

How Pronounced is the U-Curve? Revisiting Income Inequality in the United States, 1917-1945

Vincent Geloso, Phillip Magness, John Moore and Philip Schlosser *

August 1, 2018

Abstract

We offer a revision to the top income share series of Piketty and Saez (2003) for the United States, focusing upon the period from 1917 to 1945. Our revision corrects for distortions in the original series that arise from data construction problems with tax records, as well as incorporating an improved fiscal income denominator derived from national accounts. Our corrections substantially alter both the shape and magnitude of the left side of the inequality U-curve for the United States, including a 6.7 percentage point reduction on average to Piketty-Saez's top 10% income share estimates prior to 1944.

Keywords: Inequality, Top Incomes, US Economic History

1 Introduction

Interest in the study of economic inequality has undergone a marked revival in recent years, most notably focusing upon long term measurement of top income shares. Many of the empirical advances in recent decades derive from the estimation methods developed by Thomas Piketty (2001), who initiated his work with top incomes in France and then expanded it for the United States in collaboration with Emmanuel Saez (2003) (henceforth PS). This line of

*The authors are respectively affiliated with Texas Tech University, Berry College, Northwood University and University of Detroit Mercy. Acknowledgments go to Art Carden, Dale Matcheck, Stephen Miller, Richard Sutch, Peter Lindert, Jeffrey Williamson, Edward Stringham, Ben Powell, Jeremy Horpedahl, David Splinter, Russell Sobel, Antony Davies, and Joshua Hall for their helpful comments on early versions of this paper. We also extend our thanks to Rachel Kutsinger for her data research assistance in the preparation of this project.

research has since produced data sets spanning most of the 20th century for more than two dozen countries, measuring both fiscal income and the closely related distribution of wealth (Atkinson et al., 2011; Solt, 2016). Generally, these data derive from historical tax records. The use of a single historical series with a consistent data source constitutes a significant improvement over earlier estimation techniques developed by Simon Kuznets (1953) as well as attempts to assemble distributional data from numerous types of indicators (with different methodologies) to get a general impression of the evolution of economic inequality (Lindert and Williamson, 1980).

The tax-derived estimates by Piketty and Saez produced a highly influential time series of distributional data for the top income shares in the United States. They depict a century-long pattern of changing income concentration whereby top distributional shares follow a distinctive U-shaped curve. In this account a period of high inequality during the early 20th gave way to a mid-century trough during World War II, followed by a rebound in the present day from an inflexion point located in the 1980s. The major movements in this series coincide with the rise and decline of extremely steep marginal income tax rates on top earners during the middle decades of the century, which are strongly implied to play a causal role in the witnessed distributional shifts. While a U-curve pattern is a commonality shared by almost all countries for which we have century-long measures, the depicted shape for the United States is particularly pronounced. Using the PS estimates, top U.S. income shares exhibited a significantly larger mid-century decline than most (though not all) western countries, and similarly rebounded at a faster rate in the late 20th century (Lindert and Williamson, 2016; Atkinson et al., 2011).

In addition to becoming the standard depiction of U.S. income inequality over time, the PS fiscal income series has provided a baseline for multiple subsequent empirical studies including a directly derivative measure of U.S. wealth inequality and subsequent attempts to expand

distributional estimates to lower income fractiles by linking them to national accounts (Saez and Zucman, 2016; Kopczuk et al., 2010; Lindert and Williamson, 2016; Piketty et al., 2017).¹ Although the PS technique is widely employed, the derivation of distributional data from tax records is not without its difficulties. For the period after roughly 1962 (when IRS microfiles become available), there are many large-scale surveys that are used to compare inequality measures according to the source material used to generate them (Feenberg and Poterba, 2000; Sommeiller, 2007; Autor et al., 2008; Burkhauser et al., 2012a; Frank, 2009, 2014). In some cases, the resulting depictions of top income and wealth shares show important discrepancies that result from choices in the selection, construction, and adjustment of source data (Kopczuk et al., 2010; Burtless and Svaton, 2010; Burkhauser et al., 2012b; Meyer and Sullivan, 2013; Armour et al., 2014; Kopczuk, 2015; Mechling et al., 2017; Bricker et al., 2016; Auten and Splinter, 2017).² These differing treatments of tax data and alternative measures tend to depict recent increases in both income and wealth inequality in the United States, but at very different magnitudes that show considerable sensitivity to data construction. These include a number of proposed adjustments to the PS series that reduce the rate of the depicted rise in income concentration between 1980 and the present as seen on the right side of the U-curve.

In this paper, we turn our attention to the comparatively neglected left side of the U-curve, encompassing the years preceding and leading into the mid-century U-curve trough. Sensitivity to data construction is likely more pronounced in this period due to the comparative incompleteness of older tax data sources, particularly under the pre-World War II income tax

¹As of this writing, the Piketty-Saez series has been cited in over 3,000 scholarly works on income inequality.

²It is also worth mentioning Reynolds (2006; 2012) who has probably been the most critical of using tax data for the United States to measure inequality. He argues that most (though not all) depicted changes in income inequality since the 1980s were the results of changes in tax-reporting behavior instead of underlying differences in income growth across the distribution, an improper accounting of in-kind compensation and non-taxable transfers, and the impact of the major Tax Reform Act of 1986 - the latter of which is echoed elsewhere (Mechling et al., 2017; Auten and Splinter, 2017). More controversial are Reynolds's criticisms of "top-coding" (extremely high income individuals earning above a certain threshold being given a fixed income level in reported data in order to protect their privacy) - we are agnostic on that latter issue as it does not affect our period of study.

regime. The statistical discrepancy arises from an unparalleled expansion of the federal income tax base between 1940 and 1945, extending its reach to most working Americans. Prior to 1940, actual tax filings often comprised less than 10% of all income earning tax units in a typical year due to an extremely high eligibility threshold and corresponding filer exemptions; most wage earners were not required by law to file a return. By the end of World War II, this figure ballooned to almost 90%, where it remains to this day. This shift profoundly affected the recording of federal income tax data, as well as the distributional characteristics of reported and unreported income.

Given the small size of the tax-eligible public prior to the war, the calculation of distributional estimates in the top income decile is highly sensitive to data construction. A series of adjustments are unavoidable in order to make these data comparable to postwar income records. PS only lightly examined these issues in their original series and did so by way of large and sometimes arbitrary assumptions that, as we shall see, introduced problematic elements into their results. We accordingly propose a series of revisions and improvements to their series that account for the unique data challenges affecting tax records in this period. The cumulative result is a corrected series covering 1917 to 1945, which may then be linked to the modern PS series after 1946.³

Before undertaking these corrections, we must carefully consider the underlying construction of the inequality series across the entire first half of the 20th century. To this end we revisit a series of adjustments that were made in the original PS series and that appear to have the effect of augmenting the overall height of top income shares, as well as accentuating the timing of specific movements in the U-curve. These adjustments are substantial at times, including an average reduction of over six percentage points annually from the top 10% income share during the pre-war period.

³All our measurements, original data, and calculations sheets are available upon request.

The purpose of this paper, we must emphasize, is *not* to assert that the United States was an exception to century-long international patterns in inequality. Neither do we question the idea that top income shares fell then rose over the 20th century. We accept the general occurrence of the "great levelling," subject to further investigation of its shape and magnitude. We show instead that the left-side of the PS U-curve likely overstates the original height and ensuing decline of top income shares, with implications for distributional pattern over time and for other derivative works that rely upon the PS series. Our adjustments do not change the fact that inequality fell between 1929 and 1945. What they do is attenuate this decline so that, when combined with other proposed corrections to the right-side of the U-curve, we see a much shallower "tea saucer" of inequality in the United States over the the course of the twentieth century. We confirm an inequality peak in 1928-1929, which is actually much more acute to those years than in PS. However, the "great levelling" (an expression we borrow from Lindert and Williamson [2016]) becomes a comparatively gradual story where the Great Depression is a more readily visible factor. This draws into question numerous inferential claims that wartime income tax policy and its subsequent entrenchment as a permanent feature of the mid-century tax system played a major role in the observed levelling.⁴

To keep the discussion manageable, we concentrate upon the core income distributional statistics from which top fractiles are calculated, namely the top 10%, 5%, and 1% income

⁴The idea of a curvilinear evolution of the distribution-wide income inequality for the United States is well supported. The closing of the wage gap between white and African-American workers (Higgs, 1977; Margo, 2016), between men and women (Goldin, 1990) and between northern and southern states (Mitchener and McLean, 1999) were strong contributors to the reduction of gaps in wage income during the early and middle 20th century (Lindert and Williamson, 2016). Alone, these forces would have been sufficient to cause a reduction in inequality. In addition, the comparatively pronounced U depicted by the PS series for the United States is in acknowledged tension with other closely related metrics and historical observations. As Piketty (2014, p. 152) notes "the shocks of the twentieth century struck America with far less violence than Europe." Piketty's (2014, fig. 5.1 and 5.2) estimates of the capital-to-national income ratio for the United States is also substantially flatter across the entire 20th century than the visible U-shaped swing that appears in Western Europe during the same period. This latter evidence is consistent with the United States' comparative insulation from the physical destruction of both World Wars.

shares over time.⁵ As our corrections are centered upon the core methodological assumptions regarding data treatment made by PS in the first three steps of their calculations, we do not examine the composition of incomes within the series. Our main findings therefore pertain to changes to the overall level of income shares among top earners.

Our revisions to the PS series proceed in three steps, after outlining the original data construction of PS in section 2. In sections 3 and 4, we revisit a pair of adjustments undertaken in the original PS study to account for issues of data quality and consistency in the pre-World War II period. While acknowledging the need for these adjustments on account of tax record limitations, we find evidence of further distortive effects that they each introduced to the PS series, generally causing them to overstate the depicted level of inequality and to accentuate its patterns of change over time. Specifically, these entail the use of an adjustment multiplier for tax filings that inadvertently captures the wartime distortions on the U.S. labor market through the selection of its base year.

After resolving the labor market issue, we turn to a second crucial adjustment in PS to account for the IRS's change in statistical reporting from "net income" to Adjusted Gross Income (AGI) in 1944. Using distributional data for pre-war tax deductions as a benchmark, we show that the PS adjustment on this point likely skews their series upward while also obscuring the effects of changing deduction patterns across tax brackets between 1917 and 1943. To address these problems we replace the original PS deductions adjustment with an improved estimate of AGI-comparable income derived from itemized tax returns. The cumulative result of these first two steps is a deductions-responsive harmonized series reflecting peacetime labor market conditions and dating back to 1917, the earliest year that tax records

⁵This restriction to income shares above the top 10% is consistent with both the constrained pre-war tax base and the absence of IRS microfile data prior to 1962. Subsequent attempts to extend distributional estimates to the full income range via national accounts are similarly constrained in this period, thus Piketty, Saez, and Zucman (2018, p. 566) actually link their own pre-1962 estimates to a baseline taken directly from the 2003 PS series. As such, our corrected series automatically implies a need to carry these adjustments over to subsequent studies of income and wealth that rely upon the original PS series.

permit.

In section 5, we investigate the assumptions made by PS regarding the size of the fiscal income denominator used to calculate distributional shares of income prior to 1944, along with an adjustment they apply to 1944-46. We show that PS adopt a denominator (based on a uniform adjustment to the personal income series from national accounts) that is consistently too low in the pre-war years, and that misses several important accounting corrections that are necessary to reconcile its measures with reported tax income and corresponding tax units. These considerations lead PS to overestimate top income shares before World War II by a sizable margin. We construct an improved tax-comparable denominator from national accounts to be used in place of the PS denominator. Combining this step with the aforementioned data corrections from our first two steps, we then construct cumulative adjustments for the top 10, 5, and 1 percent income shares reflecting the improved fiscal income denominator. Taken together, these corrections revise the locus of the timeline and pattern for the U-curve's levelling while also reducing top distributional shares throughout this period.

As a final point of inquiry (section 6), we explore one under-acknowledged issue of tax data quality that is specific to the pre-World War II period in the form of systemic income under-reporting due to issues of tax avoidance.⁶ To account for this issue, we employ the method used by Gene Smiley (1998; 2000) in his quantification of 1920s tax avoidance and extend it through 1941 in order to estimate the number of expected federal tax returns, holding tax rates constant. As a result, we can measure the income share of the top 10% without the variations caused by changes in tax rates and ensuing avoidance strategies. Our resulting estimate is meant to be similar in design to PS but without the variations in income reporting generated by the tax changes. Our results here suggest somewhat lower top income shares in this period and different patterns in the distributional evolution of top incomes during the

⁶There are other data quality issues, but discussing these would move us too far away from the discussion on methodological assumptions which is our aim here.

interwar period. While offered as a stand-alone adjustment and constrained only to the top 10% shares, these findings attest to the importance of accounting for further data quality issues with tax records in this period, consistent with the implication of our results from the first three steps.

The cumulative effects of our corrections and adjustments indicate a need to revise several longstanding assumptions about the shape and magnitude of income distribution changes in the first half of the 20th century. In reaching this conclusion, we note that historical tax data for the United States in the pre-World War II period should be considered very cautiously when evaluating inequality. The strong and often causal depictions attributed to the U-curve's patterns stand in marked contrast with IRS data sources that are of questionable quality, accuracy, and completeness. Due to the paucity of detail in pre-war tax statistics, seemingly trivial assumptions in the handling and adjustments of these data have large impacts on the resulting income distributions, as displayed in the PS series. All of these points converge to suggest that the left-side of the U-curve of income inequality in the United States warrants further refinement affecting both the magnitude and duration of its depicted leveling. We remain cautious about reaching strong conclusions over the level of inequality in this period, but by transparently displaying the trade-offs in the data, the assumptions underlying its adjustments, and our revisions to the overall series where data permit, one can see the need to exercise greater caution before hailing a pronounced U-curve of income inequality as a key stylized fact of American economic history, let alone attributing this pattern to fiscal policy design.

2 How Piketty and Saez arrived at their estimates

Before proceeding, it is necessary to understand how PS arrived at their original distributional estimates, including the reliability of their adjustments in light of the revisions that we propose and lay out. There are basically three interlocked steps to deriving the series constructed by PS prior to 1960 and presented in their work.⁷ The first step is to assemble distributional data from IRS Statistics of Income (SOI) reports by income class in their raw form. From there, Piketty and Saez (2007, p. 195) use a Pareto interpolation technique to assess how income is distributed above the top 10% threshold by matching their estimates to the closest corresponding income bracket in the annual SOI tables. This technique also permits them to calculate further distributional shares for the top 5%, 1%, and 0.1% of reported earners. No alteration is made to the IRS data in this first step, which for the purposes of this paper we call PS model 1. As such model 1 yields a “raw” unadjusted measure of the income distribution. As PS recognize, model 1 is a problematic measure due to dual issues of income under-reporting and the inconsistencies introduced by accounting changes in the IRS data collection after 1943.

