2.0%

All adults

1.4%

Bottom 90% working-
age, post-tax

0.9%

0.7%

Bottom 90% pre-tax
working-age

1946 1950 1954 1958 1962 1966 1970 1974 1978 1982 1986 1990 1994 1998 2002 2006 2010 2014
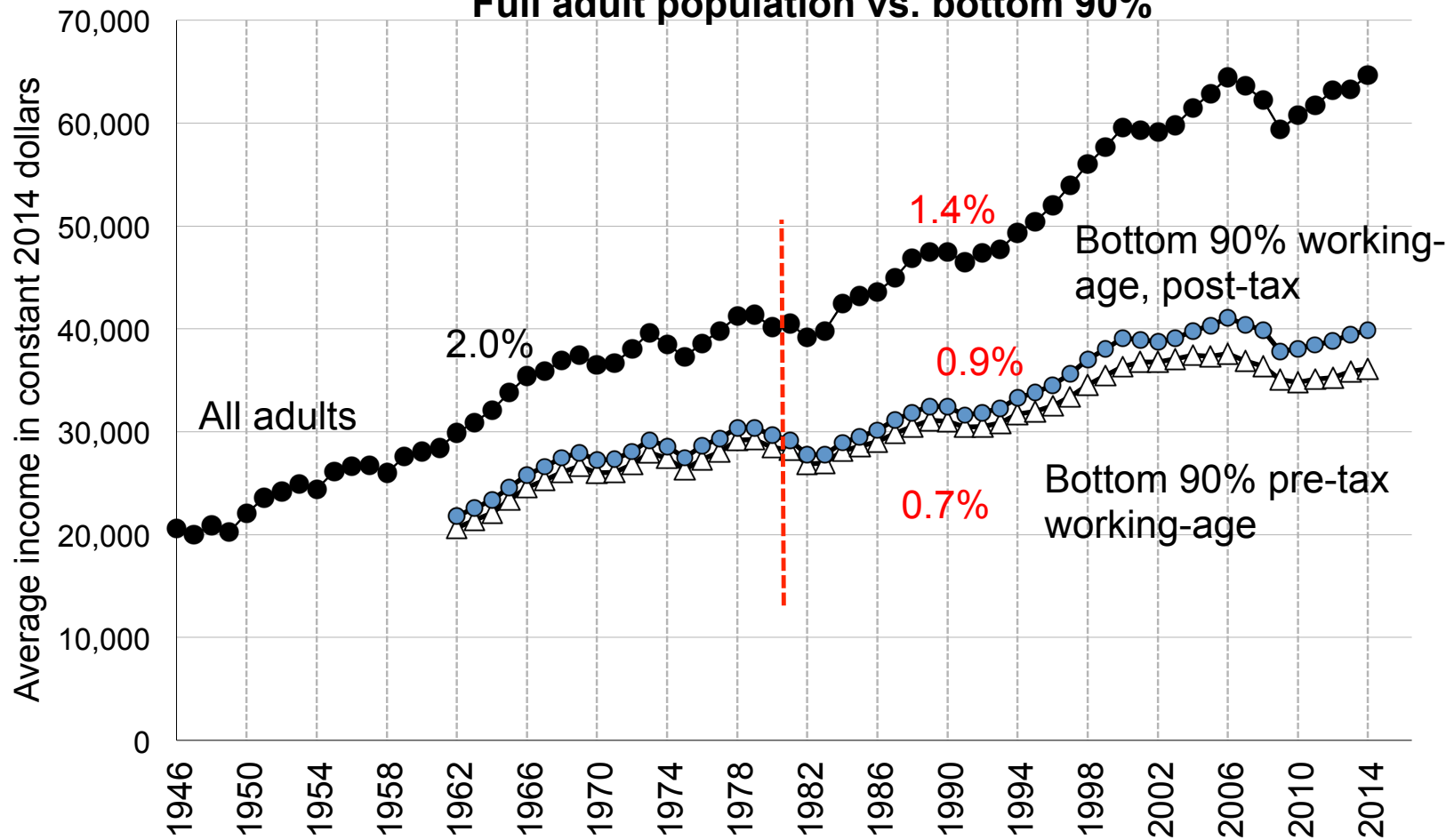
Source: Tables II-B6 and II-C6.

