The Triumph of Injustice:
How the Rich Dodge Taxes and How to Make Them Pay

Emmanuel Saez and Gabriel Zucman

Online Appendix (Last update: October 13, 2019)

This document describes the sources and methods used to compute the data series and figures presented in our book *The Triumph of Injustice*. It also presents a number of Appendix series and figures and describes their construction. This appendix is organized as follows. Section 1 presents the general methodological principles we have followed in this research. Section 2 provides additional methodological details and references for a number of figures included in the book. Section 3 describes our appendix tables and figures (not included in the book). Section 4 provides a list of data files and programs available online allowing to replicate all figures, appendix figures, and appendix tables.

1. Income, wealth, and taxes in the United States: methodology

1.1. General methodological principles

The income and wealth inequality statistics presented in the book (e.g., Figure 1.1. and Figure 5.3) are taken from the Distributional National

Accounts of Piketty, Saez and Zucman (2018), updated to 2018 (see section below 1.2. below for a description of the method used to project to 2018).

Estimates of taxes paid update and supersede those in Piketty, Saez and Zucman (2018). They are based on an enhanced view of tax incidence, described in detail in Saez and Zucman (2019b). Here we describe the main changes compared to Piketty, Saez and Zucman (2018).

Treatment of sales tax

Our treatment of consumption taxes is improved compared to Piketty, Saez and Zucman (2018). PSZ (2018) allocated consumption taxes (sales taxes, excises, and other indirect taxes such as business licenses, motor vehicle licenses, etc.) proportionally to factor income minus saving. One limitation of the original PSZ approach is that it fails to take into account that a large fraction of the consumption basket of the wealthy is not subject to sales taxes (namely, most services).

We now allocate consumption taxes as follows. First, we assume that 70% of consumption taxes are paid by consumers and 30% are paid by factors of production (labor and capital).

For the fraction of consumption taxes paid by consumers, we allocate these taxes following the estimates of ITEP based on a state-by-state

distributional analysis of sales and excise taxes (see https://itep.org/whopays/ October 2018). We apply the estimated ITEP State tax rates by bins of disposable cash income plus individual income taxes (federal plus state) minus food stamps (as groceries are generally exempt) and scale the implied amount of sales taxes paid to match 70% times the total amount of indirect taxes recorded in the NIPA for all levels of government (including federal excise taxes and local sales taxes).

We allocate the fraction of consumption taxes paid by factors of production (30% of the total) proportionally to factor income. By construction, this means that in our series a non-trivial amount of sales taxes is allocated to the wealthiest households (even if their consumption is a very small fraction of their income).

Total consumption taxes at all levels of government add up to about $850 billion in 2018 (4.8% of national income), which can be decomposed as follows:

- $400b for State & local government sales taxes ($300b for states, $100b for local governments);
- $200b for State & local government excise taxes ($160b for states, $40b for local governments);
- $100b for other state & local government indirect taxes (motor vehicle licenses, other license taxes, severance taxes, etc.);

- $100b for federal excise taxes (gasoline, diesel fuel, tobacco, air transport, etc.);
- $50b for federal custom duties

The $850 billion total can also be decomposed roughly as follows:

- $500b of State and local sales and excise taxes captured by the ITEP analysis;
- $100b of federal excise taxes;
- $250b of other taxes (such as motor vehicle licenses, business licenses, other indirect taxes not captured by the ITEP analysis).

Our assumptions are equivalent to following ITEP for the $500bn in state and local sales and excise taxes captured by the ITEP analysis, applying the ITEP distributions to the $100bn in federal excise taxes, and allocating the $250b residual (30% of $850bn) proportionally to factor income.

Treatment of corporate taxes

We allocate the corporate tax to corporate equity owners (excluding S-corporations), which in our view is the only meaningful way to allocate existing corporate taxes. Of course, should the corporate tax rate change in the future, pre-tax incomes might change (e.g., wages might increase)—but that's different question from the question of who pays existing taxes. See Saez and Zucman (2019b). Note that in the 1950s and 1960s, the corporate

tax generated a large amount of tax revenue (almost as much as the individual income tax in the early 1950s), and equity ownership was concentrated. This explains why effective tax rates at the very top are higher in the book than in PSZ (2018) in the 1950s-1960s. PSZ in effect allocated part of the corporate tax to bond-holders and owners of non-corporate businesses further down the wealth ladder, while our revised treatment allocates all the corporate tax to the persons who effectively paid these taxes (namely, shareholders).