The first adjustment made by PS (which we refer to here as model 2) attempts to correct for some of the issues created by the low number of tax filers in the pre-war period relative to the necessary thresholds to calculate distributional shares of earned income. Due to the small eligible tax base in the pre-war period, total filed tax return income actually dipped below the cutoff threshold for the top income decile in some years. This pattern did not substantially change until the dramatic tax base expansion primarily affecting lower level income earners under the wartime revenue measures of 1941-45.

⁷There are further adjustments that they make in order to more accurately incorporate capital gains into their calculations. However, since these adjustments are made after those that relate to Models 1 to 3, any modifications to these models will automatically reduce the capital gains-adjusted estimates as well. The adjustments are discussed in steps in Piketty and Saez (2007, 194-98).

To account for some of these problems, PS (2007, pp. 195-96) model 2 introduces an upward adjustment to the total number of married joint filers in lower income brackets (typically incomes under \$5,000) based on the assumption of a stable ratio between married and single filers in the same years (married filers enjoyed a higher exemption level than single filers, therefore the assumption is that single filer returns more accurately capture the share of tax filers in low-income brackets and can be used to adjust the total). The results of this adjustment apply to the PS income share estimates prior to 1940. Notably, almost all of the effect is seen in the top 10% income share estimate; the top 1% falls well within the threshold for both single and married filers, and therefore is largely unaffected by this step.

The PS model 2 adjustment is a necessary work-around to the problems created by the high eligibility threshold in this period. Some discretion must nonetheless be exercised in selecting the appropriate base year from which to calculate the married/single filer ratio. With only a limited number of suitable options given the underlying data quality issues, PS select 1942 to perform this adjustment. While the extrapolation method they use is adequate given the data limitations, the selection of a wartime year presents its own unintended complications as we discuss in a later section.

The next PS modification seeks to address changes to accounting practices that the IRS implemented in 1944, converting the pre-war figures to adjusted gross income (AGI). Prior to 1943, income earnings were reported as "net income." AGI encompasses a more expansive definition of income earnings than net income, extending its scope to include certain specified tax deductions above net income. Since not all eligible deductions are AGI-inclusive, accurately performing this conversion requires detailed distributional records for specific deductions categories. Importantly, PS do not utilize distributional records of deductions when harmonizing net income to AGI.

For sake of convenience in following their steps, we refer to the PS deductions adjustment

as model 3 in their series (Piketty and Saez, 2007, pp. 195-196). PS's method of performing their model 3 adjustment is substantially more arbitrary than the married filer adjustment in model 2. It also extends to the entire range of reported earnings, adding to the income shares of both the top 1% and top 10%. To perform their model 3 shift, PS utilize a set of stable and evenly-rounded multipliers for the entire pre-war period, which they then apply across the board to adjust the income share of each distributional fractile upward until 1941. They then reduce these multipliers for 1942 and again for 1943 until they link the resulting series with the unadjusted AGI data from 1944. Notably, the multipliers do not appear to reflect multiple changes to the tax code's permitted deduction categories between 1917 and 1943, or shifts in deduction patterns across income brackets. The multipliers applied to each income share and year are depicted in Figure 1 while Figure 2 shows each level of data treatment (models 1, 2, 3) resulting from the PS adjustments to the top 10% and top 1% time series.⁸ A key point that will be of relevance later in our discussion is that the deductions adjustments in model 3 are assumed in PS to tier upward across income brackets while remaining a fixed constant within each income bracket. The resulting adjustment therefore assumes uniformity across affected years, with the effect of simultaneously shifting the calculated distributions of the previous steps in a less-equal direction. Historical deduction patterns, however, were not uniform in either level or distribution across income brackets.

The model 3 adjustment has a dramatic effect upon the overall shape of the PS U-curve by unambiguously augmenting the depicted levels of pre-war inequality. In a given year, this adjustment alone adds between three and six percentage points to the top 10% income share. It adds roughly one to three percentage points to the top 1% income share. The

⁸We have attempted in vain to identify the empirical basis of these multipliers, particularly given the evenly-rounded numbers used and their constant application across multiple years of data. As best as we can tell, they are educated guesses. See Piketty and Saez (2007, p. 196). In their datasheet notes, PS indicated that they used what appeared to be "reasonable" adjustments. Concurring annotation may be found in their comp1398.xls file where the notes are in French : *je me suis contenté de reprendre des valeurs raisonnables à partir des tabcomp reproduits sur la feuille Comp1665.*

Figure 1: Deductions-AGI Multipliers

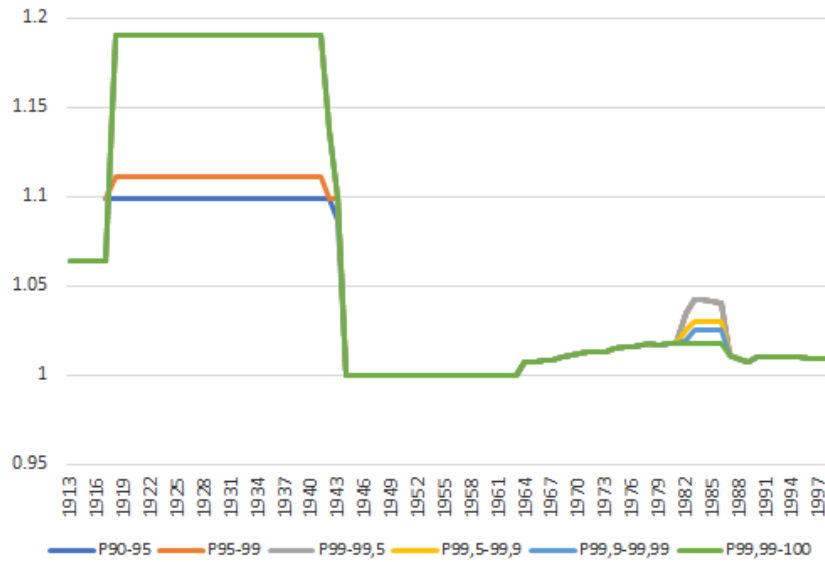
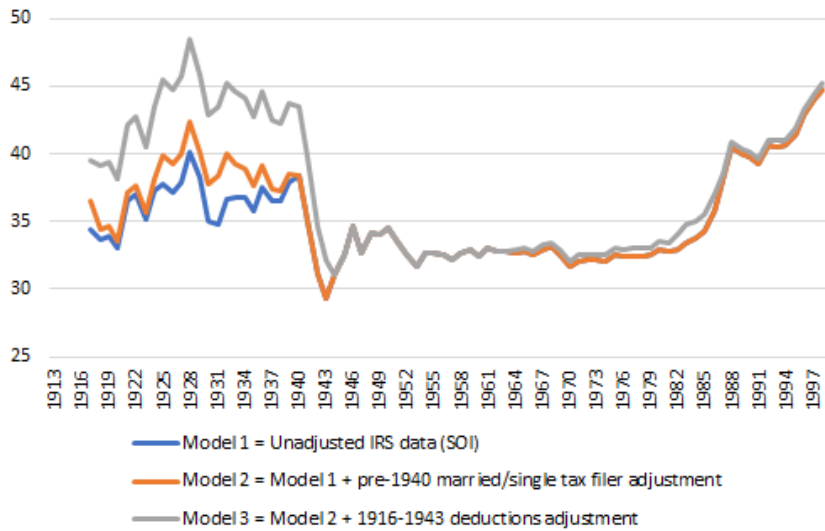


Figure 2: P90-100 Income Share, SOI plus Adjustments



most pronounced features of the first half of the U-curve, including its starting height and its sharp drop during World War II, are largely products of this single adjustment. As we will discuss in subsequent sections, substantial issues of data quality complicate the accuracy of this adjustment, and thus its effects upon the resulting U.⁹

To complete their distributional calculations following the steps of models 1-3, PS take the income estimates of the richest fractiles and divide them by the fiscal income of the total population as recorded in national accounts. For the period of 1913-43, their denominator series is equal to 80% of total personal income (minus transfers). As the national accounts for personal income produced by the Bureau of Economic Analysis begins in 1929, PS extrapolated the 1929 level backwards using the movements in an earlier series of personal income (those produced by Kuznets (1941, 1945, 1953)). This step yields corresponding top income shares from the adjusted tax data from previous steps.

Before continuing to our proposed revisions to the PS series, we must point out a final assumption that accepts the general reliability of the IRS data. This assumption applies to the present inquiry only, and for the sake of simplicity. Any set of data that is self-reported, as is the case with tax records prior to payroll deduction, will be affected by the incentives surrounding self-reporting as well as other associated complications of accounting. In the early days of the Internal Revenue Service, enforcement of tax laws can be considered comparatively lax (United States House of Representatives, 1916). Issues of avoidance, evasion, and simple inaccuracy were widely acknowledged in the pre-war period and became a recurring theme of tax enforcement policy discussions at the time. The numerous and sizable changes in the tax regime created extensive challenges for enforcement that likely extend to the data quality of

⁹The datasheets for the step-by-step calculations were provided by PS themselves and pertain to their 2003 article, covering the years 1916 to 1998. Saez has since extended this series through 2015, while also making a number of minor refinements to the published original series. For purposes of consistency in tracking the steps performed while estimating the pre-World War II period, we adhere to the original results and accompanying datasheets from the published 2003 version of their paper.

self-reported income returns.¹⁰ Addressing these additional problems exceeds the scope of the present study. That much noted, IRS records may systemically under-report actual income earnings during the decades preceding payroll deduction, and particularly at or near the minimum tax eligibility threshold where audit risks were low.¹¹ As evidence of this problem we note recent findings of substantial discrepancies in a comparison between federal income tax returns and independently collected state income tax returns for the same years.¹² This pattern of federal under-reporting is most visible in Wisconsin, where a relatively complete pre-war time series of distributional data from a stable state income tax regime exists (Geloso and Magness, 2017). We mention this because it is necessary to bear in mind that there are additional complications related to data quality that go beyond the assumptions employed in generating top income shares from IRS data.

3 Adjustments for "Missing" Tax Filers

Prior to 1940, the federal income tax had an extremely high eligibility threshold in most years – even sitting above the top decile of income earners. As a result, tax records only partially capture tax filers in the lower ranks of the top decile. To accurately measure the top 10% of income earners, an adjustment is necessary to effectively swell the number of filings at the bottom levels of tax-eligible income brackets to account for the “missing” returns. The PS model 2 adjustment (Piketty and Saez 2007, p. 195) attempts to account for this problem

¹⁰The federal government enacted major income tax revisions in 1916, 1917, 1918, 1921, 1924, 1926, 1932, 1934, 1936, 1940, 1941, 1942, 1943, 1944, and 1945. These include large changes from top marginal tax rates in the high seventies to the mid-twenties and back up.

¹¹Also worth noting, as this relates to the issue studied by Auten and Splinter (2017) regarding the role of the 1986 Tax Reform Act in causing artificial increases in inequality measurements, that business income under-reporting remained a major issue after the introduction of payroll deductions.

¹²Troiano (2017) estimates that the introduction of withholding, third-party information reporting, and information sharing between state and federal governments all increased reported top incomes disproportionately.

by extrapolating the missing returns from known ratios of single to married filers. Since the single filer threshold was much lower than married filers, it consistently encompassed the top decile of earners even in years where married filers sat inside the minimum threshold. PS accordingly assume that the ratio of married filers to single male non-head of household filers is stable across the entire pre-World War II period. After ascertaining this ratio, they are accordingly able to estimate a multiplier for the missing returns and add these to the lower brackets as applicable (typically those below \$5,000).

While the technique is necessary, it nonetheless requires the selection of specific tax reporting years to establish a base for the multiplier, as well as the calculation of a suitable married/single male filer ratio. PS select 1924 and 1932 as their base years, extrapolating a multiplier share from both as an internal check on the estimates. The married/single ratio requires a year in which reporting in both filer categories reliably exceeded the top decile's cutoff threshold, which limits the available options to the tax base expansion after 1940.

PS select 1942 to calculate their ratio, but in doing so they risk introducing a substantial distortion to this adjustment. Given the disruptions to the labor market caused by the United States' entry into the war, it is unlikely that the 1942 ratio accurately reflects the conditions of the previous two decades. In fact, we may reasonably expect that war-related labor distortions are most acutely realized among young and single men at the lower end of the tax eligibility threshold. 1942 is therefore a problematic choice.

Since the United States did not enter the war until mid-December 1941, this year is likely a closer approximation of peacetime labor markets among an admittedly slim number of choices. Due to the large influx of filers from the ongoing tax base expansion, the IRS did not fully tabulate returns under \$5,000 after 1939 and instead estimated them by statistical sampling. Samples were especially pertinent for the new 1040A form, which allowed earners below \$3,000 with no itemized deductions to file a simplified tax return. In selecting 1942, PS appear to

have misread the sampling methodologies for each year and erroneously concluded that its 1040A sample was more comprehensive than the same for 1941. The opposite is actually true, as the 1941 SOI report utilized a larger 1040A sample than the following year.¹³

When the married/single filer ratios of 1941 are compared to 1942, the wartime distortions on the labor market immediately become apparent with a surge in single filers in several of the lowest income reporting brackets. To address this distortion and other potential problems introduced by the war, we recalculated the married/single filer ratio using 1941 as our base year. The resulting missing income multipliers with a 1941 ratio tend to be lower across the board than if 1942 is used. This swap generally reduces the estimated income shares for the top 10% of earners in pre-war years, with the sharpest changes occurring in the 1930s when a high minimum filing threshold further accentuated the effect of these missing incomes. Before examining the full effects of this adjustment though, we first need to address the handling of deducted income for the same years, inclusive of this correction. The cumulative results are therefore presented after taking this additional step.

4 Linking the net income series with the adjusted gross income series

The next issue that needs to be discussed relates to the manner in which tax return data before and after the 1943-44 accounting switch are linked. As mentioned above, the IRS

¹³PS indicate in annotation to their datafile Corrections1840.xls that “Year 1941 not used for extrapolations because form 1040A statistics based on a sample that is not fully representative (see SOI 1941 pp. 51-52).” This is incorrect, and points to a separate special tabulation of 112,472 1040A forms from six states that were only used to study patterns in taxable and nontaxable returns under this newly adopted filing option. A more comprehensive set of aggregate 1040A data, however, are reported elsewhere in the 1941 SOI and utilize the IRS’s standard sampling technique for all incomes under \$5,000. The 1040A sample in 1941 consisted of 516,000 returns taken from all 64 IRS collection districts, or 5% of all 1040A forms (SOI 1941, p. 6). By comparison, the 1040A sample in 1942 is actually the smaller of the two and consists of 455,000 returns, or 3% of all 1040A forms (SOI 1942, p. 7).

moved from net income to AGI in 1944. As such, the two sets of reported income figures are conceptually different and require harmonization - which is what the PS series attempts to do with its model 3 multipliers. This adjustment is largely responsible for adding most of the depicted height to the first half of the U-curve. Given how basic the correction method used in PS model 3 is relative to its importance to the overall series, a strong case exists to improve upon it by using published deductions records instead.