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

National income per adult

CPS income per household (CPI)

Fiscal income per tax unit (CPI)
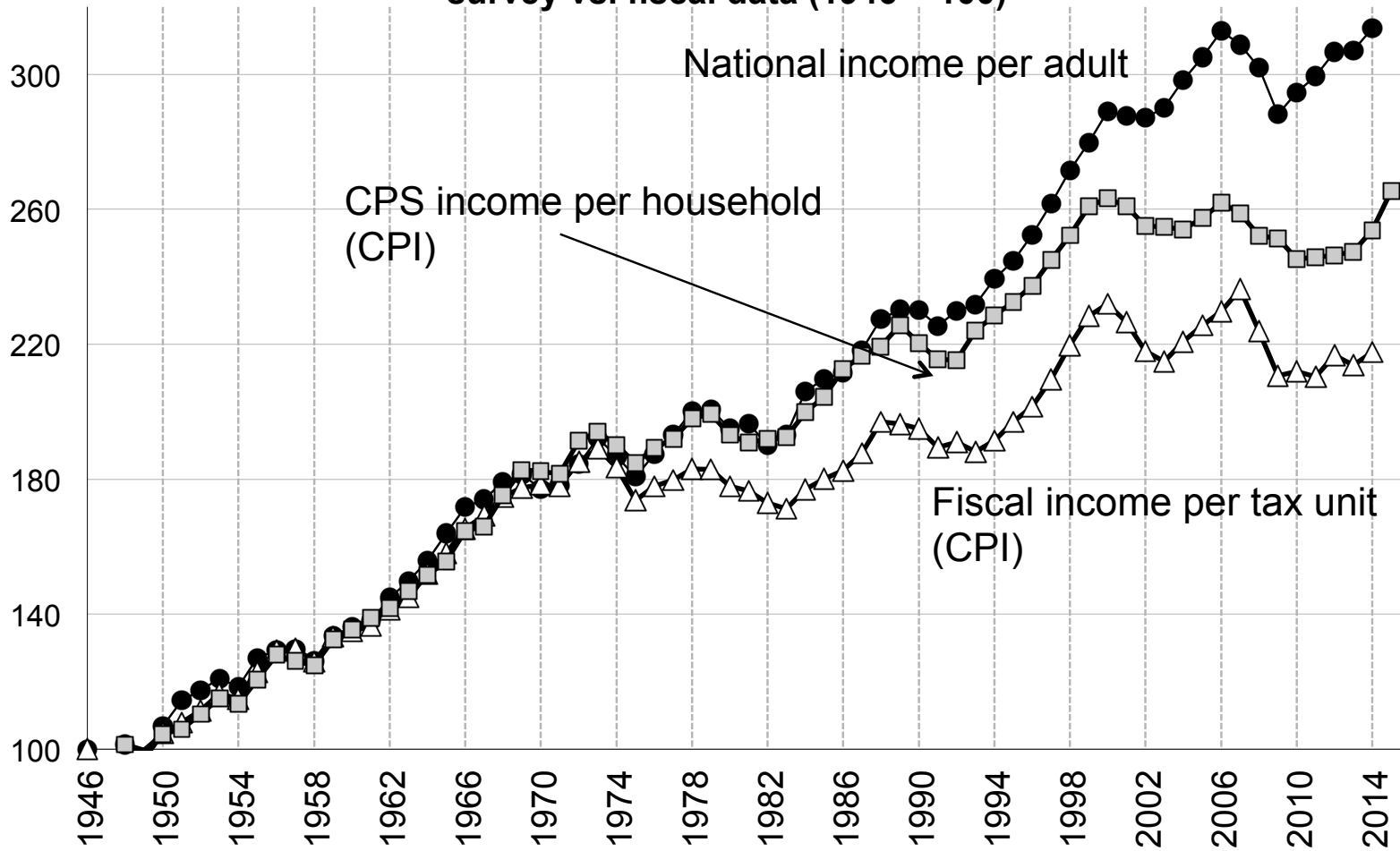
Source: Appendix Table I-A0 and Census Bureau (2016).

