Top Income Inequality in the 21st Century: Some Cautionary Notes

Fatih Guvenen[*]   Greg Kaplan[†]

April 2, 2017

Summary

We revisit recent empirical evidence about the rise in top income inequality in the United States, drawing attention to four key issues that we believe are critical for an informed discussion about changing inequality since 1980. Our goal is to inform researchers, policy makers, and journalists who are interested in top income inequality. Our analysis is based on a reexamination of publicly available detailed statistics from two administrative data sources: (i) Internal Revenue Service (IRS) data on *total* incomes (labor income plus capital income), reported in Saez (2012), and (ii) individual-level micro data on *labor* income (wage plus self-employment income) from the U.S. Social Security Administration (SSA), reported in Guvenen et al. (2014). Reexamining these statistics reveals four findings:

In the final two decades of the 20th century (1981–2000):

1. The rise in top income inequality revealed in SSA and IRS data track each other very closely, except for the two-year period between 1986 and 1988 (see point 2). This is despite non-trivial differences between the two data sources in the definition of income (labor vs total) and the unit of analysis (individual vs. tax unit).

2. Between 1986 and 1988, IRS data show a large jump in top income shares, which is not evident in SSA data. This jump has been previously noted by researchers and

---

[*]University of Minnesota, FRB of Minneapolis, and NBER; `guvenen@umn.edu`
[†]University of Chicago and NBER; `gkaplan@uchicago.edu`

is likely a consequence of income shifting from the corporate sectors to pass-through entities in the wake of the Tax Reform Act of 1986.

So far in the 21st century (2001–2012):

3. IRS and SSA data reveal diverging patterns in top income shares – the IRS data show a steady increase, whereas the SSA data show no increase at all. The difference is due to the increasing importance of income accruing to pass-through entities (partnerships and S-corporations), which is included in the IRS measure of total income but not in either the IRS or SSA measure of labor income.

4. Moreover, the bulk of this growth in income from pass-through entities was concentrated at the very top of the distribution – above the 99.99th percentile, a group that contains only about 12,000 households. The share of incomes above the 99th percentile (around $372,000 in 2012) but below the 99.99th percentile (around $7.2 million in 2012) has barely changed in the last two decades.

1   Background

Top income inequality in the United States, as measured by the share of incomes accruing to individuals in the top percentiles of the income distribution, has risen dramatically in the last half century. Figure 1, which is based on data from Saez (2012) is a typical example of the evidence that is used to illustrate this trend. The figure shows that after bottoming out in the 1960s, the top 1 percent and top 0.1 percent shares of total income (excluding capital gains) have been rising steadily for the last 40 years.[1] This trend has motivated research on top income inequality to a phenomenal extent; a vast number of recent papers, too numerous to list, study its policy implications, theoretical underpinnings and statistical deconstruction. That top income inequality is increasing in the United States has become almost gospel in academic and policy debate on the topic of income distribution.

Our goal in this note is to highlight some features of the recent evidence on top incomes that researchers, journalists, and policy analysts should bear in mind as the discussion of this issue evolves. We do not deny that dramatic changes have taken place in the distribution of incomes; the evidence in Figure 1 is indisputable. Instead, in typical economist fashion, we want to emphasize some caveats that we think are particularly important for guiding

---

[1]Throughout the paper, we exclude capital gains from our definition of total income.

Figure 1: Top Income Shares, 1917-2012
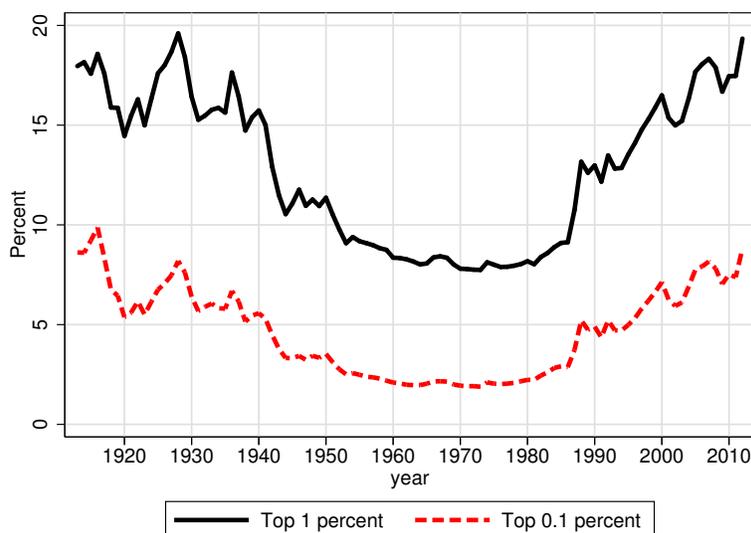
the search for useful theories of changes in top inequality, for framing further empirical investigation of top incomes, and for contextualizing the debate over the appropriate policy response. The caveats that we emphasize pertain to subtleties in both words of the phrase "top incomes".