To understand the significance of these adjustments to the PS series, we must first look to the differences between net income and AGI. Roughly speaking, the IRS defined net income as gross income minus total eligible tax deductions, as permitted in the tax code, in a given year before 1943. Total deductions differed from year to year, particularly as they concerned the treatment of reported property and capital asset losses and eligible work-related expenses. But they also included a subset of regular itemized deductions that comprise the bulk of the difference between net income and AGI. As per the 1944 standards, AGI encompasses gross income minus "above the line" source-specific deductions for trade and business expenses, expenses on travel and lodging connected to employment, eligible rents and royalties for the years permitted, certain depreciation and depletion amounts allowed in payment to life tenants and beneficiaries of property held in trust, and certain allowable losses from the sale or exchange of property as defined by the tax code in specific years. However, AGI still retains specific itemized deductions including charitable contributions, state and local tax payments, interest payments, casualty losses, and, after 1942, medical expenses that are not counted as part of net income in those same years.¹⁴ Several complications arise from near-constant

¹⁴AGI-inclusive deductions are explained in detail on pp. 20-21 of the annual SOI report for 1944. The 1944 tax code changes that brought about this accounting shift also introduced the option of taking a standard deduction in place of these itemized AGI-inclusive categories. We further note from an accounting perspective that the shift to AGI and the accompanying standard deduction likely affected post-war tax planning patterns by incentivizing different deduction practices. These effects may explain residual fluctuations in estimated distributional shares during the war itself, relative to the years that preceded and that followed the 1943/44 accounting conversion.

changes to the pre-war tax code that preclude simply adding total amounts of all claimed deductions to net income. Instead, an accurate deductions adjustment needs to differentiate these AGI-inclusive itemized deductions from the other deductions categories (e.g. work-related expenses) that comprise the difference between AGI and simple Gross Income.

IRS records of the deducted amounts are inconsistently complete in the pre-war period, varying by both category and year. Fortunately, full distributional data on charitable giving across tax brackets exist from 1922 onward. Remaining deductions records improve substantially after 1928 to include state and local tax payments and interest payments. A new AGI-inclusive medical expense deduction was also adopted into law for the 1942 and 1943 tax years, with complete distributional statistics reported. Combined, these deductions categories encompass the largest components of AGI-inclusive deductions.¹⁵ We use these extant records to construct an alternative to the PS model 3 multiplier, but first a number of additional issues justifying this substitution warrant mention.

For interrelated reasons, it is likely that the model 3 multiplier adjustments used by PS in lieu of a direct deductions adjustment are incorrectly tiered. First, as shown in Figure 1, PS held their multiplier levels constant from shortly after the start of their series until 1941. They then reduced their multipliers precipitously until connecting with the AGI transition in 1944. Second, PS scale their multipliers upward toward the highest income fractiles, such that tax filers within the 99th percentile, or the top 1%, are consistently assumed to deduct at a fixed rate that is higher than tax filers at the 90th percentile threshold. The justification offered by PS is that these deductions are small and that they primarily benefited the richest income earners (2007, p. 196). This likely explains the general design of the their multipliers including their decision to hold them constant across the affected years, as seen above in

¹⁵A handful of other eligible categories such as casualty losses from fires were not consistently reported. We exclude these by necessity due to inconsistent data, but note that they comprise a very small portion of claimed deductions in reported years and are unlikely to have a noticeable effect upon the overall adjustment.

Figure 1. Observable patterns from IRS deduction records create a substantial complication to this assumption.

A closer examination of total deductions by income bracket suggests that the recorded deduction patterns were actually quite volatile and uneven prior to the 1943 accounting switch. As can be seen in figure 3, depicting a sample of years with complete deductions data, most income categories saw deductions fall as a percentage of total income prior to 1943. However, some income categories including those around the P95 threshold actually increase. Moreover, in some years like 1937, AGI-inclusive deductions represented a greater share of income for filers at the P90 cutoff (8.28%) than those at the P95 cutoff (5.80%).

As such, it is not true that deductions patterns monotonically increase with income levels, or that they hold stable from year to year within the same income categories. To see a snapshot of these effects at the time of the net income/AGI accounting shift, we may compare AGI in 1944 with net income in 1943 by income categories. This allows one to see if the assumption holds across the two years under consideration. If the PS assumptions are correct, then plotting the ratio of 1944 AGI over 1943 net income should more or less follow a flat line with an uptick at the end of the income ladder.

This is not what we find, as can be seen in figure 4 where the *opposite* pattern is observed. The adjusted gross income of taxpayers earning \$1,500 or less in 1944 was more than twice as high as the reported net income in these same income brackets for the previous year. Moreover, the ratio falls as we move further up the income ladder. In other words, at the time of the net income/AGI shift, deductions were becoming substantially more important in lower income brackets. While we do not expect these ratios to hold across all deductions in all prior years, the severe skew of the 1943/44 data shows that we cannot reliably assume that pre-war deduction patterns simply tiered upward to the highest income percentiles.

Using charitable deductions as an illustration (as they constitute the earliest broken down

Figure 3: Deductions as percentage of net income categories

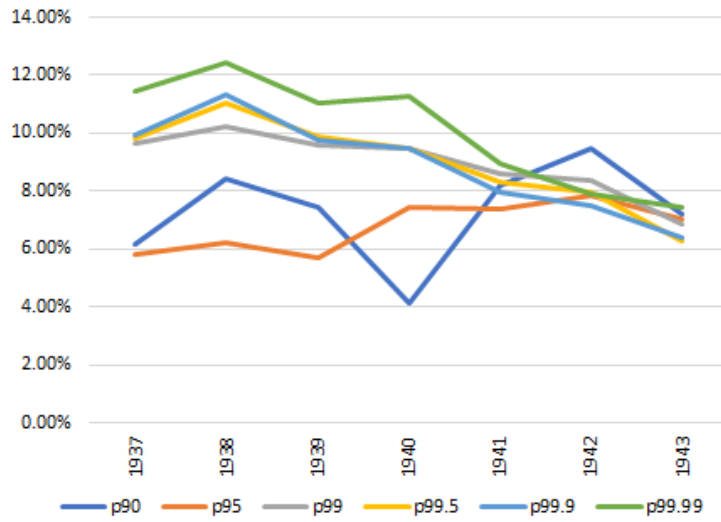
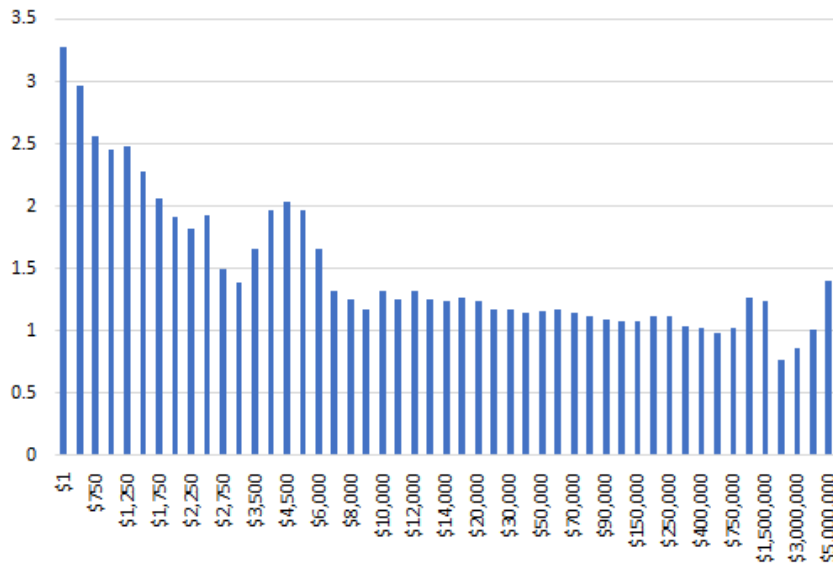


Figure 4 Ratio of 1944 AGI to 1943 NI by tax bracket



deductions - going back to 1922), further complicating patterns appear that run counter to the PS model 3 multiplier in the two decades preceding the war. Between 1922 and 1926 (there is a data break in 1927 for filers below \$5,000, requiring alternative estimation on which more is said below) filers with incomes below \$1,000 had charitable deductions equal to between 4.18% and 8.34% of their net income. This proportion was much higher than for most other income brackets. In 1926, the proportion at the \$1,000 level was the largest of all the income categories. Since PS presented their adjustments as the way to pass from model 1 to model 3 separate of the missing filer adjustment in their model 2, it is best for us to recalculate from model 1 with the charitable deductions in those years in order to assess the effect of this other method of turning net income into gross income. In figure 5 below, we can see that using this partially adjusted measure of income actually *lowers* the raw level of top income shares, *ceteris paribus*.¹⁶ Compared to the additions made with Model 3 by PS, these are clear signs that the harmonization of net income and adjusted gross income is inadequate. We understand why a shortcut as in PS model 3 was used for their original estimates.¹⁷ Nonetheless, given the doubts raised by figures 3, 4, and 5 there is value in attempting to generate an alternative to the Model 3 adjustment.

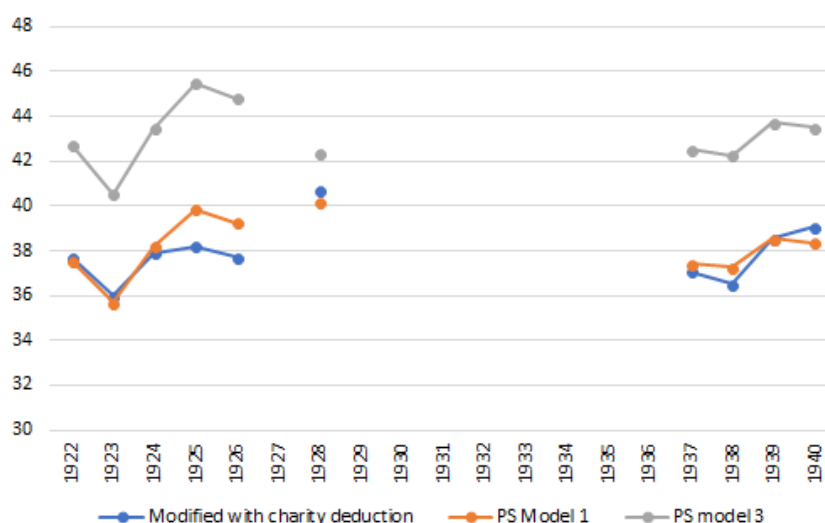
The third reason for doubting the model 3 adjustment is probably the most important: its cumulative effect is too large. While the SOI does not provide a detailed breakdown of deductions by income categories for all years, it does provide the aggregate totals for deductions by type in most years (1922 through 1927 and 1933 through 1943).¹⁸ These

¹⁶We note that our findings here are consistent with a well-established literature attesting to the crowd-out effects of New Deal spending on certain forms of charitable giving during the Depression era (Gruber and Hungerman, 2007). IRS-SOI records similarly reveal that charitable giving was not necessarily concentrated at the highest levels of eligible tax filers in certain years prior to the AGI accounting shift.

¹⁷It is worth pointing out the possibility that adjusting from net income to AGI might induce some re-ranking of filers between tax brackets. Aware of this issue, PS (2007, p. 196) examined the possible effects of the AGI conversion in later years where they may be externally checked against additional data from IRS micro-files. As they note, "using the micro-files for 1966-99, we have checked that this re-ranking has small effects on our final results and thus we do not attempt any correction for that re-ranking effect."

¹⁸For 1928 to 1932, deductions for brackets below \$5,000 are not included in cumulative totals for two of

Figure 5: Income Shares, Piketty-Saez AGI adjustment vs. Charitable Deductions Adjustments



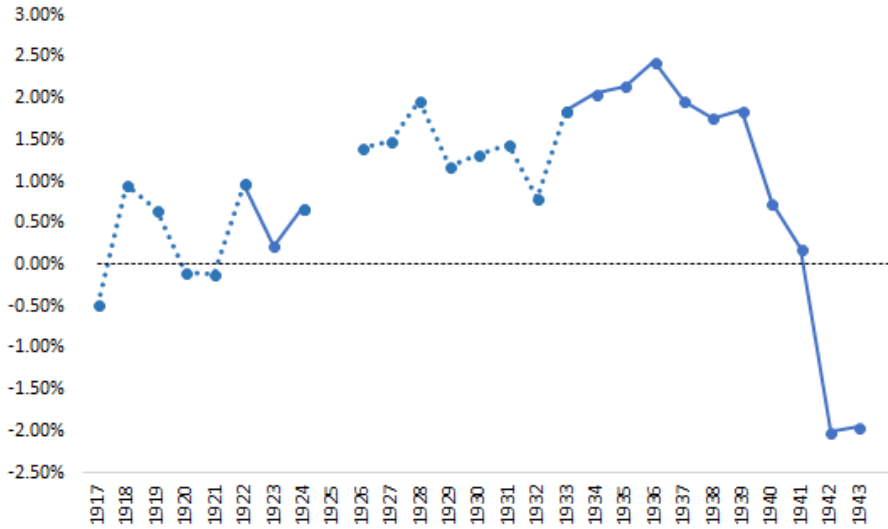
amounts can be compared to the total amounts of deductions that the PS adjustments imply (net of model 2). If the model 3 adjustments are valid, the two should match. However, this is not the case as can be seen in figure 6.¹⁹ Prior to 1940, the PS adjustment is consistently too large.²⁰ This pattern persists until the tax base expansions after 1940. As a result, the PS Model 3 adjustment likely has the effect of biasing the distributional shares of top income earners upward in the years prior to 1940, essentially ensuring the pronounced post-1940 decline of the U pattern over an extremely short period of time as opposed to a longer multi-year trend. Moreover, after 1940, their adjustments are likely *too small*. PS scale AGI eligible deductions downward from 1940 to 1943 in their multiplier, even though observed deductions patterns were moving in the opposite direction. A large share of this result came from making

the three AGI-inclusive categories. We resolve this issue below.

¹⁹1925-32 are omitted due to the need to estimate deduction totals for filers under \$5,000 for these years, using SOI records. Our deductions adjustment estimates the size these missing data by extrapolation from known SOI records, and suggests that PS overestimate the size of their multiplier in these years as well.

²⁰To illustrate the amplitude of this point, the year 1936 is informative. In that year, the overshooting of AGI by PS Model 3 is close to \$1.3 billion (the second highest amount for the period extending back to 1922). This amount is equal to 2.44% of their income denominator (the highest of the same period).

Figure 6: Difference between AGI-eligible deduction totals and implied deductions from PS model 3 multiplier



health care expenses eligible for deduction in 1942, thereby removing them from reported net income.

However, these are not irremediable flaws. They can be supplemented by more reliable data. In place of PS Model 3, we utilized SOI tabulations of claimed deductions by income bracket to estimate missing income found in the AGI-inclusive deduction categories. These encompass charitable contributions, state and local taxes, and interest payments for all years prior to 1944, as well as the medical care deduction that was in place for 1942 and 1943. Collectively, these itemized categories account for most of the difference between net income and AGI, as well as the majority of all deductions claimed including those that are not considered a part of AGI as per the post-1944 accounting definitions.