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

National income per adult

Fiscal income per adult
(national income deflator)

Fiscal income per tax unit
(national income deflator)

Fiscal income per tax unit
(CPI)
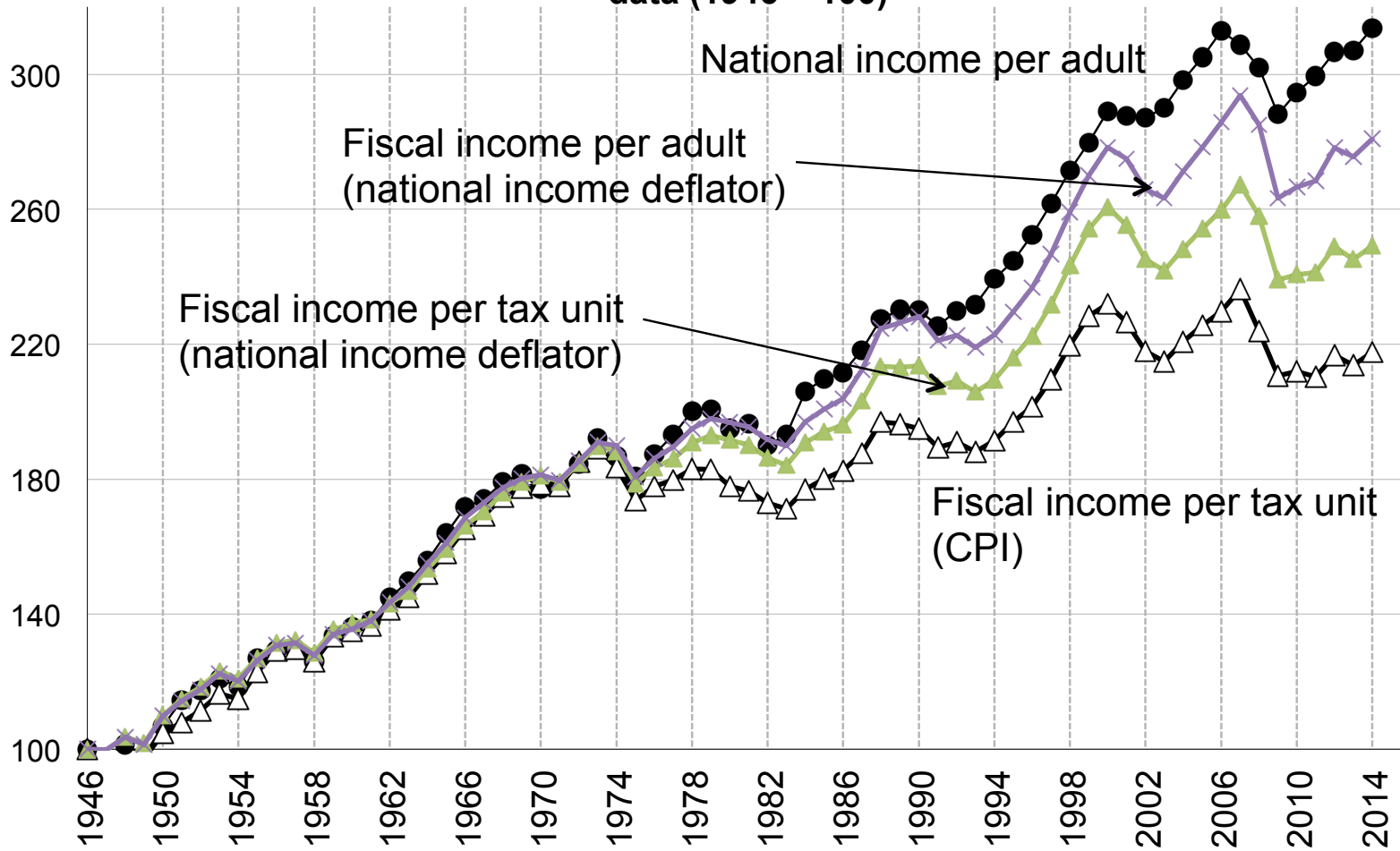
Source: Appendix Table I-A0.