First, the precise definition of "income" is critical for the inferences one makes about recent trends in top inequality. Much of the remainder of this note focuses on illustrating the differences in trends for alternative measures of income. We start in Section 2 by highlighting the different trends in top income shares calculated from Social Security Administration (SSA) data on wage, salary, and self-employment income versus those calculated from Internal Revenue Service (IRS) data on total incomes as in Figure 1. The SSA data show a much smaller increase in top income shares relative to the IRS data and no increase at all since 2000. We then show in Section 3 that the difference in the trends in the two series is due to the different definition of income. The increase in top income shares observed in the IRS data since 1981 is mostly a reflection of an increase in entrepreneurial income, and most of that increase is due to income accruing to flow-through entities, such as S-corporations and partnerships. This is important because the income accruing to these entities is sensitive to changes in tax policy, such as the Tax Reform Act of 1986 (TRA86) (Feenberg and Poterba (1993); Gordon and Slemrod (2002)). Indeed, the increase in top

income shares in the two years immediately following TRA86 is larger than the entire increase in top income shares over the period from 1981 to 2000. It is thus difficult to know how much of the increase in measured income of flow-through entities reflects a true increase in incomes versus income shifting between the corporate and personal sectors.

Second, the precise definition of the "top" of the income distribution is critical for the discussion of top income inequality. Researchers variously adopt the top 10 percent, 1 percent, 0.1 percent, or 0.01 percent as their preferred definition of the top of the income distribution. This choice is often determined primarily by which of the definitions produces statistics that better highlight the points they wish to make. In Section 4, we show that almost all of the post-2000 increase in the top 1 percent income share is due to an increase in the income of the top 0.01 percent income share. This means that the recent increase in top income inequality is primarily a phenomenon that is about incomes above the 99.99th percentile, rather than a phenomenon that is about incomes between the 99th and 99.99th percentiles.

To put these differences in perspective, observe that in 2012, the 90th percentile of the total income distribution was \$112,000, the 99th percentile was \$372,000, the 99.9th percentile was \$1.55 million, and the 99.99th percentile was \$7.2 million. This is a wide range of income levels even within the top 1 percent. Households with earnings at each of these income levels come from different backgrounds, work in different occupations, obtain their income from different sources, face different investment opportunities, and are subject to different types of risk. It is unlikely that a single mechanism simultaneously describes the evolution of incomes at this wide range of levels.

A simple but often underemphasized point is that changes in top income shares can be driven by changes in incomes at either the top or the bottom of the distribution. In Section 5, we undertake a decomposition to better understand which of these changes are more important for the observed trends. We find that both slower than normal income growth in the bottom percentiles, as well as faster than normal income growth in the top percentiles, contribute to the rise in top income shares over the past 30 years. The relative contributions depend on assumptions about what constitutes "normal" income growth, but with assumed normal growth rates in the range of observed average income growth over the last century, stagnation of incomes in the bottom percentiles plays an important role in explaining the growth in top income shares.

We end in Section 6 with some thoughts about the relative benefits of alternative sources of data on top incomes as well as issues facing research on top income inequality.