Due to sporadic and inconsistent record keeping by the IRS, it is not possible to isolate the full range of AGI-inclusive deductions in the earliest years of the income tax system. Recognizing these data limitations, we are nonetheless able to estimate and construct a consistent deductions adjustment for most pre-war years using the three aforementioned categories as well

Figure 7a: Adjusting Top 10% and Top 5% Income Shares for Deductions versus PS Model 3 adjustment

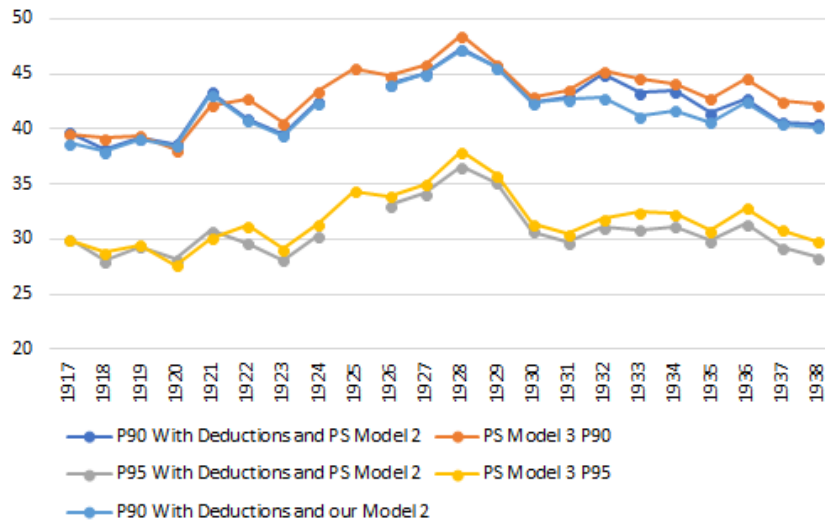
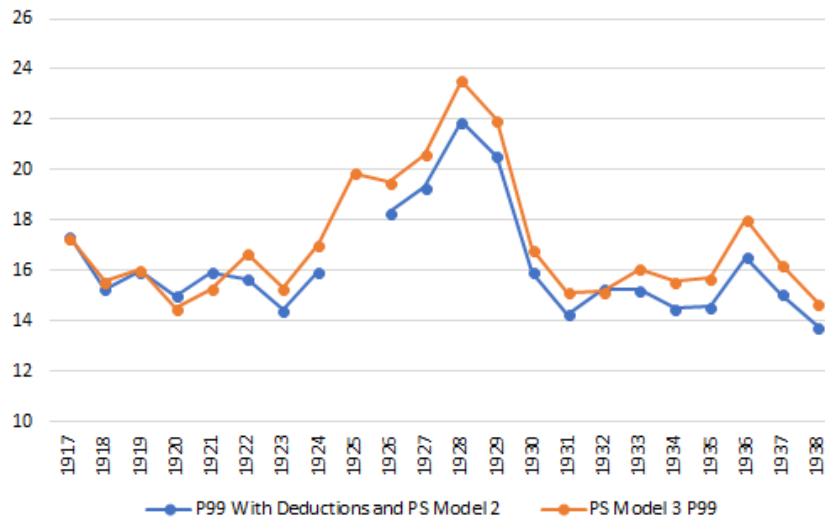


Figure 7b: Adjusting Top 1% Income Shares for Deductions versus PS Model 3 adjustment



as the late addition of medical expenses. To do so we begin by directly tabulating deductions from SOI records covering 1937 to 1943, which contain a complete distributional breakdown by income category for all of the aforementioned AGI-inclusive deductions categories. For earlier years, the IRS consistently recorded charitable deductions from 1922-1936, state and local taxes from 1927-1936, and interest payments from 1929-1936. Taxes and interest payment prior to 1927 and 1929, respectively, were reported together as part of miscellaneous deductions dating back to 1922, and can be extrapolated for these years.²¹ Finally, for the years 1917-1921, the IRS reported a separate total income tabulation that represented "gross income, less losses plus other deductions applicable" to specific income streams such as business travel expenses. This tabulation included income from a separate "general deductions" category, encompassing the main deductible income sources that were later included in AGI. This supplemental tabulation accordingly functions as a precursor to AGI for these years, subject to minor accounting differences, and allows us to approximate a consistent income share estimate, including deductions, going back to 1917.²²

²¹Several steps were taken to extrapolate and harmonize a consistent series of AGI-inclusive deductions prior to 1937. Our steps begin by identifying portions of the 1922-36 series that are incomplete in the SOI records. With the exception of a separately reported tabulation of charitable deductions for all brackets in 1928, the IRS only reported cumulative totals for charitable deductions claimed in income brackets under \$5,000 for the years 1926 to 1936. We accordingly impute them from the known totals using distributional ratios for reported years. To get a comparable series for state and local taxes and interest payments, we estimated the distributional share for these lower income brackets by using the ratios for the same brackets in known years under the closest comparable tax regime (thus 1937 for all three categories between 1933-1936, 1928 for charities in 1927 and between 1929-1932, and 1926 for taxes and interest taken together for 1926-1932) and joined them to complete figures for income brackets over \$5,000. Since state and local taxes are reported as part of miscellaneous deductions prior to 1927 as are interest payments prior to 1929, we take the ratios of taxes and interest payments to total miscellaneous deductions for 1929, by each tax bracket, and use this to impute taxes and interest payment figures as a share of all miscellaneous deductions for 1926-1928. Finally, we use reported aggregate amounts for taxes and interest payment deductions for 1922-1924 to estimate the share of each deduction from miscellaneous deductions for those years. When assembled in full, these steps give us a complete total deductions adjustment by tax bracket encompassing the three main AGI-inclusive categories for 1922-1943, along with medical deductions for 1942-43. The lone exception is 1925, which is omitted from our adjustment due to lack of reported aggregates necessary to calculate a suitable ratio for tax and interest payments from miscellaneous deductions.

²²The total income tabulation equaled net income plus state and local tax payments, interest payments, charitable contributions, and a small amount of miscellaneous general deductions for 1918 to 1921 (the miscellaneous category included non-business related "bad debts" that are not encompassed in the 1944 definition of

Collectively, these series allow us to estimate an AGI-inclusive deductions adjustment by income class going back to 1917, excluding only 1925 due to insufficient data for that year. We then integrate the distributional totals of eligible deductions into the reported income totals of their corresponding income classes, inclusive of the revised PS model 2 "missing filer" adjustments discussed in the previous section. By integrating these figures into reported net income, we thereby obtain SOI-based estimates of the actual effect of eligible deductions by income class each year.²³ The resulting aggregate gives us a functional approximation of the main AGI-inclusive deduction categories, capturing both their distributional skews and their yearly fluctuations in response to surrounding economic events and tax regime changes. We use this figure in place of the fixed model 3 multiplier from the original PS series, yielding a deductions-responsive income distribution estimate for most years in the pre-war period that can then be linked to the post-1944 AGI-derived estimates from the original PS series.

Figure 7 above shows the effects of adjusting for deductions across the top 10%, 5%, and 1% income fractiles. This figure compounds the married/single ratio year adjustments to

AGI, however these appear to account for a very small portion of the difference in the years in question). The tabulation is conceptually the same for 1917 with the exception of charitable contributions, which were omitted and not reported separately. Although these factors likely lead to a slight under-count in earned income for 1917, it remains a better approximation of the other AGI-inclusive deduction categories by tax bracket than the rough guesswork of the PS multiplier. We accordingly use reported figures for 1917 to calculate an adjustment ratio for net income in that year, and link it to the AGI-comparable total income tabulations for 1918 through 1921. Note that our 1917 adjustments (which amount to a difference of 0.85 percentage points between ours and PS, exclusive of other adjustments) likely overestimate the level of inequality in that year. The reporting discrepancies of the early IRS figures for these years are detailed on pp. 26-27 of SOI for 1919. Increasing complexity of the tax code with respect to the categorization of loss deductions and work-related expenses unfortunately limits the accuracy of this approach in later years, following the separation of charitable contributions from general deductions for reporting purposes in 1922 as well as the addition of new income source-specific deductions that fall outside of AGI in the 1920s and 30s. We therefore adopt the adjustment method described in the previous footnote for the years 1922 onward, as data reporting permits.

²³In this step we also supplemented reported income figures for the years 1941-1943 with additional income bracket data from the 1040A form, which was tabulated separately by the IRS and therefore not included in the original PS calculations. The 1040A form allowed earners with incomes less than \$3,000 and no deductions to file a simplified form reporting their "gross income," which in this case requires no further adjustment to harmonize with AGI. The addition of these records does not substantially affect resulting income shares beyond a small fraction of a percentage point, however we included these unadjusted figures in addition to our deductions adjustment to the regular deductions-eligible 1040 returns to improve the comprehensiveness of income filer accounting in affected years.

model 2 to the effects of deductions, but only for the top 10%.²⁴ We find that the shares are overstated on average by 1.31 percentage points for the top 10%, 0.95 points for the top 5% and 0.8 points for the top 1%. The greatest differences are concentrated during the period from 1932 to 1939, with the largest discrepancy reaching 3.4 percentage points in 1933 within the top 10% share.

5 Fiscal Denominators and Tax Units

5.1 Correcting Fiscal Denominators

To complete our corrections to the PS series, we turn next to the total fiscal income denominator over which top tax-reported incomes are divided to arrive at estimated distributional shares. The sum of all incomes on all tax records does not equal total personal income earned in the same year.²⁵ This distinction has been well-covered in the work of Kuznets (1953), Goldsmith (1951; 1954; 1957) and, more recently, Fixler and Johnson (2014). As such, it is necessary to make adjustments to either the numerator or the denominator to bring the two terms into methodological alignment before using them to calculate distributional shares.

PS begin their own denominator construction by taking the official personal income series (NIPA) from the Bureau of Economic Analysis, and subtracting transfer payments. By comparing this adjustment to tax-reported AGI, they discover a relatively stable ratio between the two terms lasting from shortly after World War II to the present day. Based on this observation and recognizing a residual between the two terms that derives from the small number

²⁴This is because the Model 2 adjustment affects largely the top 10% share alone. Higher fractiles generally fall within the exemption threshold for married filers.

²⁵For example, there will be differences between "fiscal income" and "personal income" because of imputed rent, interest and dividends received from pension plans, life insurance carriers, non-profit institutions, non-taxable employer and employee contributions to non-monetary forms of compensation, capital and inventory adjustments - see Fixler and Johnson (2014) for greater details. An additional residual difference may also be expected from under-reporting on income tax returns.

of persons falling below the postwar tax eligibility threshold, they then adjust AGI upward by 20% thereby producing a denominator for 1944 to the present.

Two complications limit the use of this approach prior to World War II. First, the official NIPA series only begins in 1929. In their work, PS accordingly extend the NIPA series back to 1913 by linking it to the movements of earlier personal income estimates prepared by Kuznets for the NBER.²⁶ Second and more importantly, the high tax eligibility threshold before World War II precludes a direct comparison between total tax-reported income and the national accounts series, even with extensions. Whereas the postwar approach relies upon a stable and expansive tax base extending to most income earners, the percentage of income earnings captured in tax records rapidly drops as one moves back further in time from 1945. The prewar denominator must accordingly be taken directly from the NIPA personal income series, as extended backward using Kuznets.

To account for residual differences between NIPA and conceptual tax-reported income, PS apply a uniform adjustment equal to 80% of personal income minus transfers across the entire series between 1913 and 1943 and then link these results to the denominator they derived by the aforementioned method from 1944 to the present.²⁷

The selection of a uniform 80% adjustment to the extended NIPA series raises a number

²⁶PS indicate that they used the work of Simon Kuznets on national accounts that covers 1913 to 1939 to extrapolate backwards (Kuznets, 1941, 1945). The data for personal income contained in their sheet ComptaNat.xls reflects subsequent tabulations from these two series undertaken by Kuznets in 1953 (p. 571) for his own estimation of top income shares.

²⁷PS (2007, p. 169) do not offer a justification for selecting 80%, other than that it appears to roughly approximate the 77-83% ratio of their postwar total income series to the NIPA personal income series after transfer payments are removed. Moreover, PS actually adjust their denominator as derived from the postwar method for the years 1944 through 1946 using a different set of adjustments to AGI. In their 2007 paper (p. 169), PS state that they assume "to non-filers a fixed fraction of filers' average income" at 50% in 1944-45 and 20% thereafter. Their code files (see "AGI vs Personal" in their ComptaNat.xls) use slightly different proportions though. They set the adjustment equal to 50% in 1944, 60% in 1945, 25% in 1946, and 20% for all years thereafter. To account for this discrepancy we extend our full analysis through 1948, allowing it to be linked to the remainder of the PS series using the alternative denominator approach permitted by the postwar tax base. Other than this step, the construction and reliability of PS's postwar denominator exceeds the scope of the present study

of complications for the resulting PS distributional estimates. As we shall see, calculated top income shares are extremely sensitive to discretionary choices in the construction of the denominator. As a point of reference, it is useful to compare the PS approach and adjustment ratio to earlier denominator estimates provided by Kuznets (1953) in his own attempt to measure top income fractiles for the same period. Rather than applying a fixed adjustment ratio, Kuznets attempted to remove several categories of nontaxable income on a year-by-year basis to produce a comparable series.²⁸

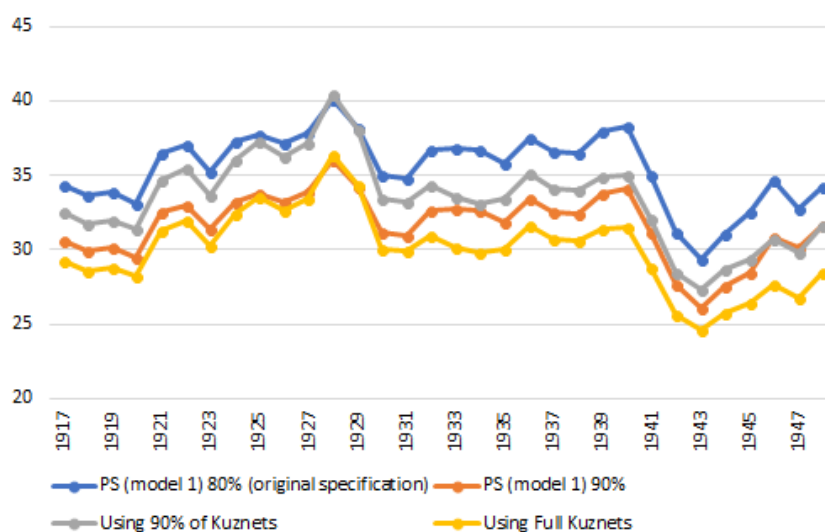
The Kuznets denominator is consistently higher than PS. Including adjustments, Kuznets's series comes to between 110% and 126% of the denominator series generated by PS for the years 1916-1948. PS attribute this discrepancy to an insufficient retraction of items from the residual difference between the two measures, yielding a denominator that is too high. While PS are correct that the Kuznets adjustments may be insufficiently complete in cumulative, they defend their own methodological choice for years prior to 1944 by arguing that using the Kuznets series would yield an implausible result where average non-filer income potentially exceeded average filer income. We note that in doing so they base this decision on calculations from the postwar period. However it is not apparent that their criticism holds prior to the 1943-44 tax code and IRS accounting revisions.