Figure S.20: Ratio of personal pre-tax income to personal factor income
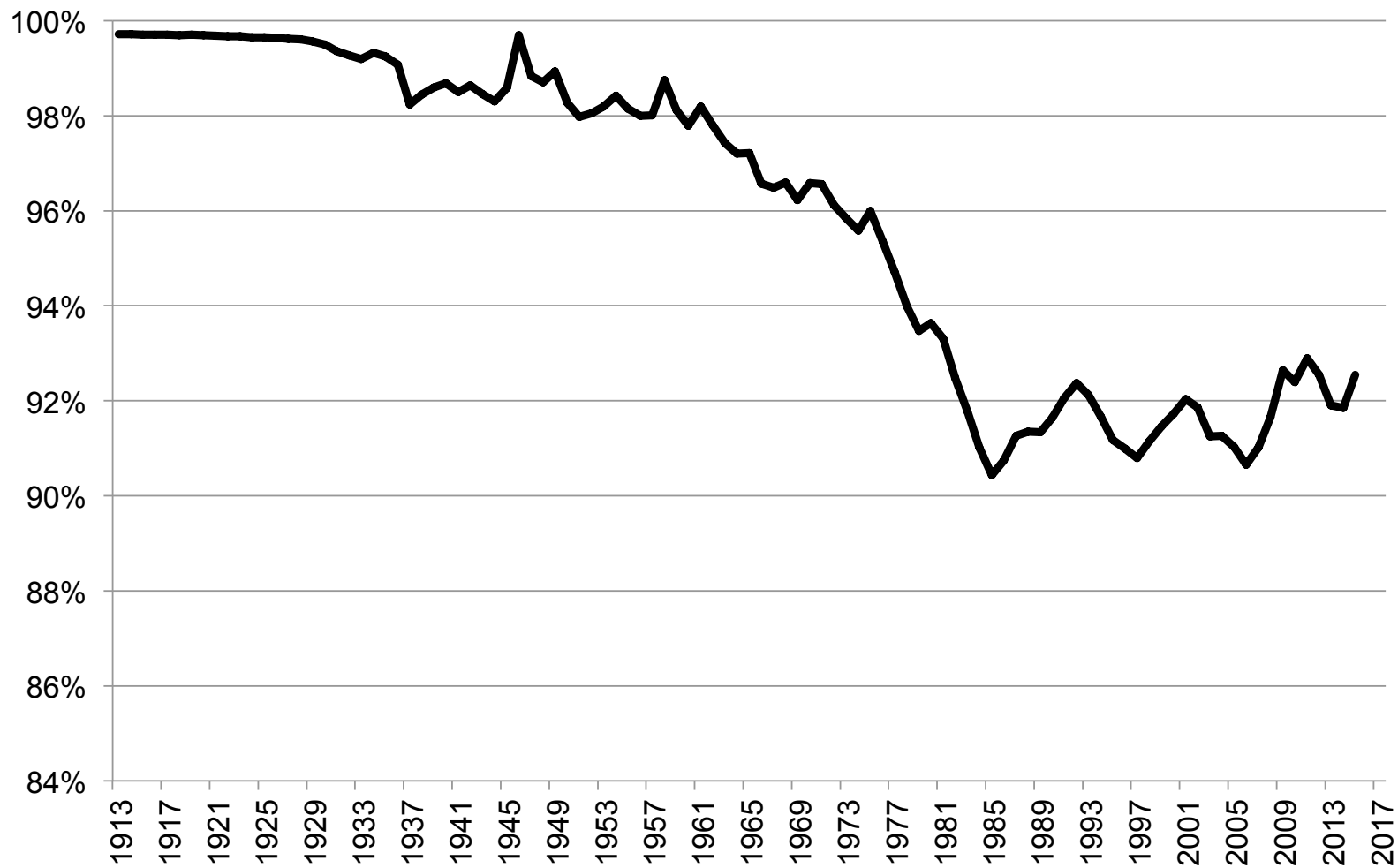
Source: Appendix Table I-A7.