Figure 2: Top Income Shares, Alternative Sources

(a) Top 1 percent

(b) Top 0.1 percent
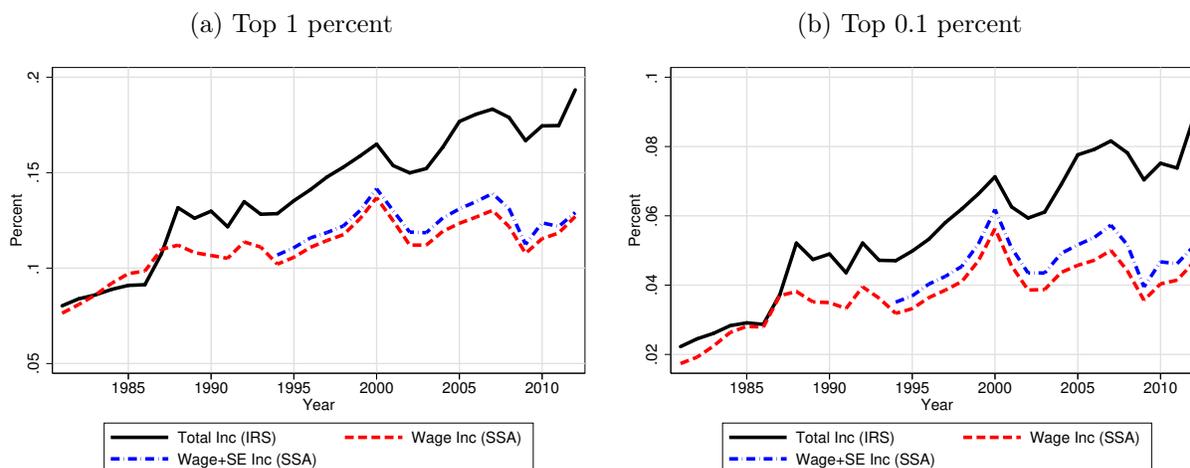
Notes: IRS data are for total income excluding capital gains from Saez (2012). SSA data are from Guvenen et al. (2014). The dashed red line is wage and salary income from W-2. The dash-dot blue line is wage and salary income from W-2 plus self-employment income from Schedule SE.

2 Two Data Sources for Top Income Shares

We will frame our discussion of these issues by comparing two empirical time series for top income shares, shown in Figure 2 (top 1 percent in the left panel, top 0.1 percent in the right panel). The solid black line is the top income share from Saez (2012), defined as the share of aggregate household income (including all types of labor and capital income, but excluding capital gains) that is earned by the top 1 or 0.1 percent of households. The series is constructed using data provided by the IRS. This is the same series as in Figure 1, displayed from 1981 onward. The dashed red line is a closely related measure of top income shares from Guvenen et al. (2014), which is defined as the share of aggregate individual income from wages and salaries that is earned by the top 1 or 0.1 percent of individuals. The series is constructed using data provided by the SSA.[2]

The series in Figure 2 are plotted starting in 1981 because the SSA data are only reliably available from 1981 onward. For this reason, we will focus our remaining analysis on the period from 1981 to the present. Although this is a much shorter time span than what

---

[2]The IRS data come from Table A1 from Saez (2012). The SSA data are based on data from a 10 percent random sample from the Master Earnings File. See main text and Guvenen et al. (2014) for more details of the data. Neither series includes non-wage (fringe) benefits that have been growing over this period.

is displayed in Figure 1, almost all of the increase in top income shares has taken place during these three most recent decades.

From 1981 to 1987 the two series show similar levels and trends for top income shares, for both the top 1 percent and the top 0.1 percent. After 1987, however, striking differences between the two series emerge. The IRS series show a distinct jump between 1987 and 1988, followed by a steady, almost linear, increase in the income shares of both top groups. But the SSA data show only a very small increase in the income shares of these groups, almost all of which took place during a brief period in the late 1990s. Since 2000, if there is any trend at all in the SSA data, it is downward.

The different trends in these two series for top income shares post-2000 have important implications, particularly with regard to appropriate policy responses. With this in mind, the remainder of this note focuses on understanding why and where these two measures of top income inequality differ in recent years. We delay our thoughts on the relative merits of the two data sources and the implications of the different trends until Section 6, at which point the likely reasons for the discrepancy will be clearer.

Two main differences between the IRS and SSA income data could account for the different recent trends in Figure 2: the unit of observation and the definition of income.

In the IRS data, the unit of observation is a tax unit, whereas in the SSA data, the unit of observation is an individual. Moreover, the two series in Figure 2 are each based on selected subsamples from the respective data sources. The SSA series in Figure 2 comes from Guvenen et al. (2014), which is based on a subsample of individuals who meet the following two criteria: (i) the individual must be between 25 and 60 years old, and (ii) the individual must have annual earnings that exceed a time-varying minimum threshold that is equal to the earnings one would obtain by working for 520 hours (13 weeks at 40 hours per week) at one-half of the federal minimum wage of that year.[3] Excluding individuals with low annual earnings is common practice in empirical analyses of earnings inequality. The IRS series is constructed by Saez (2012) by scaling the number of tax units up to the total population using U.S. Census Bureau population estimates and setting income for non-filers equal to 20 percent of average income.