Population restriction

Our estimates of effective tax rates (as a % of pre-tax income) across the distribution (e.g., Figure 1.2) restrict the population to adults earning more than half the federal minimum wage in pre-tax income. People with very low pre-tax income (below half the federal minimum wage) earn transfer income (temporary assistance, SNAP, supplemental security income, veteran benefits, etc.), which is not part of pre-tax income. They pay sales taxes on that transfer income when it is consumed. As a result, they have high (sometimes very high) tax rates as a fraction of their pre-tax income. We avoid that problem by restricting the population to adults with more than half the minimum in pre-tax income. This is equivalent to winsorizing at roughly the 9^{th} percentile of the pre-tax income distribution (with little variation over time).

Top 400

The smallest group we consider is the top 400, a group that owns about 3% of total wealth and hence could account for a non-negligible fraction of tax revenue. Studying the top 400 is important because several tax proposals would increase taxes specifically on the top 400. Recognizing the economic importance of this group, the IRS also publishes statistics on the top 400 (their statistics, however, cover the top 400 tax units with the highest fiscal income, while our statistics cover the top 400 adults with the highest pre-tax income, whether this income is subject to the individual income tax or not). Forbes magazine attempts to measure the wealth of the top 400 wealthiest Americans and has made this group salient and concrete in popular discussions.

In our statistics, the top 400 in pre-tax income and the top 400 in wealth are essentially the same (since pre-tax income includes the entire return on wealth, rather than the realized return on wealth). Thus, the top 400 in our data can be understood as the wealthiest 400 Americans or the 400 Americans with the highest economic income. Our estimates of the wealth of the top 400 is not based on Forbes magazine (it is based on capitalizing income, see Saez and Zucman, 2016), but is consistent with Forbes magazine data in recent years (see Saez and Zucman, 2016). In 2018, for example, our estimate of the wealth of the top 400 is exactly the same as

the Forbes estimates, $7.2 billion on average. Saez and Zucman (2019c) discuss the limitations of the Forbes estimates.

Although PSZ did not provide estimates for the top 400, the micro-files created by Piketty, Saez and Zucman (2018) can be used to analyze the top 400. By construction however, in PSZ the top 400 highest pre-tax income earners are tax units with large amounts of fiscal income. They include individuals with particularly large realized capital gains who may not have commensurately large economic annual incomes (for example, a business owner who sells his business which represents income earned over decades rather than just one year of income). PSZ (2018) did not attempt to address the re-ranking issue involved in moving from the distribution of fiscal income to the distribution of pre-tax income.

In Saez and Zucman (2019c), we attempt to address that issue. We estimate the fiscal income of the top 400 by triangulating across publicly available sources. Using three different sources and methodologies, we estimate that the top 400 have a fiscal income that is about 45% of their true economic pre-tax income (defined as wealth times the average macroeconomic return to wealth) vs. 70% economy-wide; see Saez and Zucman (2019c) Section 3.1 and Table 3 for complete details. This implies that the effective tax rate of the top 400 (out of pre-tax income) is lower than the effective tax rate of

the group just below and the effective tax rate of the top 400 in the original PSZ (2018) data.[1]

Therefore, we reduce the fiscal income of the top 400 such that the top 400 tax units collectively have a fiscal income /pre-tax income ratio of 45% at most each year on average (see program top400.do). Note that in some years, especially in the 1960s, 1970s, and 1980s this ratio was already below 45% in which case we do not adjust the data. After having made this adjustment for fiscal income, we adjust the amount of income taxes paid proportionally (see program top400.do).