To make the case that Kuznets' denominator (1953, pp. 570-577) is too high, PS cite only that in 1948, the average income of non-filers would be higher than those of filers (2007, p.

²⁸PS incorrectly report that Kuznets only adjusted his personal income series for imputed rent. This is incorrect as Kuznets also adjusted his series downward to account for other differences that go unacknowledged in PS. The most notable is an adjustment for tax-exempt state and local government employees prior to 1939, of which more is said below. This tax exemption adjustment was substantial as compensation of state and local governments overshadowed imputed rent in every year, sometimes by as much as a four to one ratio. It is also important to point out that Kuznets argued that he overestimated his own deductions, indicating that the actual denominator should be greater than he reports. In discussing the "dissimilarity" (1953, p. 407) between the Wisconsin data and the IRS data for the state, he mentions that he may have overestimated some adjustments to the top income shares (p. 406). He also pointed, in a footnote, that limited state level data for Delaware (p. 406) overestimated his corrections and as such meant that he likely overshot his inequality estimate.

169) if they used that series. They deducted reported income from personal income (minus transfers) to get the income earned by non-filers and divided this result by the difference between the estimated number of tax units and number of tax returns. Given that tax filers tended to be high-income earners, we may intuitively conclude that they should have had higher average earnings than non-filers. The finding that non-filers had higher average incomes than filers in 1948 therefore implies that the Kuznets series is too high for the purposes of a fiscal income denominator. In this case it would overestimate the denominator needed to calculate the income share of each fractile. As such, their argument appears reasonable. However, while PS are correct for 1948, this problem does not appear to apply before 1943. It is only after 1943 (when automatic payroll withholding is introduced to the tax code) that non-filers start showing income earnings that are greater than those of filers under the Kuznets method. Throughout the 1917 to 1941 period, the ratio of the implied income of non-filers to filers varies between 18% and 46%. It is only in 1942 that the ratio starts increasing while remaining below 100% (at 86%) and in 1943 that it surpasses 100%. The pre-1941 proportions are well in the range of their later estimates for non-filers when they imputed to non-filers a fixed fraction of filers' average income of "50% in 1944–45, and 20% thereafter" (2007, p. 169). As such, the Kuznets series yields very plausible results based on PS's own benchmark for imputing the income of non-filers. It is only after the 1943-44 AGI shift that this breaks apart to support the contention made by PS. As an added complication, PS offer their critique based on Kuznets' unadjusted personal income figures. Once we retract Kuznets' adjustments for imputed rent and tax exempt government employees, the ratio of the income of filers to non-filers varies between 16% and 43% between 1917 and 1941. Subject to certain important modifications, the Kuznets denominator appears to more accurately capture the yearly movements of tax-comparable fiscal income, as derived from national accounts, than the fixed 80% adjustment ratio used by PS.

Figure 8: Top 10% income shares (PS model 1) depending on fiscal denominator



Even if we are to discount Kuznets' adjustments in favor of PS, only slight variations to the 80% adjustment ratio can potentially change the estimated top 10% income share by several percentage points. The extreme sensitivity of estimated top distributional shares to the denominator selection may be seen in Figure 8. Using the "raw" data of PS model 1 (exclusive of any other adjustments to the data in our previous steps), a full Kuznets denominator would reduce the top 10% income share by an average of 5.42 percentage points between 1917 and 1948. Shifting to a middle ground where we take 90% of the Kuznets estimate to include more adjustments, we obtain a reduction of 2.08 percentage points for the same period. Even altering the PS adjustment ratio from 80% to 90% of personal income minus transfers would yield an average reduction of 3.83 percentage points. As such, the justifications for the appropriate denominator choice warrant a more thorough consideration than PS provide.

Considerable evidence suggests that the PS denominator, when set at a uniform 80% of personal income minus transfers, is consistently too small. Some of it is provided by PS

themselves. For the period from 1944 to 1974, their ratio of fiscal income to personal income (excluding transfers) is equal to 82.7% on average. From 1947 to 1974, this average jumps to 83.04% and remains stable around that level.²⁹ After 1974, the ratio falls constantly to a low of just over 77% in the late 1990s. This is not what is implied by the claim made by PS (2007, p. 169) that the postwar ratio "fluctuates between 77% and 83%". Rather, the data suggest that there is a stable plateau (averaged at 83.04%) from 1947 to 1974 and a transition to another regime (in the high seventies) up to the 1990s. As such, their adopted ratio implies - without explanation - that a pre-war "average" of 80% jumped to a new plateau roughly 3 percentage points higher in the postwar period. If we applied the 1944-1974 ratio of 82.7% to the 1917-1943 period instead of the 80% used by PS, even this modest modification reduces the top 10% income share by 1.35 percentage points.

Other empirical analyses point in the same direction. As part of an effort to study the characteristics of the taxable income base of the 1940s, Selma Goldsmith (1951) undertook a series of comprehensive accounting adjustments designed to bring the BEA personal income series (now NIPA) into comparison with other measures of fiscal income. This approach yields a much more comprehensive adjustment than the Kuznets retractions. Starting with national accounts, Goldsmith accordingly removes all duplicated, nontaxable, and non-money income categories that are present in the personal income series to yield an estimate of total fiscal or money income retained by consumers. Although they are geared for comparison with field survey data on consumer income, the adjustments she makes for 1941 illustrate parallel implications for tax-reported income accounting. Here Goldsmith begins from a comparable baseline as PS, reporting personal income exclusive of transfer payments at \$91.6 billion (PS begin from \$93.6 billion before the 80% adjustment). Goldsmith's accounting reconciliation reduces this number to \$86.1 billion in consumer money income or 94.03% of

²⁹The years between 1944 and 1946 appear to be a war-related outlier: they fluctuate between 78.7% and 79.9% before jumping 82.5% in 1947. For the period from 1947 to 1974, the ratio is higher – at 83.04%.

the initial level of personal income minus transfers. Goldsmith's adjusted amount is some \$11.3 billion above the PS denominator for 1941 after their 80% ratio adjustment, or \$74.9 billion.

In order to resolve the empirical problems introduced by the PS 80% adjustment, we adapt Goldsmith's reconciliation approach to the full period of 1917 through 1948 (the inclusion of 1944-1946 allows us to link our results to the remainder of the PS series after 1946 while also accounting for additional wartime discrepancies). In doing so we follow additional steps that Goldsmith devised specifically to bring the NIPA series into comparison with tax-reported income from SOI.³⁰ This approach was subsequently adopted in BEA analysis for comparison of the NIPA series with IRS data in the postwar period and extends to the present day literature on tax revenue accounting.

To extend the reconciliation approach to the prewar and wartime period we must also address two time-specific components of tax accounting. The first pertains to tax-exempt military pay subject to the Combat Zone Tax Exclusion. This exemption originated in 1918 as a World War I soldier pay benefit until being rescinded in the 1921 tax year. The Military Pay Act of 1942 revived the policy, which persists to the present day as a feature of the tax code. (Gould and Horowitz, 2011) At its peak in 1945, affected military pay amounted to \$22.6 billion, or almost 14% of personal income minus transfers. As a result, the amount of income that needs to be removed from the denominator for reconciliation is larger between 1942 and

³⁰Goldsmith's (1951, pp. 359-62) reconciliation technique removes the following from personal income: transfer payments, other labor income, tax exempt military pay subject to the Combat Zone Exclusion, imputed rents, the value of goods and services received in-kind, home-grown food and provisions used on farms, changes in farm inventories, non-corporate non-farm inventory valuation adjustment, tax exempt interest on state and municipal bonds (unreported on tax returns), accrued interest on unredeemed government bonds, and a handful of smaller accounting adjustments pertaining to life insurance providers, non-profit organizations, and retained fiduciary income. Items to be added from SOI include an offset for employee contributions to social insurance, a net gains adjustment for income from the sale of capital assets and property, an adjustment for the estimated taxable income share of Alaska and Hawaii (which were excluded from national accounts totals prior to the 1960s), and a handful of miscellaneous tax-reported income sources that fall outside of the national accounts definition. Further extensions of this method include Goldsmith et al (1954), (Pechman, 1956), (Farioletti, 1958), and (Joint Economic Committee, 1965)

1946 than surrounding years - a discrepancy that is absent from PS.

The second issue stems from a quirk of two 19th century Supreme Court decisions that effectively exempted most state and local public employee salaries from federal income taxation in the early years of its operation. A combination of revenue and equity concerns, as well as changing judicial treatment of the matter, induced Congress to extend tax eligibility to these employees under the Public Salaries Tax Act of 1939. For prior years, this exemption introduces a sizable segment of missing income that affects the SOI-derived portion of the PS series. This distinction potentially carries substantial implications for the accuracy of all measured income in PS, above and beyond our corrections to the SOI statistics in the previous two sections.³¹ Absent a means to reliably estimate their distributional characteristics in the numerator, the income of these employees must be subtracted from the denominator side for each year prior to 1939.

We construct our tax-reconciled denominators in two components. For 1929-1948 we begin with the published NIPA tables, remove affected non-taxable and non-money income lines directly, and add missing income sources that fall outside of national accounts as per Goldsmith. The state and local public employees adjustment is further retracted from our results for all years prior to 1939. Data for almost all terms are available with the exception of a handful of smaller categories where reliable estimates do not exist. We also adjust the entire series for the estimated income earnings of Alaska and Hawaii, which were subject to tax reporting in

³¹See *McCulloch v. Maryland*, 17 U.S. (4 Wheat.) 316 (1819) and *Collector v. Day*, 78 U.S. (11 Wall.) 113 (1871). A study produced by the Department of the Treasury in 1937 (Congress, 1939, p. 26) estimated that the exemption affected some 2.6 million state and local employees with \$3.6 billion in untaxed earnings (by comparison the IRS received 6.3 million tax returns in the same year with unadjusted reported income at \$23 billion). Furthermore, income brackets for eligible public employees appear to have clustered near the currently estimated cutoff level for the top decile income share, suggesting they may have implications for the position of the top 10% relative to other top income fractiles as well as the premises of the model 2 adjustment in PS. Unfortunately the 1937 Treasury study is the only year for which distributional estimates of affected persons by income bracket appear to exist, although congressional attention to the tax problems created by state employee exemption dates to at least 1926

SOI even as they were not included in national accounts prior to the 1950s.³² Goldsmith also includes an adjustment for net sales of capital assets and property within the reconciliation, however we separate this step from our calculations on account of the distinct treatment of capital gains in PS and reintegrate it at a later point so as to preserve and follow their steps for calculating average income, as used in obtaining top income fractiles.

For 1917-1928, the reconciliation steps are conceptually identical, including the removal of state and local government employee income as well as the Combat Zone Tax Exclusion for eligible years during and after World War I. In place of NIPA, which begins in 1929, we use the older Kuznets NBER series to perform the adjustment. Corresponding data for retraction are available for most lines from Kuznets' work and related NBER studies from the period. We supplemented these figures with additional data series from the Department of Agriculture (farm inventories and products for consumption) and other federal data sources such as the Statistical Abstract of the United States. For remaining items where few records exist, we provide estimates linked to reported amounts from the 1929 NIPA series.³³

³²Omitted lines include two small adjustments for retained fiduciary income and for property income from non-profit organizations, both of which were estimated for specific years separately of national accounts in Goldsmith's calculations and are not readily available prior to the 1940s. We calculated imputed farm rent, which is not separately recorded in NIPA, using a ratio that follows Kuznets' similar estimates of total imputed rent for his adjustments. Prior to 1950, there are no consistent personal income data for Hawaii and Alaska (? , p. 36). At that point in time, these states were still territories that represented 0.51% of the national economy Federal Reserve at Saint-Louis (2018). As such, the personal income denominator excludes those two areas, which is problematic given that the IRS' SOI data includes them. A survey of the different editions of the SOI show that PS used the total national reported figures without removing Hawaii and Alaska. This means that there is a need to insure geographic uniformity in the top income shares. The easiest way to create this uniformity is to augment the denominator to include the income stemming from the two territories up to 1943. A 1955 report from the Bureau of Labor Statistics which estimated per capita personal income in Hawaii in 1939 at \$525 (not adjusted for inflation)(Ige, 1955) which meant that Hawaii represented 0.30% of personal income in the 48 continental states. That is roughly the same proportion that Hawaii represented in 1950 (0.34%), and is consistent with Goldsmith's separate estimates for Alaska and Hawaii personal income for 1944-46. As such, we augmented the personal income minus transfers series by 0.51% (the proportion of the US which the territories represented in 1950) to arrive at a geographically consistent definition. This has the effect of reducing the top 10% income shares (as measured through PS Model 3) by 0.2 percentage points in the 1917 to 1943 period. Although this is not a sizable adjustment, it adds to the other elements which suggest that PS set a denominator that was too low.

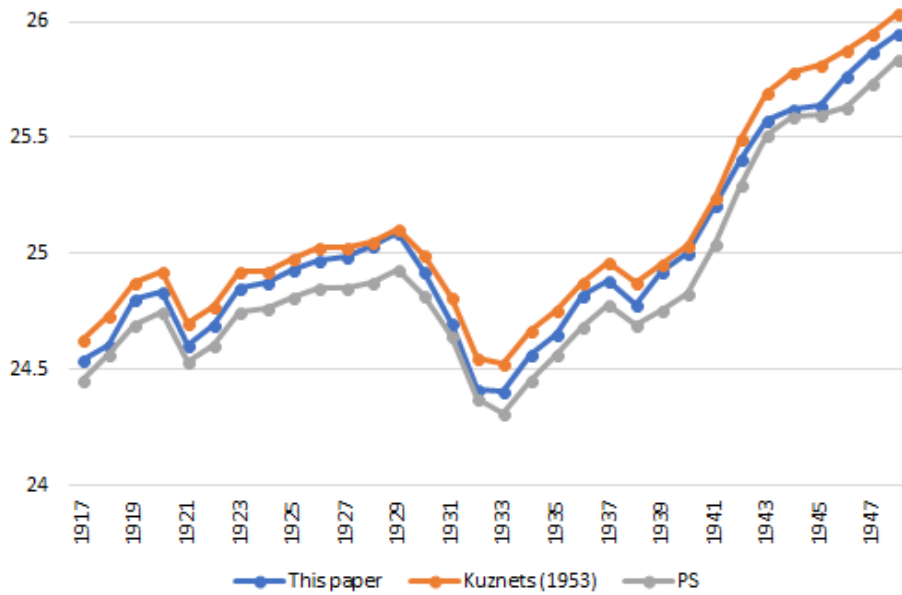
³³Transfer payment statistics are the most important incomplete line for these years, although they represent a substantially smaller portion to remove from national accounts prior to the establishment of Social Security

To complete the denominator series, we then harmonize the tax-reconciled 1917-1928 component with the reconciled NIPA series at 1929.³⁴ Once harmonized, these steps yield a national accounts-derived personal income baseline that is directly comparable to tax-derived income from IRS records, covering 1917-1948. The results of the reconciliation further show that the PS 80% denominator is consistently too small across the entire pre-war period. For example, the directly comparable 1929-43 portion of the reconciled denominator varies between a low of 83.6% and a high of 95.1% of personal income minus transfers. These results further show that the fixed ratio of PS fails to capture yearly variations within the retracted components of personal income, including pertinent factors such as Depression-era inventory fluctuations and overlooked policy factors such as the state and local government employee tax exemption. The reconciled denominator does converge on the 80% ratio for a brief period in 1944 and 1945, but this is a direct effect of the aforementioned Combat Zone Tax Exclusion. When incorporated into distributional calculations, the larger denominator derived from the reconciliation approach generally reduces estimated top income shares albeit at different magnitudes that are more reflective of contemporary events such as the Depression. Figure 9 shows the difference in our resulting denominator which is systematically between those of PS and Kuznets' and is higher, on average between 1917 and 1943, by 11.3% to that of PS. In figure 10, we show the cumulative effects of our modifications from the previous sections

in 1936. Rather than impute this subtraction backwards at a fixed number as PS do, we use a tabulation of government employee retirement payments to account for their largest component in this period. An undetermined amount of transfer income was also included with "other labor income" in the early Kuznets series. Taken together, these two sources yield a plausible transfer payment estimate that is consistent with the 1929 NIPA series equal to 2.25% of personal income. We found no reliable estimates for miscellaneous personal in kind income (e.g. lodging, meals for domestic workers, and non-money financial intermediary services) for this period and accordingly estimated it by taking a fixed ratio from the 1929 NIPA series. The same approach was used for the transfer offset adjustment to account for social insurance contributions by employees, although this number represented a very small amount during this period (0.17% of personal income in 1929). Note that these estimates likely overstate the amounts necessary to remove in the associated categories as we move further back in time, thereby erring against our claim of more modest reduction to the denominator. We adopt them out of an abundance of caution, recognizing that this implies slightly higher top income shares than likely existed for the years affected by estimation.