A consequence of the different units of observation is that the two data sources deal with the very bottom of the income distribution in different ways. As shown in Figure 3, the trends in these two alternative units of observation differ substantially over the period 1981-2012. Trends in top income shares are determined by trends in incomes at both the

---

[3]In 2012, the minimum earnings threshold corresponded to annual earnings of $1,885.

Figure 3: Units of observation
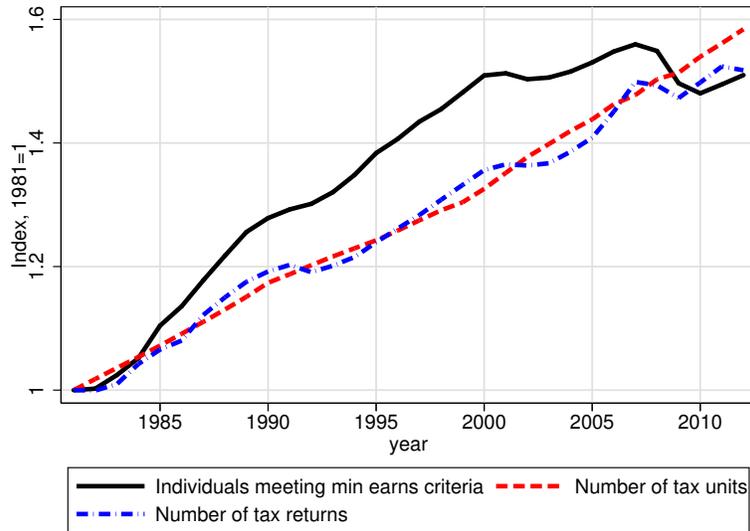
Notes: Data from Saez (2012) and Guvenen et al. (2014).

top and bottom of the distribution, so one might be concerned that the different unit of analysis drives some of the differences between the IRS and SSA series. Indeed, in the context of comparing measures of income and wealth derived from IRS data with those derived from Survey of Consumer Finances (SCF) data, Bricker et al. (2016) conduct a comprehensive investigation of the importance of using tax units versus families as the unit of analysis for top income and wealth shares. They find that these differences in the unit of analysis can explain virtually all of the difference in trends in top income and wealth shares between the SCF and IRS data. However Bakija et al. (2012) and Smith et al. (2017) compute top income shares in the IRS data with different assumptions about the bottom of the distribution to Saez (2012) and find a trend similar to the black line in Figure 2.

So although in principle, different units of analysis can affect measured trends in top income shares, we do not think this is a particularly important factor in reconciling the IRS and SSA trends in Figure 2. Rather, it is the second difference between the two series – the definition of income – that explains the different trends in top income shares. The IRS series is based on a broad measure of total income reported on tax returns, whereas the SSA series is based on labor income reported on W-2 forms. The labor income reported on W-2 forms includes wages, salaries, tips, bonuses, overtime, commissions, realized stock