Our 45% estimate of fiscal income relative to full economic income for the top 400 is based on triangulating publicly available sources and it could be refined in future work. This triangulation is the best approximation we could create using public sources. Given the importance of the question—how much do billionaires really pay in taxes?—it would be desirable to mobilize internal data to provide better estimates in the future. For example, linking the Forbes 400 to income tax data would allow for a direct

[1] We have not found evidence that the ratio of fiscal income to pre-tax income falls for groups just below the top 400, and hence did not correct data below the top 400. The ratio of fiscal income to pre-tax income seems to be falling only at the very top, where a large fraction of wealth is in equities issued by C-corporations that do not distribute all their profits (e.g., Amazon, Berkshire Hathaway, Facebook, etc.)

estimation of the fiscal income of the 400 richest.[2] Linking the Forbes 400 to corporate tax data would allow to directly estimate how much they pay in corporate taxes. Of course, a well-enforced wealth tax would be an even better source to study this question in depth (and cross-check the Forbes 400 estimates). We hope that our estimates, imperfect as they may be, will generate more and better research, and we will adjust our numbers when better estimates become available.

Note that we do not attempt to reattribute the income earned by foundations controlled by living founders to their founders. In particular, the pre-tax income of the top 400 fully excludes foundation capital income (which is non-negligible in some cases, e.g., Bill Gates). This tends to bias our estimated top tax rates upwards. Reattributing the income of foundations to the people who control them is left to future research.

Denominator used to compute tax rates: treatment of capital gains

In Piketty, Saez and Zucman (2018), effective tax rates were computed by dividing taxes paid by pre-tax national income. This raises the problem that when people realize a lot of capital gains such as in 2000 or 2007, peak years for the stock-market, effective tax rates mechanically rise (because realized capital gains trigger taxation but capital gains are excluded from

[2] Similar linking for research purposes has already been done in the context of estate tax data by Raub, Johnson, and Newcomb (2010).

pre-tax national income), making it hard to interpret year-to-year changes in effective tax rates.

We improve the computation of effective tax rates compared to PSZ (2018) by adding pure capital gains to our income denominator when computing tax rates. The challenge is to include only the "pure" component of capital gains (due to pure asset price increases), as capital gains generated by the accumulation of retained earnings are already included in pre-tax national income via the inclusion of corporate retained earnings in pre-tax income.

What fraction of realized capital gains should be included in the denominator of effective tax rates? One possibility is to remove from realized capital gains the average flow of corporate retained earnings. On average over 1962-2018, corporate retained earnings (domestic + foreign) have amounted to 3.8% of national income; realized capital gains to 3.6% of national income; realized taxable capital gains (i.e., included in AGI) to 2.8% of national income (50% or more of realized capital gains were exempt before 1987). The fact that retained earnings are slightly higher than realized capital gains on average can be due to several factors, such as the fact that some retained earnings are in corporations owned by tax-exempt shareholders (pensions), and permanent deferral (e.g., due to basis step-up at death).

In our series, we consider that realized taxable capital gains amounting to 3.0% of national income reflect retained earnings. 3% of national income corresponds to 80% of the average flow of retained earnings, thus taking into account that part of retained earnings accrue to non-taxable investors (pensions, foreigners, etc.). We consider that any realized taxable capital gains above 3.0% of national income reflects pure price effects. We add these pure-price capital gains to our income denominator. In 1986 we add 50% of realized taxable capital gains to the denominator to account for the shifting of capital gains in 1986 (before the abolition of the 60% capital gains exemption in 1987).

In practice, this correction makes a difference only for the years 1925, 1927, 1928, 1929, 1986, 1987, 1988, 1996, 1997, 1998, 1999, 2000, 2001, 2004, 2005, 2006, 2007, 2008, 2012, 2013, 2014, 2015, 2016, 2017, and 2018 (as in other years realized taxable capital gains are less than 3.0% of national income); see Appendix Figure 1.11. In 2018 for example, the amount of realized taxable capital gains is around 5.0% of national income (about $900 billion out of a total national income of $17,600 billion). This is 2 points in excess of 3.0%, so we include 40% of realized taxable capital gains (2.0%/5.0%) in our pre-tax income denominator for the purpose of computing effective tax rates.