³⁴Full reconciliation calculations and accompanying sources are available in our data appendix file.

Figure 9: Log of different income denominators



with the denominators produced here. For the top decile, our adjustments through this step remove an average of 4.68 percentage points between 1917 and 1943.³⁵ For the top centile, the corrected level is lower by 2.08 percentage points.³⁶ These are sizable differences that affect the evolution of inequality.

5.2 Consistency in tax units

There is one additional issue that needs to be addressed with regards to the denominator. Certain conceptual differences between national accounts and tax data also carry over to associated population statistics, which play an important role in establishing the number of tax units to determine top income fractile thresholds and to calculate distributional shares. Recall that most government employees at the state and local level did not have to file federal tax returns until the Public Salary Act of 1939. As Kuznets rightfully pointed out, it is

³⁵With the Kuznets denominator, the difference stands at above 7 percentage points.

³⁶With the Kuznets denominator, the difference stands at above 3 percentage points.

Figure 10a: Comparisons of PS Model 3 with cumulative effect of deductions and denominator adjustment (P99)

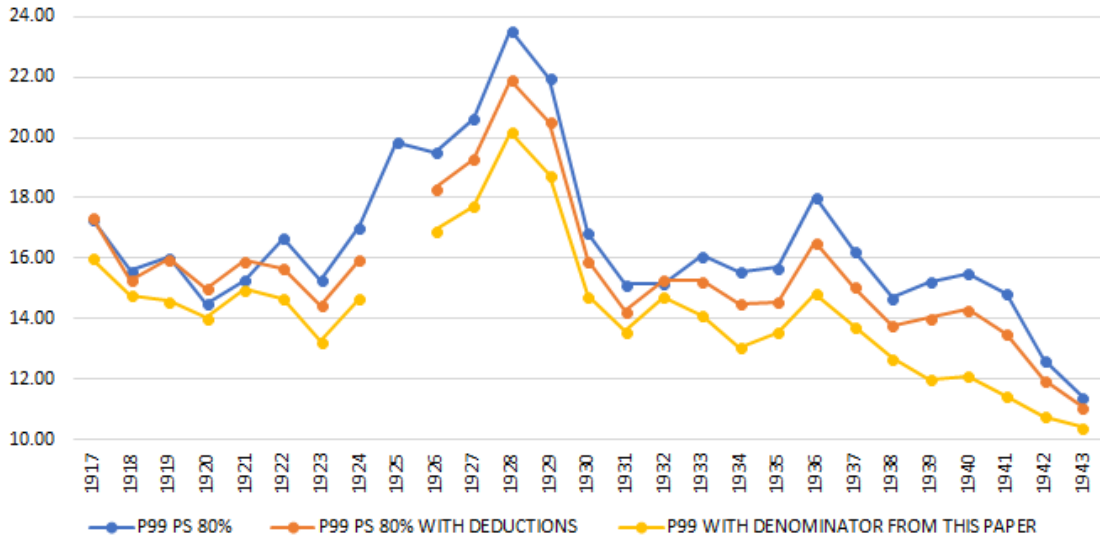
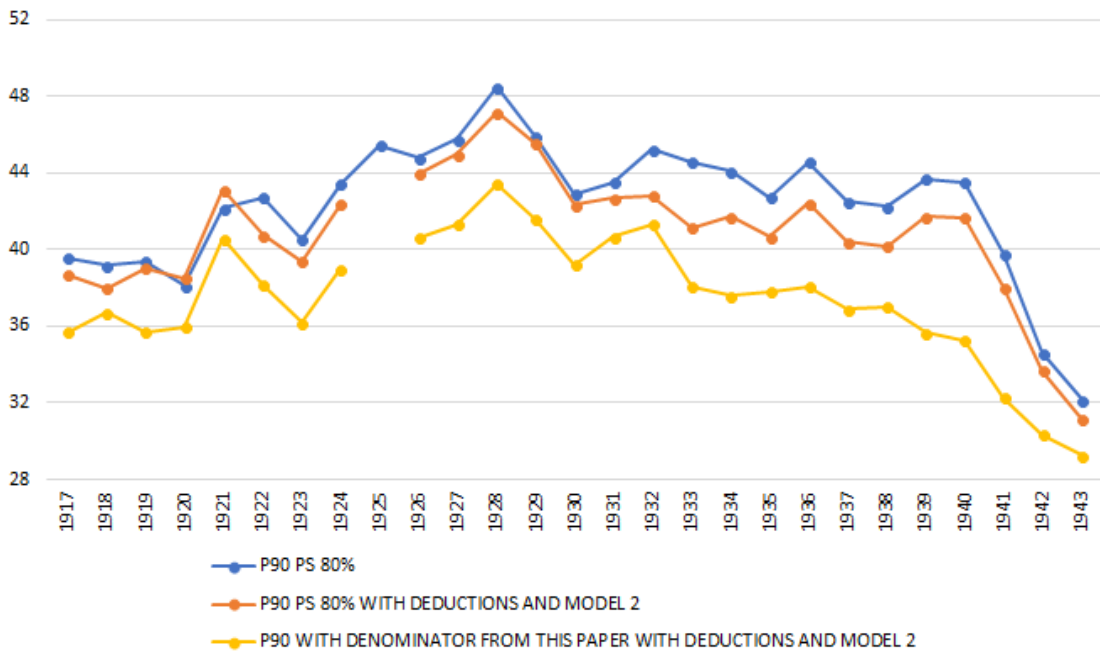


Figure 10b: Comparisons of PS Model 3 with cumulative effect of deductions, married/unmarried and denominator adjustment (P90)



necessary to remove their income from national accounts up until 1939 when the Act is passed (1953, 578). If the income from state and local government employees is removed directly from the denominator (as per Kuznets and ourselves) in order to create a consistency in income definitions, the same step should be extended to ensure consistency in the population definition used to estimate the associated number of tax units.

This is not a negligible adjustment. Take 1917 as an example. In that year, state and local government employees were equal to 3.48% of all tax units as estimated by PS (Kuznets 1953, 578-579). This is enough to increase the average income per definition-consistent tax unit by 3.6% and reduce the income of the top decile in PS model 1 by 0.35 percentage points for that year. This is a problem that only grows in amplitude over time. The number of affected state and local employees grew from 3.48% of all tax units in 1917 to 4.8% in 1929 and 5.1% in 1938.³⁷ As such, an increasingly larger share of the national accounts population baseline is inconsistently aligned relative to the tax data population. Tax units in affected years must therefore be adjusted downward to facilitate comparison on common terms.

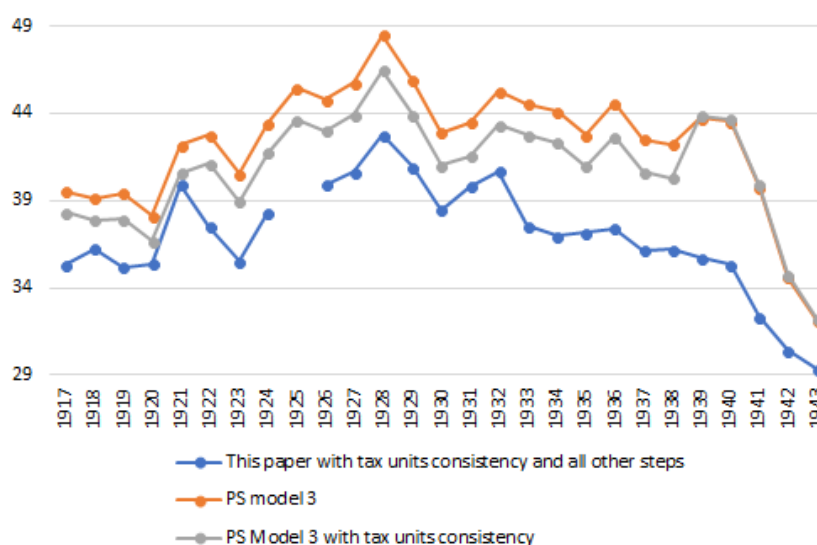
There is another definition problem in tax units that must be addressed to reconcile the two population counts. In their data files PS state that Hawaii and Alaska are not included in their tax units numbers prior to 1950. For the same years, Alaska and Hawaii are included in the total income definition taken from SOI.³⁸ We accordingly adjust the number of tax units upward to reflect population shares that are consistent with our tax-reconciled denominator.

As can be seen from figure 11, the cumulative effect of these two consistency adjustments to population, and thus tax units, removes an average of 1.39 percentage points from the

³⁷To estimate the number of tax units associated with tax-exempt class, we took annual figures of state and local government employees from NIPA and predecessor accounts and adjusted them slightly downward to correspond to the 1937 Treasury Department study of tax-exempt persons. We then used this figure to estimate associated tax units as per the PS method, and subtracted them from total tax units for the corresponding year.

³⁸In their file "Pophouseholds.xls", they explicitly state the "1900-1949 estimates exclude Alaska and Hawaii" without noting that their tax data include the income reported in those territories.

Figure 11: Effect of all adjustments and consistency in definition of tax units



income share of the top income decile (see the difference between the two lines that concern PS Model 3) for the period from 1917 to 1943. When combined with the previous adjustments to the inequality figures, our series for the top income decile is on average 6.68 percentage below that of PS model 3.

6 Extending the Smiley Adjustments to 1941

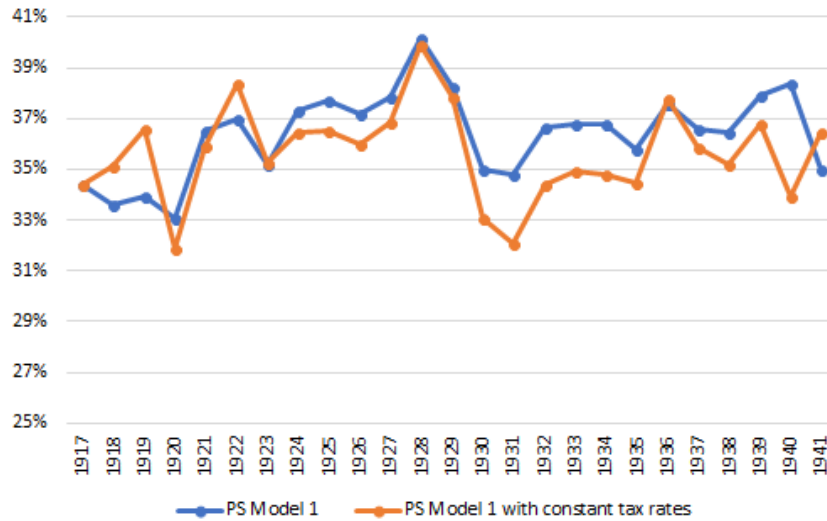
Before presenting these results in cumulative though, we discuss an additional pertinent complication that affects the underlying quality of pre-war tax statistics. As a further avenue of inquiry, independent of the previous section’s adjustments, we call attention to a recurring problem of tax collection and reporting in the pre-World War II era. Given the sizable and frequent changes in the tax code in this period as well as the self-reported nature of income filings, we may reasonably expect filers to have behaved strategically in attempting to shelter or mitigate their tax burdens. On top of lax enforcement issues and tax resistance (Beito,

1989), the dramatic changes in federal tax rates over time likely created artificial changes in measured inequality. In his work, Smiley (1983; 1995; 1998; 2000) argued that the wide swings in tax rates (from 77% in 1918 to 25% in 1925 to 81% in 1941) would have prompted tax filers to alter their reporting behavior. In years with lower tax rates, tax-eligible households would exit tax-exempt shelters and provide a more accurate representation of their full income. As such, a change in reporting behavior would overstate the change in inequality. The opposite pattern might be expected in years of high tax rates, where affected filers seek out various tax shelters to reduce their overall liability. The result would display different income earnings on paper than high income filers actually realized. These patterns effectively obscure the sources and total amount of income that may be discerned from tax records, altering observed distributional patterns.

Concerned primarily with the period from 1915 to 1929, Smiley proposed to correct for changes in reporting by running a regression that would allow him to re-estimate the number of anticipated tax returns for each tax bracket by holding tax rates constant.³⁹ Smiley was also interested in measuring the level of inequality of reported and unreported income. While this method did remove artificial variations in top income shares caused by changes in reporting during the period studied by Smiley, it is not directly comparable in nature to the PS series who presented income inequality through reported income (rather than Smiley's addition of reported and unreported income). Indeed, to properly estimate the evolution of the PS series, the appropriate question is not one of identifying the overall level of inequality, but rather one of determining the long-term patterns of inequality as measured by tax data, absent the variations in tax rates. In other words, we need to ascertain how much of the movements observed in the PS series (model 1) over time are the result of changes in reporting behavior in response to tax rate changes. We therefore modify the Smiley method in order to reflect

³⁹The details of the regression can be found in the appendix to Smiley's paper. With the permission of Smiley, who generously shared his data and appendix, we can share that appendix upon request.