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

Average income in constant 2014 $

Eduction lump sum per child

Education proportional to disposable income (excluding health)

25,000

20,000

15,000

10,000

5,000

0

1962 1966 1970 1974 1978 1982 1986 1990 1994 1998 2002 2006 2010
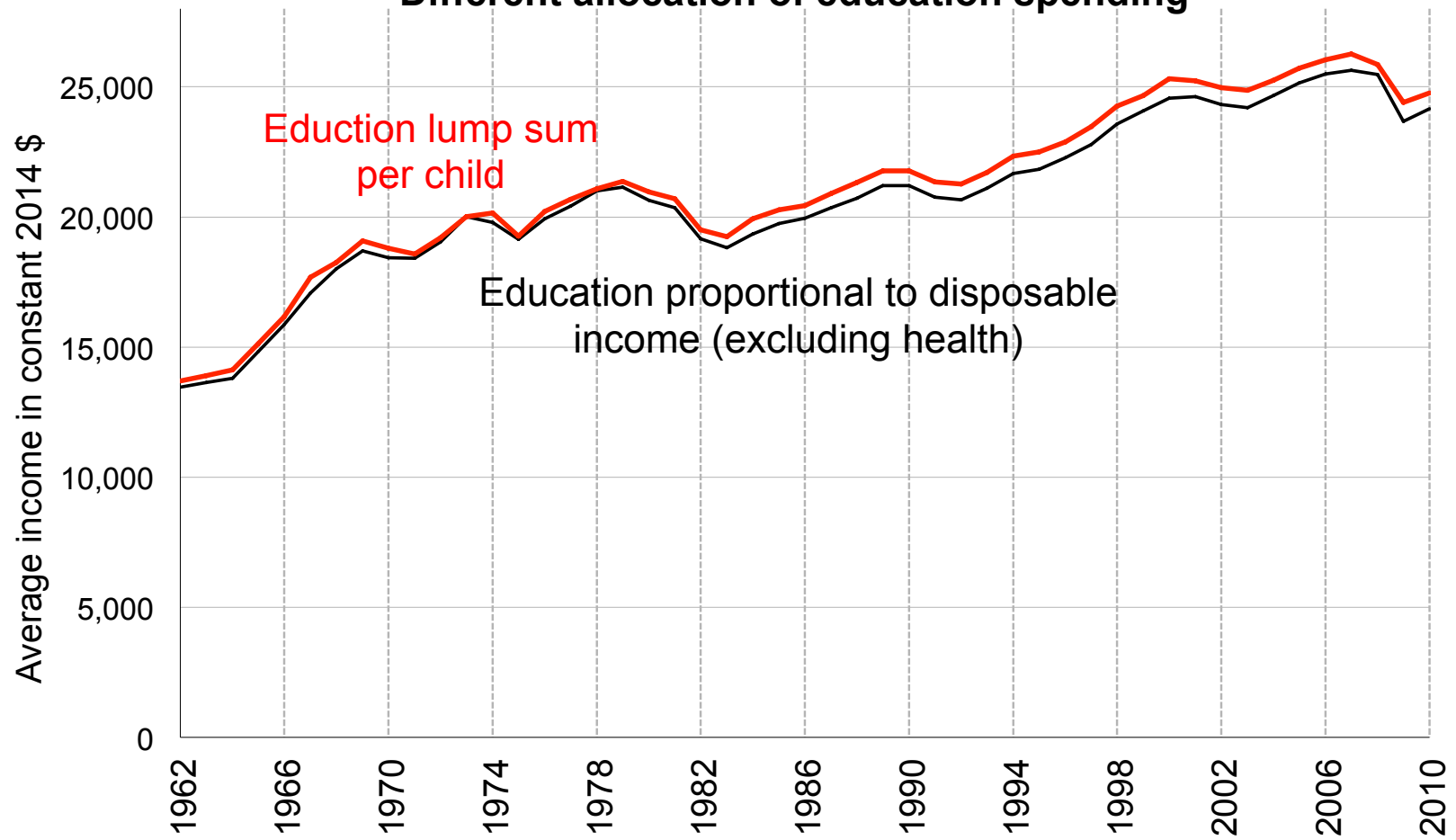
Source: Appendix Tables II-C3d.