Treatment of taxes before 1962

We slightly change the treatment of taxes before 1962 compared to Piketty, Saez and Zucman (2018). First, we adjust all our pre-1962 computations to be consistent with our new post-1962 assumptions on the incidence of sales and excise taxes (see above). Moreover, we now ensure continuity in 1962 for the individual income tax. In PSZ (2018) we applied the raw distribution of federal income taxes to the pre-1962 years. But in 1962, the distribution of income taxes across adults ranked by pre-tax income in DINA is slightly less concentrated than the distribution of federal income taxes across tax units ranked by fiscal income (due to re-ranking when moving from fiscal to pre-tax income). We now explicitly deal with this discontinuity, which slightly reduces the effective tax rates at the top before 1962.

Top income and taxes during the Great Depression

Our treatment of top incomes and tax rates during the Great Depression is improved compared to Piketty, Saez and Zucman (2018). The main difficulty in this period is that in 1931-1932, the macro flow of equity income is zero or negative, with positive dividend payments offset by negative retained earnings. In our benchmark estimates, all groups of the distribution get the same rate of return on equity wealth (equal to the macro flow of equity income dividend by equity wealth). This implies that people who do receive sizable dividends (and hence are ranked at the top of the distribution) get zero equity income, which is inconsistent. In reality, people with large

dividend payments owned still profitable businesses (and so should get a positive return on equity wealth), while people who owned loss-making businesses should get a negative return (and be ranked at the bottom of the distribution). To address this re-ranking issue, we apply heterogeneous returns to equity in the early years of the Great Depression: we assume that at the top (top 1% and above) the return to equity wealth was 5% in 1931 (close to the macro return observed in 1930; see PSZ Table I-B7), and 3% in 1932 and 1933 (close to the macro return observed in 1934). This implies that equity returns down the income distribution are negative (such that the macro equity return is about 0). This increases the top 1% and above shares in 1930 and especially 1931 and 1932 (which fell too much in PSZ and are now much more consistent with the Piketty-Saez fiscal income series).

1.2. How we project income and taxes to 2018

The procedure we use to create income and tax statistics for the most recent years is the following. First, we age the fiscal income files using the most recent available data. Second, we apply the distributional national account methodology to the aged fiscal income files using current aggregate national account data and projections for future years.

Aging fiscal income files

The distributional national accounts methodology of Piketty, Saez and Zucman (2018) starts from annual individual income tax data files (fiscal income files). To create fiscal income files, we proceed as follows.

For years 2013, ..., 2016, we use the aging technique described in Saez and Zucman (2018) based on the 2012 public use file (the latest year available) and tabulations based on the internal files for years 2013, ..., 2016. The tabulations provide detailed statistics by income groups and demographic groups for each fiscal income component. We update the 2012 public use file to match the tabulated statistics from internal data. Saez and Zucman (2018) show that this aging methodology produces distributional statistics very close to the statistics obtained using directly the internal data.

For 2017, we start from the 2012 public use file. We inflate weights uniformly to match the total adult population (aged 20 and above). We then inflate each fiscal income component to match the aggregate levels for 2017 for each fiscal income component published by the IRS Statistics of Income. This strategy works well because the 2017 distribution of fiscal income is very similar to the 2012 distribution of fiscal as shown by the Piketty and Saez (2003) series for 2017.

For 2018 and after, we start from the 2017 aged fiscal income file and assume that adult population growth is .9% per year. We then inflate each

income component assuming that each aggregate fiscal income component grows at the same rate as nominal GDP: 5.3% from 2017 to 2018, 4.5% from 2018 to 2019, and 4.1% from 2019 to 2020. These projections for nominal GDP growth are from Federal Reserve open market committee, 26 Sept. 2018. This again amounts to assuming that the distribution of fiscal income in 2017 and after remains stable. Preliminary fiscal income statistics for year 2018 (updated Piketty-Saez series) show indeed that the fiscal income distribution and its composition are very stable from 2017 to 2018. This stability assumption is the most natural benchmark to use for the 2019 and 2020 projections (years for which there is no data yet available). Naturally, all these assumptions will be revisited as more data become available in order to produce even more accurate statistics. Because of the large 2018 tax reform, we recomputed individual income taxes for 2018 and after using the NBER TAXSIM simulator (Feenberg and Coutts, 1993).