Figure 12: Effect of keeping tax rates constant (PS model 1 results) for top 10% income share



this necessity over the entire pre-war period. First of all, we extended our examination from 1917 to 1941 (rather than up to 1929) and estimated elasticity coefficients for reporting in each income category.⁴⁰ These coefficients gave us corrected numbers of tax reports as if 1917 tax rates had applied throughout the period. By applying the ratio of the corrected to uncorrected numbers of returns, we also adjusted the total income of each category. Then, we assumed that the level in PS Model 1 for 1917 was valid and attributed to that point the movements of our corrected inequality series, which can be seen in figure 12. As one can see, the top income share distribution is lower in most years than that displayed by the PS models. This suggests that the levels of inequality showcased in the work of PS are very sensitive to the artificial variations in income induced by changes in tax regimes. Indirectly, this means that they are more or less capturing a measure of tax incidence.

⁴⁰More than 30 result tables were generated for this purpose. We do not include them in this article but we are willing to share them on request.

7 Discussion

The sum of our corrections should not be seen as a rebuttal of the idea that there was a fall and then rise of inequality during the 20th century. Our adjusted series for the top 10%, 5%, and 1% income shares confirm the downward trajectory of income concentration in the first half of the century, particularly following the events of 1929. However, we do show that the overall level of inequality before World War II and the magnitude of its decline are likely overstated in the original 2003 PS series. Our adjustments further reveal an imprecision in the shape of the PS series for these years. Whereas PS depict a relatively sustained high level of inequality between 1913 and 1940, we find an acute peak in 1928-29 followed by a gradual levelling pattern that extended across the Depression decade.

Evaluated historically, our findings bring distributional measures of inequality into closer alignment with the economic events of the period preceding World War II. The U-curve depicted in the original PS series exhibits a sharp and sudden decline in income concentration among all top fractiles between 1940 and 1944, completely resituating the trajectory of distributional estimates for the next three decades in a mid-century trough. Using the PS model 3 series, the top 10% is depicted as dropping an astounding 12.48 percentage points in only four years. The top 5% sheds 9.23 percentage points, and the top 1% sheds 5.05 percentage points in the same brief period. While the direction of this downward movement is consistent with other levelling effects observed worldwide during World War II, its concentration over such a short time span and in a country largely removed from the physical destruction of fighting stands out as an unusual feature of the PS series for the United States. Indeed, as our corrections seem to indicate, this concentrated focus in the original PS series may have more to do with three major coinciding discontinuities affecting the source data used to calculate the top distributional fractiles: the wartime tax base expansion (1940-45), the accounting shift from net income to AGI (1943-44), and the change in calculation method used to derive the fiscal

income denominator from national accounts (1943-47). Our refinements to the underlying source data to more accurately account for these changes dramatically lessen the depicted wartime shift in the original series.

Several other indicators suggest the severity of the wartime decline is overstated in the PS series. While the wartime economic recovery certainly expanded income earnings at the bottom of the distribution, the upper ranks of the income ladder grew aggressively between 1940 and 1944, and for several years thereafter. The number of filers in the \$100,000 tax bracket and higher more than doubled from 3,452 to 8,210 in the same years. Those in the \$50,000-100,000 brackets grew from 10,673 to 28,963, while those making between \$10,000-50,000 swelled in number from 226,555 to 595,966. While postwar inflation may explain some of this growth, these numbers stand in sharp contrast with the immediate previous levelling effects of the Great Depression, when annual tax filers in the same brackets all sharply declined between 1929 and 1933. The income growth of the 1940s also occurred in a period when the IRS dramatically revised the tax base downward and intensified enforcement, albeit while leaving intact and then slightly increasing the highly progressive top marginal rates for the highest income filers.

We find it likely from these patterns that the PS series exaggerates the U-curve's inequality downswing by both overstating its magnitude and condensing the locus of its time span to 1940-44. One possible reason is the aforementioned effect of the wartime tax code revisions upon tax reporting and accounting, creating a statistical illusion of a rapid and condensed top income share decline between 1940-44 that had actually been playing out at a more gradual pace over the preceding decade. A more plausible levelling pattern emerges from our corrections to the PS series, which instead show a prolonged decline through the Depression years and continuing into the late 1940s when our adjustments may be linked with the AGI-derived portion of the original curve. For example, PS model 3 depicts a top 10% share

that shed only a small amount after the stock market crash and then stabilized at over 43% through the Depression years until its decline in 1941. Our revisions show a continuous a Depression-era decline extended over a decade, with a much smaller wartime drop that also immediately rebounds and stabilizes after 1945.

The war remains an important part of this story, including the yearly distributional swings it produced and - importantly - its downward expansion of the federal income tax base onto lower income households that were previously exempt from filing. But the larger historical pattern that emerges from our revisions shifts the focus to 1929. That peak also takes the form of an acute event roughly coinciding with the stock market boom and bust. While top income shares in the early 20th century were generally higher than the post-war era, the revised estimates from the 1920s and 30s evince a short-term spike in inequality, followed by a steady decline over the next two decades - events that are less evident in the original PS series. On the whole, the story depicted in our corrected series is more consistent with the long-positing levelling effects of the Great Depression and less consistent with a causal attribution to conscious fiscal policy design. Significantly, the downward slide in top income shares from their late 1920s peak precedes the entrenchment of highly progressive top marginal tax rates by several years, and the World War II era expansion of the federal income tax base to its modern levels by over a decade. The extremely steep marginal tax rate structure of the mid-20th century may accordingly play a lesser role in the levelling story than suggested by the original PS study, where the progressivity of the tax system serves as a posited primary causal mechanism.

A lowering of the observed level of inequality is not a novelty per se. Many other scholars have questioned the magnitude of the increase in both income and wealth inequality, albeit - notably - without denying its existence, in the second half of the 20th century (Burkhauser et al., 2012a,b; Armour et al., 2014; Bricker et al., 2016; Mechling et al., 2017; Auten et al.,

2016; Auten and Splinter, 2017). The significance of our study is that we extend data-improving adjustments to the early part of the century, resulting in similar findings that show the original PS estimates for this period consistently overstate top income shares. As such, when we link our findings with other existing series that correct PS after 1960, the depicted pattern resembles a century long tea-saucer shape rather than a pronounced U-curve (see figure 13, table 1 for the effects for selected years and table 2 for a full listing of our cumulative corrections by year).⁴¹ This flatter evolution of income inequality confirms the "great compression" thesis whereby the 1930s through the 1970s saw a reduction in inequality (Goldin and Margo, 1992; Katz and Murphy, 1992). However, it is also more consistent with narratives such as those of Scheidel (2017) who argue that inequality may be less responsive to intentional policy designs than to external levelling events and disasters. It is important to note that our corrected series brings the series closer in line with Kuznets' (1953) earlier measure of income concentration for the top 1% in these years, which suggests that his initial work may have been closer to reality than PS realized (and probably Kuznets himself, given greater limitations on his underlying source data).

A revised depiction of the century-long series for the top 10% income share appears in our figure 14. Here we incorporate the full effects of the 1941 filer ratio and pre-1944 deductions adjustments, as well as a tax-comparable denominator arrived at through the reconciliation method inclusive of associated tax unit corrections. In cumulative, our improved estimate of the top 10% income share is consistently 3.7 to 8.2 percentage points lower than PS during the Depression era, as well as 2.3 to 5.7 points lower in the 1920s. Outside of the brief inequality

⁴¹We must remind readers that we are only dealing with the PS 2003 paper that covered the years 1913 to 1998. Since then, PS have released an updated series that extends their distributional estimates to the present. To keep the discussion simple and focused upon the construction of an accurate estimate for the pre-World War II period, we intentionally limit our discussion to their initial article, which explains why our figures end in 1998. Piketty-Saez and the Auten-Splinter adjustments to their series currently extend these numbers through 2015 in subsequently released data. We include both the original and updated PS figures in our accompanying data appendix file for comparison with our revisions.

Figure 13: The U-Curve of Top 1% Income Shares with Correction, PS and Other Series

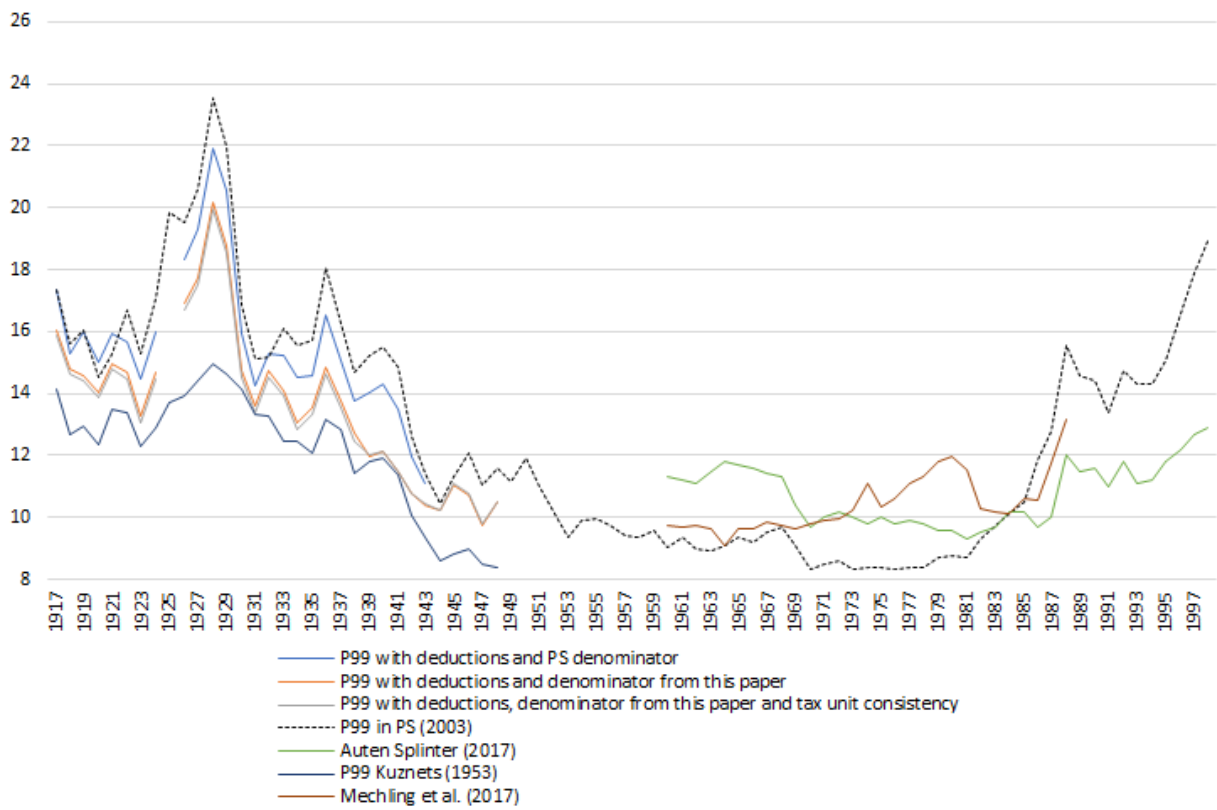


Table 1: Effect of Different Data Adjustments on Original PS Estimates (Selected Years)

Top 10% Income Shares					
	1917	1929	1936	1939	1943
Piketty-Saez Model 3	39.6	45.9	44.6	43.7	32.1
Missing Filer + Deductions					
This Paper	-0.9	-0.3	-2.2	-2.0	-1.0
Missing Filer + Deductions + Denominator Effects					
This Paper	-3.9	-4.3	-6.5	-8.0	-2.9
Missing Filer + Deductions + Tax Unit Consistency + Denominator Effects					
This Paper	-4.2	-4.9	-7.2	-8.0	-2.8
Top 1% Income Shares					
	1917	1929	1936	1939	1943
Piketty-Saez Model 3	17.3	22.0	18.0	15.2	11.4
Deductions					
This Paper	0.0	-1.4	-1.5	-1.2	-0.3
Deductions + Denominator Effects					
This Paper	-1.3	-3.2	-3.2	-3.2	-1.0
Deductions + Denominator Effects + Tax Unit Consistency					
This Paper	-1.4	-3.4	-3.4	-3.2	-1.0

Table 2: New top income shares (all corrections), 1917–1948

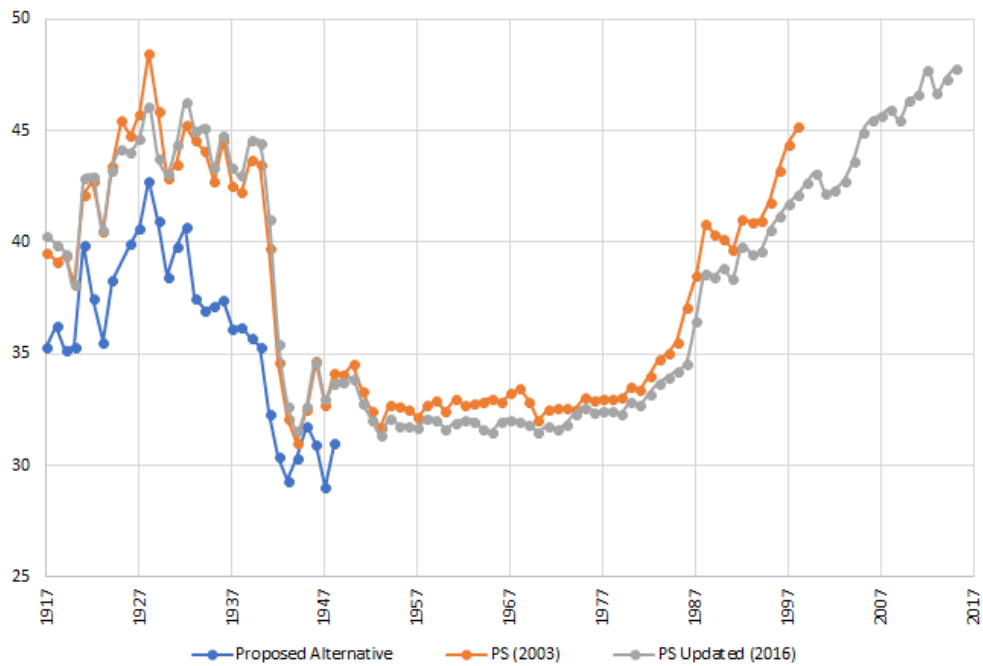
Year	Top 10%	Top 5%	Top 1%	Year	Top 10%	Top 5%	Top 1%
1917	35.3	27.3	15.9	1933	37.5	27.9	13.9
1918	36.3	26.8	14.6	1934	37.0	27.5	12.9
1919	35.2	26.5	14.4	1935	37.2	27.0	13.3
1920	35.4	26.0	13.8	1936	37.4	27.7	14.6
1921	39.9	28.6	14.8	1937	36.2	26.2	13.5
1922	37.5	27.3	14.5	1938	36.2	25.6	12.5
1923	35.5	25.4	13.1	1939	35.7	25.1	12.0
1924	38.3	27.4	14.5	1940	35.3	24.5	12.1
1925	-	-	-	1941	32.3	22.9	11.5
1926	39.9	30.1	16.7	1942	30.4	21.4	10.8
1927	40.7	30.9	17.5	1943	29.3	20.8	10.4
1928	42.8	33.3	20.0	1944	30.3	21.0	10.2
1929	41.0	31.6	18.5	1945	31.8	22.5	11.1
1930	38.5	28.0	14.5	1946	30.9	22.2	10.8
1931	39.8	27.7	13.4	1947	29.0	20.6	9.8
1932	40.7	29.3	14.5	1948	31.0	21.9	10.5

peak in the late 1920s, the top 10% income share only seldom drifted north of 40%. The depicted levelling from that peak occurs at a subtler pace over the next two decades before settling at a stable income share of roughly 33% during the mid-century trough.⁴² As such, the century long pattern takes on the characteristics of a shallow tea saucer shape rather than the pronounced U-curve.