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

Individual income taxes

Corporate taxes

Estate taxes

Sales + residential property + payroll taxes

Y-axis: % of top 1% pre-tax income
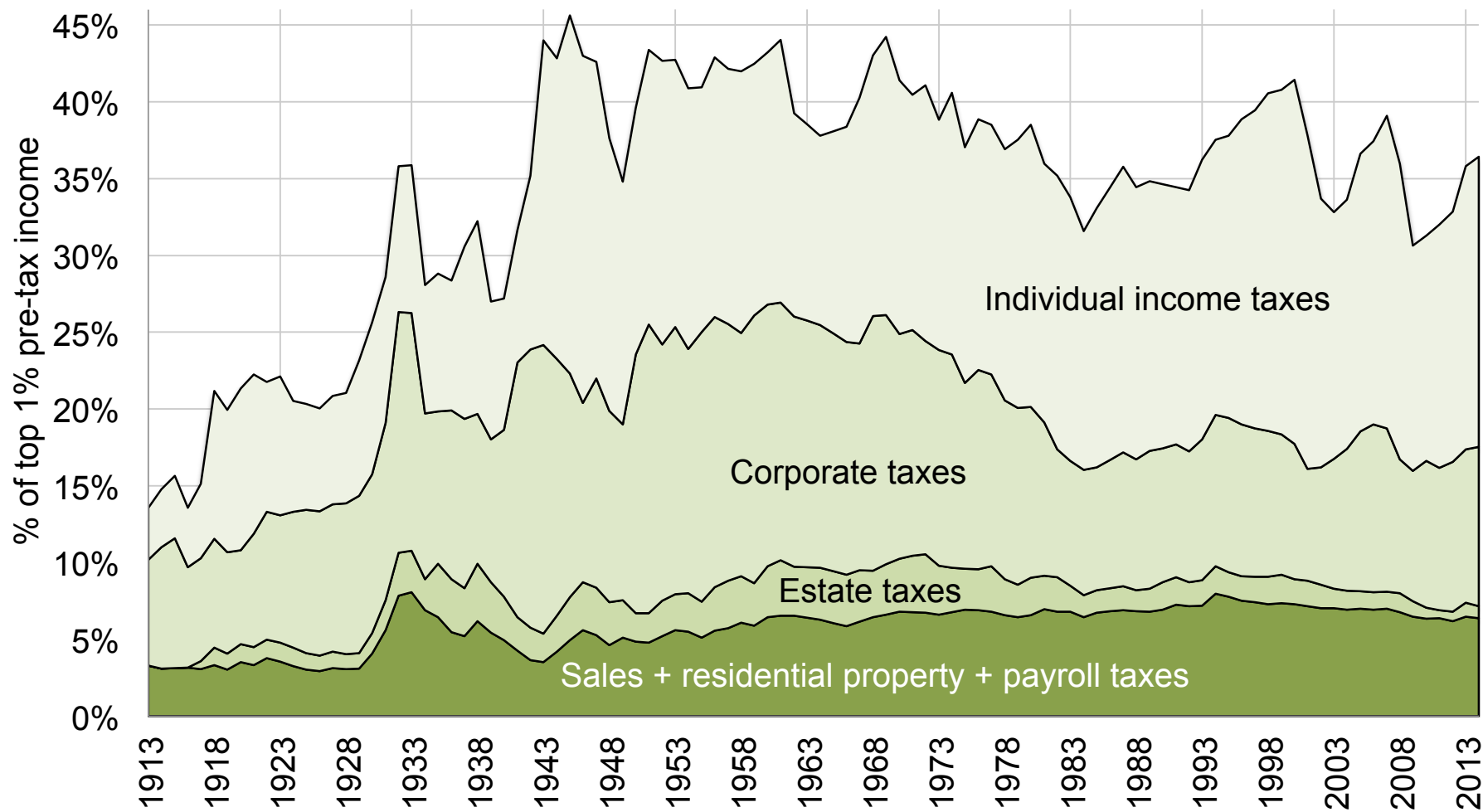
Source: Appendix Table II-G2

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

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

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

	[1]	[2]	[3]	[4]	[5]
	National income per adult	Labor income per adult	Capital income per adult	Share of aggregate per-adult income growth attributed to income from…	
				Labor	Capital
1980-2014	**1.4%**	1.2%	2.2%	63%	37%
1980-1990	**1.7%**	1.6%	1.8%	78%	22%
1990-2000	**2.3%**	2.2%	2.6%	76%	24%
2000-2014	**0.6%**	0.1%	2.2%	10%	90%