Distributional National Accounts Recent Years and Projections

Once we have built aged fiscal income files, we apply the distributional national account methodology to the files. For this, we use the current national accounts for 2016-2017, and project national accounts for 2018-2020 by inflating each aggregate income component in national accounts using the GDP nominal growth assumptions from the Federal Reserve open market committee, 26 Sept. 2018 (5.3% nominal aggregate growth from

2017 to 2018, 4.5% from 2018 to 2019, and 4.1% from 2019 to 2020). We assume that each aggregate income component grows at the same rate as nominal GDP. We inflate each aggregate wealth component of the US financial accounts using the projections also provided by the Federal Reserve open market committee, 26 Sept. 2018 (6.3% nominal aggregate growth from 2017 to 2018, 5.5% from 2018 to 2019, and 5.1% from 2019 to 2020).

2. Notes to the figures presented in the book

The figures included in the book are all available in Excel format, along with the underlying data series, in the file SZ2019MainFigures.xlsx. In turn, the underlying data series are borrowed from the file SZ2019AppendixTables.xlsx (see notes in the relevant data sheet). For example, the data used to construct Figure 1.1 are in the sheet DataF1 of SZ2019MainFigures.xlsx; in turn, these data come from Appendix Table A1 in SZ2019AppendixTables.xlsx. The book includes notes that describe the figures. Section 3 below describes each appendix table in SZ2019AppendixTables.xlsx precisely. Here we provide a number of additional notes and methodological details for a number of figures.

Figure 4.1: effect of mandatory repatriation enacted in 2017

The revenues generated from the repatriation tax are recorded by BEA in the NIPA as a one-time capital transfer in the fourth quarter of 2017 of $250.0 billion. This means they are excluded from 2018-onwards current corporate income tax payments and revenues recorded in the NIPAs. Note that the 2017 capital transfer reduces the government deficit recorded in the NIPA in 2017 by about 1.5% of national income (and symmetrically reduces corporate saving after capital transfer by about 1.5%, explaining the relatively low amount of corporate saving in DINA Table I-D.1). Note also that $250 billion is the same amount as estimated by Wrigh and Zucman (2018; Appendix Table A.19) based on JCT estimates (purged of the estimated effects of the repatriation on individual income tax revenues). It is consistent with a low effective tax rate of 8% (post-credits) on $3,125bn in previously untaxed profits. According to SOI tabulations of corporate tax forms 5471, in 2014 (the latest year available) controlled foreign corporations of U.S. firms had $2,727 billion in previously untaxed earnings & profits (of which more than $2 trillion in tax haven subsidiaries). It's likely that by end of 2017 the total previously untaxed E&P was around (possibly in excess) of $3,125 billion. In any case, the point is that the effective rate on the repatriated profits is very low (especially given that 60% of the tax payments are back loaded to 2023-2024-2015 with no interest rate charge).

Figure 5.1: computation of labor vs. capital tax rates

The key difficulty in computing labor vs. capital tax rates involves splitting income tax revenue into labor income tax vs. capital income tax. We split income tax revenue (federal + state) into a capital and labor tax component proportionally to the share of capital and labor in gross taxable income. Capital income in gross taxable income includes net realized capital gains + dividends + interest + rents & royalties + estate & trust + 30% of (sole proprietorship + partnerships + S corporation income) + the fraction of taxable pension income that corresponds to the capital share of pension income (defined as the ratio of investment income paid to pension funds to (investment income + pension contributions), about 40%-45% in recent years). To compute macro tax rates, we include pure capital gains (defined as taxable realized capital gains in excess of 3% of national income) at the denominator. We assign all pure capital gains to capital income.

Note that we do not attempt to take into account that dividends are taxed at a lower rate than wages since 2001, which means that we slightly under-estimate the true decline in the average capital tax rate since the turn of the 21st century. Note also that capital gains have historically been taxed at lower rates than ordinary income, which leads us to over-estimate the true capital tax rate in level, but conversely capital income is typically earned by individuals in high tax brackets, which leads us to under-estimate the true capital tax rate in level. We assume that the two effects offset each other, which is probably reasonable as a first order approximation.

Figure 5.2: employer-sponsored health insurance

We only include employer-provided health insurance premiums ($1,044 billion in 2017, or 6.2% of national income) in our computation of the extended labor tax rate.