Figure 4: Top Income Shares: Total Income vs Wage Income

(a) Top 1 percent   (b) Top 0.1 percent
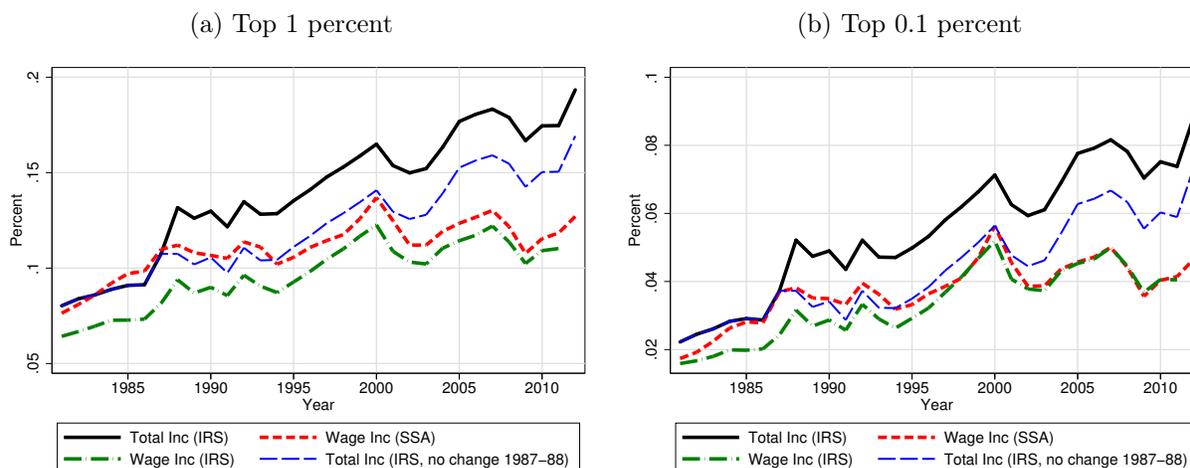
Notes: IRS data is for total income excluding capital gains from Saez (2012). SSA data is from Guvenen et al. (2014). Dashed red line is wage and salary income from W-2. Dash-dot green line is wage and salary income from IRS. Solid black line is total income excluding capital gains from IRS.

options and taxable fringe benefits. After 1994, a broader measure of labor income that includes both W-2 income as well as self-employment (sole proprietors') income reported on Schedule SE forms is also available in the SSA data. The top income shares based on this expanded measure of labor income is shown by the blue dash-dot line in Figure 2. The series closely tracks those based only on W-2 income. The main categories of income that are included in the IRS total income measure but excluded from the SSA measure of labor income are interest, dividends, rents, and all income that is reported on Schedule E (income from royalties, partnerships, S-corporations, estates, and trusts). In the following section, we investigate the importance of these differences in the definition of income.

3 Wage Income versus Total Income

A useful starting point to help clarify the role of the broader definition of income in the IRS data compared with the SSA data is to restrict the IRS definition of income to include only wage and salary income, so that it coincides with the SSA definition. Using data reported in Saez (2012), Figure 4 plots this series (green dash-dot line) alongside the corresponding series based on total income from the IRS data (black solid line) and wage and salary income from the SSA data (red dashed line), for the top 1 percent share (left panel) and

top 0.1 percent share (right panel). From the mid 1990s onward the IRS and SSA series for wage and salary income follow very similar trends. For the top 0.1 percent share, the two series lie almost on top of each other.

The data in Figure 4 are strong evidence that the absence of a trend in top income shares in the SSA data is not due to anything particular about SSA data per se, since the same conclusion is evident in the corresponding series from the IRS data. Rather, it appears that the steady increase in top income shares in the IRS data is due to sources of income other than those that are reported on W-2 forms (or Schedule SE forms, given the findings reported in Figure 2). There is, however, a small difference in the trends in the two series for top wage and salary income shares before 1995, which is likely attributable to the aforementioned differences in the way that low-income individuals are treated in the two series (see Figure 3).

Thus, the large and widening gap between the commonly reported IRS measure of top income shares and the SSA measure we report here is due to differences between the wage, salary, and self-employment income measured in the SSA data, and the broader measure of income measured in the IRS data. Given that the recent rise in top income shares is not from labor income, which components of income are driving this increase?

Four main components of income are included in the total income measure but are excluded from the wage and salaries measure: (i) entrepreneurial income, (ii) dividends, (iii) interest, and (iv) rents. Figure 5 shows the share of total income that is accounted for by each of these categories of income accruing to the top 1 percent (left panel) and top 0.1 percent (right panel). Over this period, the category that has shown the biggest increase is entrepreneurial income (solid red line). With the exception of the mid-2000s, which saw a brief rise in interest and dividend income, essentially all of the differential trend in top shares between wage and salary income and total income can be attributed to the sharp rise in the share of entrepreneurial income of top earners in total income.