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)										
1913-2014	**1.7%**	**1.6%**			**1.8%**	1.8%	1.8%	1.7%	1.7%	2.1%
1913-1946	**1.7%**	**2.0%**			**1.3%**	1.4%	0.9%	0.4%	-0.5%	-0.5%
1913-1929	**1.3%**	**0.8%**			**1.9%**	2.2%	2.0%	1.6%	1.3%	3.3%
1929-1946	**2.2%**	**3.2%**			**0.8%**	0.6%	-0.2%	-0.7%	-2.1%	-3.9%
1946-2014	**1.7%**	**1.4%**			**2.0%**	2.1%	2.2%	2.4%	2.8%	3.4%
1946-1980	**2.0%**	**2.1%**			**1.7%**	1.5%	1.1%	1.1%	1.3%	1.7%
1980-2014	**1.4%**	**0.8%**	0.0%	1.0%	**2.4%**	2.6%	3.3%	3.6%	4.3%	5.2%
2009-2014	**1.7%**	**0.7%**	0.1%	0.9%	**2.9%**	3.2%	3.5%	3.5%	3.5%	2.6%
Fraction of pre-tax income growth accruing to each group										
1913-2014	100%							20%	16%	9%	5%
1913-1946	100%							8%	3%	-2%	-1%
1913-1929	100%							32%	20%	10%	9%
1929-1946	100%	**84%**			**16%**	9%	-2%	-4%	-7%	-4%
1946-2014	100%	**49%**			**52%**	40%	23%	19%	12%	6%
1946-1980	100%	**68%**			**31%**	19%	7%	5%	3%	1%
1980-2014	100%	**34%**	0%	32%	**68%**	56%	36%	30%	19%	9%
2009-2014	100%	**37%**	1%	22%	**77%**	65%	39%	31%	18%	6%

Table S.3b: 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%
				Cumulative growth rates of pre-tax income per adult (equal split among spouses)						
1913-2014	455%	409%			516%		495%	457%	453%	719%
1913-1946	77%	93%			56%		33%	14%	-15%	-16%
1913-1929	22%	13%			35%		37%	28%	23%	67%
1929-1946	**45%**	**70%**			**15%**	11%	-3%	-11%	-2%	-50%
1946-2014	**214%**	**165%**			**296%**	302%	347%	389%	549%	871%
1946-1980	**95%**	**104%**			**79%**	65%	47%	45%	54%	75%
1980-2014	**61%**	**30%**	1%	42%	**121%**	143%	205%	237%	321%	454%
2009-2014	**9%**	**4%**	1%	5%	**15%**	17%	18%	19%	19%	14%
				Fraction of pre-tax income growth accruing to each group						
1913-2014	100%						20%	16%	9%	5%
1913-1946	100%						8%	3%	-2%	-1%
1913-1929	100%						32%	20%	10%	9%
1929-1946	100%	**84%**			**16%**	9%	-2%	-4%	-7%	-4%
1946-2014	100%	**48%**			**52%**	40%	23%	19%	12%	6%
1946-1980	100%	**68%**			**31%**	19%	7%	5%	3%	1%
1980-2014	100%	**32%**	0%	32%	**68%**	56%	36%	30%	19%	9%
2009-2014	100%	**28%**	1%	22%	**77%**	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)									
1913-2014	1.7%						1.6%	1.5%	1.5%	2.0%
1913-1946	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%	**3.2%**			**0.7%**	0.0%	-1.2%	-1.9%	-3.9%	-7.0%
1946-2014	1.7%	**1.6%**			**1.9%**	2.0%	2.3%	2.5%	3.1%	4.1%
1946-1980	2.0%	**2.2%**			**1.5%**	1.5%	1.4%	1.5%	2.1%	3.3%
1980-2014	1.4%	**1.0%**	0.6%	1.2%	**2.2%**	2.5%	3.2%	3.5%	4.2%	5.0%
2009-2014	1.7%	**1.2%**	1.1%	1.3%	**2.5%**	2.6%	2.4%	2.2%	1.7%	0.3%
	Fraction of post-tax income growth accruing to each group									
1913-2014	100%						15%	12%	7%	3%
1913-1946	100%						1%	-3%	-5%	-2%
1913-1929	100%						24%	13%	4%	7%
1929-1946	100%	**88%**			**12%**	0%	-8%	-9%	-8%	-5%
1946-2014	100%	**59%**			**41%**	32%	18%	14%	9%	4%
1946-1980	100%	**75%**			**25%**	16%	6%	5%	3%	1%
1980-2014	100%	**45%**	9%	36%	**55%**	44%	27%	22%	14%	7%
2009-2014	100%	**44%**	13%	31%	**56%**	43%	22%	16%	7%	1%