Figures 9.1 and 9.3: private health insurance

We add other private health insurance premiums (direct purchases by households, medical portion of property and casualty insurance, and ACA subsidies: $140 billion in 2017 or 0.8% of national income) to employer-provided health insurance premiums, for a total amount of private health insurance of $1,184 billion in 2017 or 7.1% of national income.

3. Appendix tables and figures

A number of appendix tables and figures (not printed in the book) are made available in the Excel files SZ2019AppendixTables.xlsx and SZ2019AppendixFigres.xlsx. The Appendix Figures are constructed using the data in SZ2019AppendixTables.xlsx. In this Section, we describe file SZ2019AppendixTables.xlsx.

SZ2019AppendixTables.xlsx is organized in four sections: A, B, C, D.

Appendix Tables A present statistics covering the entire population. Unless otherwise noted, income is pre-tax income as defined in Piketty, Saez and Zucman (2018).

Appendix Tables B restrict the population to adults earnings more than ½ the minimum wage in pre-tax income, and unless otherwise noted, pre-tax income is upgraded to include pure taxable capital gains (defined as taxable capital gains in excess of 3% of national income); see section 1.1 above.

Appendix Tables C provide detailed statistics on income, wealth, and taxes paid by generalized percentiles of the distribution of pre-tax tax income ((28 g-percentiles: p1, p2,… p99, p99.1,… p99.9, p99.91, … p99.99, p99.991, …, p99.999, with the top 400 as an extra memo line) for selected years.

Appendix Tables D provide background computations for our benchmark tax reform simulations.

We now provide additional details on each appendix table.

Appendix Table A1: Income, wealth, and their distribution in the US

- Cols. 1, 10, 19, 20: Aggregates, copied from PSZ (2018) Appendix Table I–A0.

- Cols. 2–8: Income shares, full population, matching national income (i.e., excluding capital gains), equal-split adults. Series are copied from PSZ (2018) Appendix Table II-B1 (external files).
- Cols. 11-17: wealth shares, full population, matching household wealth, equal-split adults. Series are copied from PSZ (2018) Appendix Table II-E1 (external files).
- Cols. 9 and 18: For the top 400, external files in recent years do not give accurate results due to insufficient sample sizes; the results reported in cols. 9 and 18 are thus obtained from internal files (taxgperc outputs produced by outsheet_dina.do run on Nov. 16, 2018; collated in SZ2019AppendixTables.do and exported in SZ2019Appendix.xlsx top400* raw output sheets). Note that the top 400 results reported here are for equal-split adults (like for other columns in this table), and income is pre-tax income excluding pure capital gains. (Results including pure capital gains can be found in taxgpercpeinc_kg outputs; results for tax units can be found in top400.xlsx). Note also that the top 400 by wealth has 2.8% of household wealth on an equal-split basis in 2016, 3.2% on a tax units basis, and coincidentally 3.2% in *Forbes* too. That is, in 2016 (latest year of the internal files), Forbes and PSZ (2018) have exactly the same amount of wealth for the top 400 tax units.

Appendix Table A2: Income and wealth levels in the US

Same as Appendix Table A1, with income and wealth levels (in 2018$, using the national income deflator) instead of shares.

Appendix Table A3: Aggregate taxes paid in the US

Copied from PSZ 2018 Appendix Table I-SA12b. On the right panel we report tax payments as a fraction of (national income + pure capital gains), the denominator we use for our tax rates calculations. Because pure capital gains are 0 most years and when positive add up to only a few percentage points of national income, the effect on the tax/national income ratio is negligible.

Appendix Table A4: Aggregate taxes paid in the US, including private health insurance premiums.

The source is the National Health Expenditure Accounts Table 5–6 (see raw sheet NHE and formulas in Table A4). For Figure 5.2 in the book we only include employer-provided health insurance premiums ($1,044 billion in 2017, or 6.2% of national income) in our computation of the extended labor tax rate. For Figures 9.1 and 9.3 we add other private health insurance premiums (direct purchases by households, medical portion of property and casualty insurance, and ACA subsidies: $140 billion in 2017 or 0.8% of national income, for a total amount of private health insurance of $1,184 billion in 2017 or 7.1% of national income).