We should point out that we believe the pre-World War II portion of the series may still be distorted, even after accounting for our corrections to the PS estimates. With the exception of

⁴²It warrants further mention that the residual drop in top income shares between 1940 and 1944 (followed by an almost immediate rebound and stabilizing between 1945-50) may actually reflect the administrative overhaul of tax collection procedures in this period, as opposed to a meaningful decline in inequality. These changes included the partial implementation of automatic income withholding in 1943 as part of a supplemental wartime Victory Tax. The initial withholding procedure applied only to wage incomes from monthly payroll, which tended to concentrate in lower income brackets. Other revenue streams continued to self-report annually until the full implementation of income withholding in 1944. As a result, wartime distributional estimates may reflect additional unaccounted distortions from tax reporting changes that induced irregular patterns of tax evasion and avoidance between different brackets and income streams for these years. These distortions may accordingly limit the accuracy of reported income for these years by producing the illusion of a temporary acute drop in top income shares. Scholars of long-term distributional trends should accordingly use estimates from 1942 to 1944 with extreme caution.

Figure 14: The U-Curve of Top 10% Income Shares with our proposed alternative compared with initial PS estimates



the point we made regarding artificial income reporting variations induced by changes in tax rates, we did not examine larger underlying quality issues with pre-World War II tax data. However, as Auten et al. (2016); Auten and Splinter (2017); Mechling et al. (2017) point out for later years, tax-measured inequality is bound to capture tax incidence much more than it is to capture actual inequality. Suggestive of this problem, a comparison to the state income tax system’s records in Wisconsin reveals a continuous IRS under-count of reported income and filers in the state for most years prior to 1941, particularly in the lowest tax brackets (Geloso and Magness, 2017). If this pattern occurred in other states as well, IRS data may suffer from issues of persistent under-reporting that affects the entire first half of the U-curve beyond what may be accounted for in the missing filer adjustment of PS model 2 and our corrections to it.

These and other avenues of research remain largely unexplored features of the pre-World War II federal income tax system that nonetheless weigh heavily upon the accuracy of tax

filings recorded in SOI. Our adjustments, which address issues specific to data that were captured by the SOI, should therefore be construed as a starting point for further refinements across the entire prewar period. Given the importance that the inequality debate has taken in recent years, improving these data represents a necessary margin to better understand long term dynamics in the distribution of income.

8 Conclusion

This paper presents a series of corrections in order to improve upon the pioneering long-term study of income distributions in the United States produced by Thomas Piketty and Emmanuel Saez. While their work constitutes a significant contribution to the understanding of the history of American economic inequality before World War II, a topic that had somewhat laid dormant since Simon Kuznets in the 1950s, the PS estimates for the first half of the 20th century warrant substantial revisions.⁴³ Even as PS captured elements of the levelling pattern in this period, their calculations introduced a non-negligible overstatement of income concentrations on the left side of the U-curve as well as an imprecise timeline for its depicted changes. Our findings therefore complement new work by other scholars (Meyer and Sullivan, 2013; Armour et al., 2014; Kopczuk, 2015; Bricker et al., 2016; Auten et al., 2016; Auten and Splinter, 2017; Mechling et al., 2017) who propose similar adjustments, albeit on the other end of the curve. When combined with the direction of revisions in these recent works, our own corrected estimates for the first half of the curve reveal a much shallower pattern of shifting income concentration over the 20th century in the United States. This finding, in turn, calls into question many longstanding observations about a posited causal link between

⁴³We do not infer that no research took place. There were many important articles published in the interlude (see notably Schmitz and Fishback [1983] and Lindert and Williamson [1985]). However, the attention granted to the topic has increased considerably thanks to Piketty and Saez in no small part because of the roadmap they laid to calculate top income shares.

the previously depicted U-shaped pattern and the highly progressive tax policies of the mid-century.

Our corrections are based only upon verifying the underlying assumptions behind the PS constructions. No questions were asked regarding further data quality issues beyond these initial considerations - we leave this for future research. There are obviously many points that future research should follow, notably to better control for the historical changes in tax regimes, their effects on the reliability and accuracy of self-reported income, and further adjustments to address the size of households.⁴⁴ At minimum, other derivative estimates of income composition and top wealth shares that rely upon the original PS series warrant similar revision to increase the accuracy of their findings for the early 20th century. While we believe these possibilities show promise for further improvements, we concentrate our efforts on correcting and improving foundational issues with the construction of the original inequality series.

Our results should not be taken as the final word but rather the start of a long discussion about the left-side of the American inequality U-curve. They do nevertheless provide a basis for reinterpreting the evolution of income inequality in the United States. Echoing Sutch (2017), we note that altering the shape of the distributional series also changes the interpretation of what led to the mid-century reductions in top shares and what remedies, if any, we might employ against rising inequality in the future. Difficulties of measurement are a recurring challenge of historical inequality studies that we must remain attentive to, particularly when their implications extend to depicted long-term patterns that have approached a point of stylization in the economic literature.

⁴⁴On this point see notably Auten and Splinter (2017)

References

- Armour, P., Burkhauser, R. V., and Larrimore, J. (2014). Levels and Trends in US Income and its Distribution: A Crosswalk from Market Income towards a Comprehensive Haig-Simons Income Approach. *Southern Economic Journal*, 81(2):271–293.
- Atkinson, A. B., Piketty, T., and Saez, E. (2011). Top incomes in the long run of history. *Journal of Economic Literature*, 49(1):3–71.
- Auten, G. and Splinter, D. (2017). Income Inequality in the United States: Using Tax Data to Measure Long-term Trends.
- Auten, G., Splinter, D., and Nelson, S. (2016). Reactions of high-income taxpayers to major tax legislation. *National Tax Journal*, 69(4):935.
- Autor, D. H., Katz, L. F., and Kearney, M. S. (2008). Trends in US wage inequality: Revising the revisionists. *Review of Economics and Statistics*, 90(2):300–323.
- Beito, D. T. (1989). *Taxpayers in revolt: Tax resistance during the Great Depression*. University of North Carolina Press.
- Bricker, J., Henriques, A., Krimmel, J., and Sabelhaus, J. (2016). Measuring Top Wealth Shares Estimating Top Income and Wealth Shares: Sensitivity to Data and Methods. *American Economic Review*, 106(5):641–645.
- Burkhauser, R. V., Feng, S., Jenkins, S. P., and Larrimore, J. (2012a). Recent trends in top income shares in the United States: Reconciling estimates from March CPS and IRS tax return data. *Review of Economics and Statistics*, 94(2):371–388.
- Burkhauser, R. V., Larrimore, J., and Simon, K. I. (2012b). A "second opinion" on the economic health of the American middle class. *National Tax Journal*, 65(1):7–35.
- Burtless, G. and Svaton, P. (2010). Health care, health insurance, and the distribution of American incomes. *Forum for Health Economics & Policy*, 13(1):1–39.
- Congress (1939). *Tax-Exempt Salaries Hearings before the Committee on Ways and Means: House of Representatives, Seventy-Sixth Congress First Session*. Government Printing Office.
- Farioletti, M. (1958). Some income adjustment results from the 1949 audit control program. In *An Appraisal of the 1950 Census Income Data*, pages 241–286. Princeton University Press.
- Federal Reserve at Saint-Louis (2018). Personal income. Data retrieved from Federal Reserve Economic Data, <https://fred.stlouisfed.org>.
- Feenberg, D. R. and Poterba, J. M. (2000). The income and tax share of very high-income households, 1960-1995. *American Economic Review*, 90(2):264–270.

- Fixler, D. and Johnson, D. (2014). Accounting for the distribution of income in the u.s. national accounts. In Jorgenson, D., Landefeld, S., and Schreyer, P., editors, *Measuring Economic Sustainability and Progress*, pages 213–244. NBER and University of Chicago Press.
- Frank, M. W. (2009). Inequality and growth in the United States: Evidence from a new state-level panel of income inequality measures. *Economic Inquiry*, 47(1):55–68.
- Frank, M. W. (2014). A new state-level panel of annual inequality measures over the period 1916-2005. *Journal of Business Strategies*, 31(1):241.
- Geloso, V. and Magness, P. W. (2017). Trends and Levels of Income Inequality in Wisconsin, 1919-1941: The Effect of Using Different Data Sources.
- Goldin, C. (1990). *The Gender Gap: An economic History of American Women*. New York: Cambridge University Press.
- Goldin, C. and Margo, R. A. (1992). The great compression: The wage structure in the United States at mid-century. *Quarterly Journal of Economics*, 107(1):1–34.
- Goldsmith, S. (1951). Appraisal of basic data available for constructing income size distributions. In *Studies in Income and Wealth*, pages 266–377. NBER.
- Goldsmith, S., Jaszi, G., Kaitz, H., and Liebenberg, M. (1954). Size distribution of income since the mid-thirties. *Review of Economics and Statistics*, pages 1–32.
- Goldsmith, S. F. (1957). Changes in the size distribution of income. *American Economic Review*, pages 504–518.
- Gould, B. R. and Horowitz, S. A. (2011). History of the combat zone tax exclusion. Technical report, Institute for Defense Analyses.
- Gruber, J. and Hungerman, D. M. (2007). Faith-based charity and crowd-out during the great depression. *Journal of Public Economics*, 91(5):1043–1069.
- Higgs, R. (1977). *Competition and coercion: Blacks in the American economy 1865-1914*. Cambridge University Press.
- Ige, T. H. (1955). Working conditions and workers’ wages. *Monthly Lab. Rev.*, 78(12):1422–1426.
- Joint Economic Committee (1965). *The distribution of personal income: a study of statistics on the size distribution of personal income in the United States*. US Government Printing Office.
- Katz, L. F. and Murphy, K. M. (1992). Changes in relative wages, 1963–1987: supply and demand factors. *Quarterly Journal of Economics*, 107(1):35–78.

- Kopczuk, W. (2015). What do we know about the evolution of top wealth shares in the United States? *Journal of Economic Perspectives*, 29(1):47–66.
- Kopczuk, W., Saez, E., and Song, J. (2010). Earnings inequality and mobility in the United States: evidence from social security data since 1937. *Quarterly Journal of Economics*, 125(1):91–128.
- Kuznets, S. (1941). *National Income and Its Composition, 1919-1938*. National Bureau of Economic Research.
- Kuznets, S. (1945). *National Product in Wartime*. National Bureau of Economic Research.
- Kuznets, S. (1953). *Shares of Upper Income Groups in Income and Savings*. National Bureau of Economic Research.
- Lindert, P. and Williamson, J. (1980). *American inequality: A macroeconomic history*. Academic Press.
- Lindert, P. H. and Williamson, J. G. (1985). Growth, equality, and history. *Explorations in Economic History*, 22(4):341–377.
- Lindert, P. H. and Williamson, J. G. (2016). *Unequal Gains: American Growth and Inequality since 1700*. Princeton University Press.
- Margo, R. A. (2016). Obama, Katrina, and the Persistence of Racial Inequality. *Journal of Economic History*, 76(02):301–341.
- Mechling, G., Miller, S., and Konecny, R. (2017). Do Piketty and Saez Misstate Income Inequality? Critiquing the Critiques. *Review of Political Economy*, 29(1):30–46.
- Meyer, B. D. and Sullivan, J. X. (2013). Consumption and Income Inequality and the Great Recession. *American Economic Review*, 103(3):178–183.
- Mitchener, K. J. and McLean, I. W. (1999). US regional growth and convergence, 1880–1980. *Journal of Economic History*, 59(04):1016–1042.
- Pechman, J. A. (1956). Yield of the individual income tax during a recession. In *Policies to Combat Depression*, pages 123–148. NBER.
- Piketty, T. (2001). *Les hauts revenus en France au 20e siècle : inégalités et redistribution, 1901-1998*. Éditions Grasset.
- Piketty, T. (2014). *Capital in the twenty-first century*. Belknap Press of Harvard University Press Cambridge, MA.
- Piketty, T. and Saez, E. (2003). Income Inequality in the United States, 1913-1998. *Quarterly Journal of Economics*, 118(1):1–39.

- Piketty, T. and Saez, E. (2007). Income and Wage Inequality in the United States, 1913-2002. In *Top Incomes Over the Twentieth Century*, pages 141–226. Oxford University Press.
- Piketty, T., Saez, E., and Zucman, G. (2017). Distributional national accounts: methods and estimates for the united states. *Quarterly Journal of Economics*, 133(2):553–609.
- Reynolds, A. (2006). *Income and wealth*. Greenwood Publishing Group.
- Reynolds, A. (2012). The misuse of top 1 percent income shares as a measure of inequality.
- Saez, E. and Zucman, G. (2016). Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *Quarterly Journal of Economics*, 131(2):519–578.
- Scheidel, W. (2017). *The great leveler: Violence and the history of inequality from the stone age to the twenty-first century*. Princeton University Press.
- Schmitz, M. and Fishback, P. V. (1983). The distribution of the income in the great depression: Preliminary state estimates. *Journal of Economic History*, 43(1):217–230.
- Smiley, G. (1983). Did Incomes for Most of the Population Fall from 1923 through 1929? *Journal of Economic History*, 43(01):209–216.
- Smiley, G. (1998). New estimates of income shares during the 1920s. In Haynes, J. E., editor, *Calvin Coolidge and the Coolidge Era: Essays on the History of the 1920s*, pages 215–232. Library of Congress.
- Smiley, G. (2000). A note on new estimates of the distribution of income in the 1920s. *Journal of Economic History*, 60(04):1120–1128.
- Smiley, G. and Keehn, R. H. (1995). Federal Personal Income Tax Policy in the 1920s. *Journal of Economic History*, 55(02):285–303.
- Solt, F. (2016). The standardized world income inequality database. *Social Science Quarterly*, 97(5):1267–1281.
- Sommeiller, E. (2007). *Regional income inequality in the United States, 1913–2003*. PhD thesis, University of Delaware.
- Sutch, R. (2017). The One Percent across Two Centuries: A Replication of Thomas Piketty’s Data on the Concentration of Wealth in the United States. *Social Science History*, 41(4):587–613.
- Troiano, U. (2017). Do Taxes Increase Economic Inequality? A Comparative Study Based on the State Personal Income Tax. Technical report, National Bureau of Economic Research.
- United States House of Representatives (1916). *Hearing before the Committee on Expenditures in the Treasury Department on H. Res 221, pertaining to the collecting of the income tax*. Government Printing Office.