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

Adult individuals

	1962			2014		
	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%	-16.7%	10.7%	17.1%	-59.1%
P50-P90	48.7%	50.4%	-3.3%	40.2%	41.8%	-4.0%
Top 10%	40.7%	37.3%	8.3%	49.1%	41.1%	16.2%
Top 1%	13.5%	11.2%	17.3%	21.2%	16.5%	22.2%
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%

Adults (equal split among spouses)

	1962			2014		
	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%	-15.3%	12.5%	19.4%	-54.6%
P50-P90	44.4%	45.5%	-2.4%	40.5%	41.6%	-2.8%
Top 10%	36.1%	32.0%	11.3%	47.0%	39.0%	17.0%
Top 1%	12.6%	10.1%	19.9%	20.2%	15.6%	22.5%
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%

Tax units

	1962			2014		
	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	16.2%	18.7%	-15.4%	10.4%	16.0%	-54.1%
P50-P90	45.7%	47.1%	-3.0%	38.9%	41.2%	-5.8%
Top 10%	38.1%	34.2%	10.1%	50.7%	42.8%	15.6%
Top 1%	13.1%	10.7%	18.9%	21.8%	17.0%	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

Adult individuals

	1962			2014		
	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%	-55.7%	8.2%	17.1%	-107.5%
P50-P90	50.8%	50.4%	0.8%	41.6%	41.8%	-0.6%
Top 10%	41.3%	37.3%	9.7%	50.2%	41.1%	18.1%
Top 1%	13.5%	11.2%	17.3%	20.8%	16.5%	20.8%
Top 0.1%	4.6%	3.5%	24.2%	9.7%	7.2%	25.8%
Top 0.01%	1.7%	1.3%	25.6%	4.5%	3.3%	27.6%

Adults (equal split among spouses)

	1962			2014		
	Share factor income	Share post-tax income	Implicit net tax rate	Share factor income	Share post-tax income	Implicit net tax rate
P0-50	17.5%	22.5%	-29.0%	10.3%	19.4%	-87.6%
P50-P90	46.3%	45.5%	1.7%	42.4%	41.6%	2.0%
Top 10%	36.3%	32.0%	11.8%	47.2%	39.0%	17.4%
Top 1%	12.5%	10.1%	19.7%	19.6%	15.6%	20.2%
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%

Tax units

	1962			2014		
	Share factor income	Share post-tax income	Implicit net tax rate	Share factor income	Share post-tax income	Implicit net tax rate
P0-50	14.0%	18.7%	-33.8%	8.9%	16.0%	-80.2%
P50-P90	47.6%	47.1%	1.1%	39.9%	41.2%	-3.2%
Top 10%	38.4%	34.2%	10.9%	51.2%	42.8%	16.4%
Top 1%	13.1%	10.7%	18.8%	21.1%	17.0%	19.4%
Top 0.1%	4.6%	3.5%	24.4%	9.9%	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

Taxes paid as a fraction of pre-tax income, 1962

	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%	5.0%	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%

Taxes paid as a fraction of pre-tax income, 2010

	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%

Taxes paid as a fraction of pre-tax income, 2014

	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%	1.1%	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 Average income	Factor income Income share	Pre-tax income Average income	Pre-tax income Income share	Post-tax income Average income	Post-tax 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%	$25,000	19.4%
Middle 40%	93,760,000	$68,600	42.4%	$65,400	40.5%	$67,200	41.6%
Top 10%	23,440,000	$305,000	47.2%	$304,000	47.0%	$252,000	39.0%
Top 1%	2,344,000	$1,270,000	19.6%	$1,300,000	20.2%	$1,010,000	15.6%
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%	$122,000,000	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.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%	130%
Middle 40%	41%	106%	42%	105%	49%	98%
Top 10%	120%	82%	121%	79%	113%	69%
Top 1%	197%	49%	205%	47%	194%	58%
Top 0.1%	318%	57%	321%	54%	299%	104%
Top 0.01%	455%	79%	454%	75%	424%	201%
Top 0.001%	637%	60%	636%	57%	617%	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.