Note that the NIPA "Private group health insurance paid by employer" (part of compensation of employees, $726 billion in 2017) does not include government contributions to privately administered health insurance for government employees (which are classified as employer contributions for employee pension and insurance funds). It does not include the employee portion of employer-sponsored private health insurance either ($278 billion in 2017).

Appendix Table A5: Effective tax rates on labor and capital.

This is constructed from PSZ (2018) aggregates, see formulas. On the construction of capital and labor rates, see notes to Figure 5.1 in section 2 above.

Appendix Table A6: Government revenue, spending, and deficits (all levels of government).

This is constructed from PSZ (2018) aggregates, see formulas. For 2019 we project a total government deficit of 7% of national income. This can be reconciled with official CBO federal budget projection of $976 billion (4.5% of GDP) for fiscal year 2019 (= ending October 1, 2019) as follows: 4.5% of GDP = 5.3% of national income, to which must be added the state and local government deficit which is project to be around $300 (= level observed in recent years) = 1.6% of national income. State and local governments have deficits because (i) despite budget balance requirements current

expenditures exceed current revenue (by about $200 billion in 2017); (ii) net investment of $100 billion (investment is outside of balanced budget rules). NB: a state and local government deficit s = -1.6% of national income means steady-state debt state and local government debt of s/g = 30% of national income (if $g = 5\%$ = 3% real growth + 2% inflation, which erodes debts), which is close to the current debt level of State and local governments. A total government deficit of $s = -7\%$ of national income implies a steady-state debt/income ratio of 140% if nominal growth is 5%, which is close to the current level of 130%. That is, the big Trump deficits are not going to increase much the debt/GDP ratio in the short-run given the relatively strong nominal growth of the economy. NB2: Maastricht targets of 60% of debt/GDP and 3% deficit/GDP can be understood as follows: 60% is the steady state debt/GDP ratio with 3% deficits and 5% nominal growth. Problem: this is probably inconsistent with 2% inflation mandate for ECB. NB3: in all these computations, debt should be understood as financial liabilities minus financial assets (i.e., net financial debt).

Appendix Table B1: Taxes paid by pre-tax income groups.

In appendix B, the population is restricted to adults with more than ½ the minimum wage in pre-tax income, and tax rates are expressed as a fraction of pre-tax income plus pure capital gains.

- Post-1962 series are obtained from external DINA files, taxgperc outputs collated and collapsed by deciles in SZ2019AppendixTables.do, and exported in SZ2019Appendix.xlsx taxrates_wide and taxrates_long raw output sheets.
- Pre-1962 series are pasted from PSZ (2018) Appendix Table II-G1, (series reported in PSZ Appendix(Distrib) are for the entire population, but the match is almost perfect in 1962 so we do not make any adjustment).

Appendix Tables B2 and B3: Composition of taxes paid by groups of the pre-tax income distribution.

Pre-1962 series copied from PSZ (2018) Appendix Table II-G2 (except for the bottom 50% which is re-computed, see formulas).

Appendix Table B4: Tax rates by deciles (and smaller groups at the top).

Post-1962 series are straight out of external DINA micro-files (raw output taxrates_wide).

Pre-1962 series are obtained by starting from 1962 tax rates by income bin x type of tax, and assuming that each of these rates follows the same evolution as the macro tax rate (i.e., the sales tax rate for P0-10 follows the same evolution as the macro effective sales tax rate; the payroll tax rate for P50-60 follows the same evolution as the macro payroll tax rate, etc.). Over 1950-1962 tax rates and inequality are stable (with effective top marginal

individual income tax rates of 87% throughout the period and stable levels of income and wealth inequality; see Appendix Table A1). As a first approximation methodology provides accurate results. The only limitation is that it slightly under-estimates the true effective tax rate for the top 0.01% (as compared to the more sophisticated estimates for the very top groups reported in PSZ 2018 and Appendix Table B1) in the early 1950s, when effective corporate tax rates were higher than in 1962. We address this by slightly adjusting the effective corporate tax rates at the top in col. 17, 18, 19. By construction, the adjusted rates for the very top groups are consistent with the effective tax rates reported in Appendix Table B1.

Appendix Table B5: Average tax rates by pre-tax income groups (selected years)

Same as Appendix Table B4, but for selected years (1950, 1960, …, 2010, 2018).