So what exactly is "entrepreneurial" income? The series for entrepreneurial income comes from Saez (2012) and is defined as the sum of profits from S-corporations, profits from partnerships, profits from sole proprietorship businesses (Schedule C income), and farm income. Profits from S-corporations and partnerships constitute income that is generated by entities referred to as "pass-through" businesses. Cooper et al. (2016) undertake a comprehensive investigation in order to trace the pass-through income to its ultimate originators and conclude that at least half of the post-1980 rise in the top 1 percent income share can be accounted for by pass-through entities. Precisely because S-corporation income is the largest category and largest driver of income growth among the top percentiles, Smith et

Figure 5: Breakdown of Top Income Shares

(a) Top 1 percent

(b) Top 0.1 percent
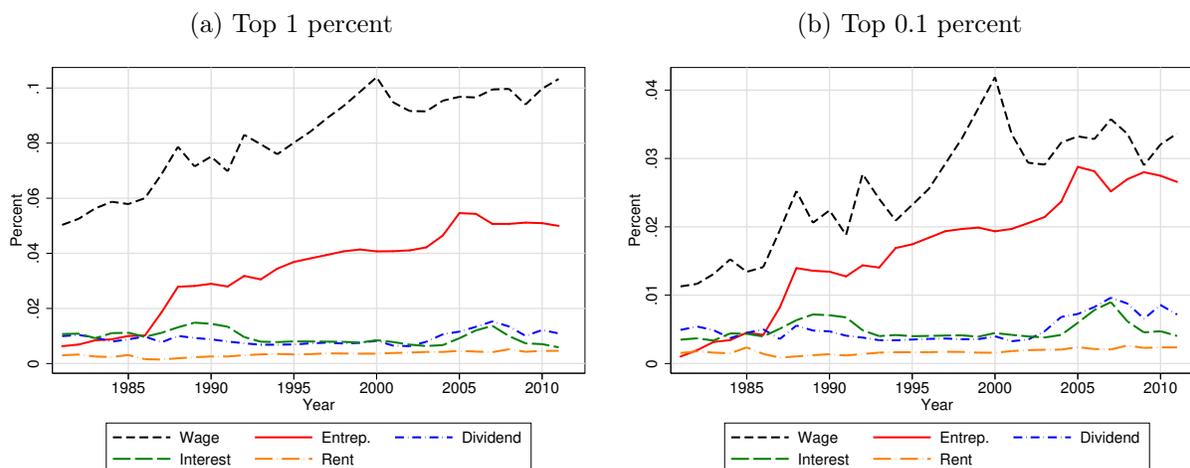
Notes: Share of total income accruing to components of income among top percentile groups

al. (2017) conduct a detailed investigation of recent trends in ownership of S-corporations. Their findings, together with the trends in Figure 4 and Figure 5, strongly suggest that pass-through entities account for almost all of the rise in entrepreneurial incomes in the top percentiles.

Moreover, it is clear from Figure 5 that the period from 1986 to 1988 is particularly important in accounting for the overall change in entrepreneurial income among the top percentiles over the last three decades. This is concerning because these are the years immediately following TRA86. It is well documented that because of the significant reduction in personal income tax rates that were part of TRA86, non-corporate pass-through entities such as S-corporations became a more attractive corporate structure than corporate entities such as C-corporations (see Auerbach and Poterba (1987), Auerbach and Slemrod (1997), and Gordon and Slemrod (2002)). For example, Auerbach and Slemrod (1997) provide a comprehensive overview of TRA86 and note the sharp increase in S-corporations in the immediate wake of the reform.

This raises the question of how much of the increase in top incomes from 1986 to 1988 reflects an actual increase in income versus how much is simply a shift of income that was previously labeled as corporate income to personal income – a concern that has been previously raised by Feenberg and Poterba (1993) and Gordon and Slemrod (2002). A back-of-the-envelope, but not unreasonable, upper bound on the effects of TRA86 on top income shares is to assume that all of the 1987-88 increase in top shares was due to income

shifting. The blue dashed line in Figure 4 shows how the IRS series for top shares based on total income would look under this assumption. With the 1987-88 jump removed, both the top 1 percent and top 0.1 percent shares from the IRS total income data (blue dashed line) lie essentially on top of the corresponding shares from the SSA wage and salaries data (red dashed line) all the way up to 2000. The difference in the 1987-88 growth in the top 1 percent income share between the IRS data and the SSA data accounts for 91 percent of the difference in the growth of the top 1 percent share over the period 1981 to 2000. For the top 0.1 percent share, the difference between the two series in the 1987-88 change is larger than the difference in the change in the top 0.1 percent share over the entire period 1981-2000. For the entire period 1981 to 2012, the 1987-88 change accounts for 35 percent and 37 percent of the difference between the two series in top the 1 and top 0.1 percent share growth, respectively.