Appendix Tables C1, … C8: Average tax rates by g-percentile of pre-tax income

These tables provide average income, wealth, taxes, and tax rates by type of tax by detailed bins of pre-tax income (128 g-percentiles: p1, p2,… p99, p99.1,… p99.9, p99.91, … p99.99, p99.991, …, p99.999, with the top 400 as an extra memo line) for selected years (1950, 1962, 1970, 1980, 1990, 2000, 2010, 2018).

These tables are constructed from external DINA micro-files (raw output full*). In recent years, the top 400 is not accurate in external files due to insufficient sample sizes; therefore, we report the top 400 line from internal outputs instead. Note that there is tiny inconsistency in the computation of the averages for P99.99-top400: we take P99.99-100 from external outputs and subtract the top 400 from internal outputs, hence the averages reported for P99.99-top400 in Appendix C Tables differ very slightly from those reported in Appendix B. This discrepancy is irrelevant for our purposes, and the Appendix B series are the correct ones.

In Table C8 (for year 2018), we report the amount of private health insurance by g-percentile and the implied tax rate. The amount of employer-sponsored health insurance premiums by bin of income is estimated for year 2015 using internal IRS W2 tax data (file fringepeincgperc2015equalsplit.xlsx). Note that W2s only cover employers with more than 250 employees; we therefore upgrade the average premiums in the bottom 60% of the distribution (see col. 15). We then scale employer-sponsored insurance premiums to match the total amount of private health insurance in 2018, as recorded in Table A4, namely $1,246 billion (see discussion of Table A4 above for details and sources).

Appendix Table D1: this table presents effective tax rate in our baseline tax reform. See the online description of the tax simulator TaxJusticeNow.org for complete details.

Appendix Table D2: this table decomposes the profits of US corporations in 2018 and computes the base of the corporate income tax (i.e., excluding S-corporations). See notes underlying the table. In 2018, US domestic corporate profits (net of capital depreciation and interest payments, before payments of corporate taxes) amounted to $1,715 billion, close to 10% of national income. Out of that total, the domestic tax base (C-corporations) was $1,308 billion (vs. S-corporations: $408 billion).

Appendix Table D3: The table reports the net income (i.e., profits), taxes paid, and effective tax rate (the ratio of taxes paid to net income) of US multinational corporations in various countries in 2016 and 2018. The remedial tax rate is the extra tax rate levied by the United States that would ensure that the profits of US multinationals are taxed at least at 25% in each of the countries where they operate. We use data from IRS form 8975 (country-by-country reports) for year 2016 (the latest year available) to simulate the static revenue from such a 25% country-by-country remedial tax. In 2016 this would have yielded $61 billion in tax revenue (of which $55 billion coming from profits booked in tax havens, and $6 billion coming from profits booked in countries generally not considered tax havens but with corporate tax rates below 25%, e.g., the UK). Assuming a 4% nominal annual growth between 2016 and 2019 gives $67 billion in static revenue in 2018. This figure under-estimates the true revenue because this

computation is based on aggregate data (not firm-level data) and there is heterogeneity in the effective tax rates paid by US firms within each foreign country.

4. Replication archive

SZ2019.zip is a replication archive that allows to reproduce the figures, appendix tables, and appendix figures of the book. It is structured as follows.

- **Directory "MicroFiles":** contains the raw Distributional National Accounts microfiles used to create the results of the book, Appendix, and Simulator. We ran our programs on a version of these micro-files ("External") created using the public-use samples of tax returns disseminated through the NBER. These "External" micro-files cannot be made available publicly; therefore we created a smaller version of these micro-files ("Online") that we make available for replication purposes. Because the "online" files are coarser than the "external files", our results cannot be perfectly replicated with the "online files".

- **Directory "Programs":** contains the Stata programs that produce the results in the book and its Appendix. To run these programs on your

computer, edit the first two lines (root directory and online macro which should be set to 1).

- **Directory "TablesFigures":** contains the results from the Book and its Appendix in a set of Excel files. These Excel files are created using (i) the Stata outputs generated by the Stata programs in "Programs", (ii) updated versions of the Excel Appendices I (macro series) and II (distributional series) of Piketty, Saez and Zucman (2018) made available in TablesFigures/PSZUpdates.