In summary, essentially all of the increase in top income shares in IRS data over the last 30 years is due to an increase in entrepreneurial income, in particular income earned by pass-through entities such as S-corporations. For the period from 1981 to 2000, the effects of income shifting due to TRA86 likely explain the majority of the additional top share growth observed in the IRS total income data compared with the SSA wage and salary data. For the period post-2000, top income shares computed using only wage, salary, and Schedule SE income are either flat or declining, whereas top shares based on a broader measure of income show a continued increase. Almost all of this recent increase is due to an increase in the share of income accruing to S-corporations, whereas capital income such as interest and dividends have played a very minor role.

4   The Top versus the Very Top

The previous section showed that the post-2000 increase in the top 1 percent and top 0.1 percent shares of total income is mostly attributable to an increase in income earned by pass-through businesses such as S-corporations and partnerships. In this section, we show that this increase in top shares is primarily due to an increase in average incomes at the very top, above the 99.99th percentile, of the distribution and is not due to a broad-based increase in average incomes among the top 1 percent. For perspective, recall that in 2012 the 99.99th percentile of the total income distribution was $7.2 million, with about 12,000 households earning above this threshold in the entire U.S. economy.

The contribution of this very small group to rising top income shares is most easily seen by recomputing the trends by excluding these households. As seen in Figure 6, when we

Figure 6: Top Income Shares Excluding Incomes above 99.99th Percentile

(a) Top 1 percent
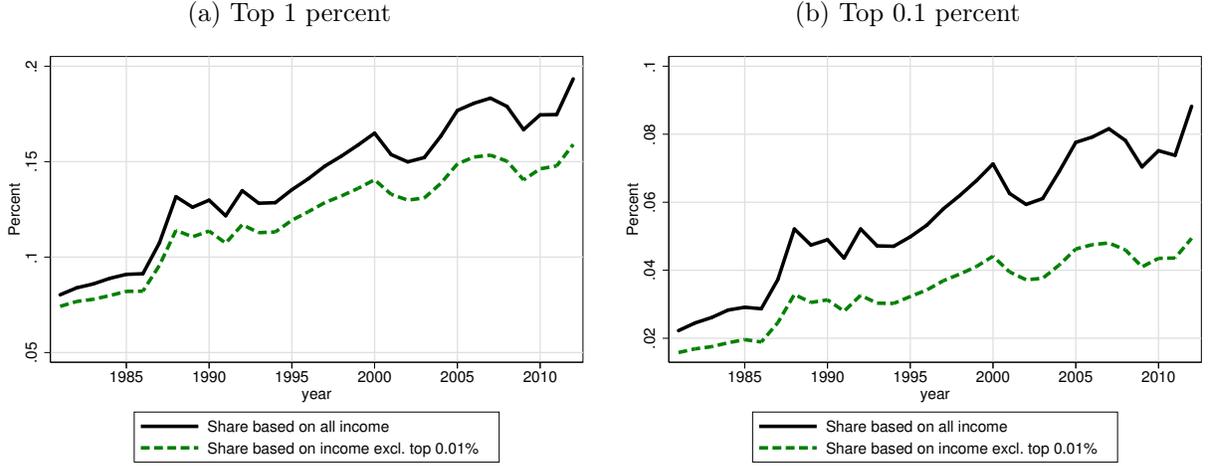
(b) Top 0.1 percent

Notes: Share of total income accruing to components of income among top percentile groups.

exclude households with income exceeding the 99.99th percentile (green dashed lines), the income shares of the top 1 percent and top 0.1 percent display a much smaller increase. From 2000 to 2012, the top 1 percent share grew by 2.8 percentage points, and the top 0.1 percent share grew by 1.7 percentage points. With the top 0.01 percent of households excluded from the analysis, the top 1 percent share grew by 1.9 percentage points, and the top 0.1 percent share grew by only 0.5 percentage point over this period.

Alternatively, we can decompose top income shares to quantify the relative importance of changes at the very top of the distribution. Let total incomes be denoted by Y, and let y_p be the pth percentile of the income distribution. Then the share of income accruing to the top pth percentile, s_p, is given by

$$s_p = \frac{E\left[Y|Y > y_p\right] p}{E\left[Y\right]}.$$

If $q > p$ is a higher percentile in the distribution, then we can decompose the top pth income share s_p as

$$
\begin{aligned}
s_p &= \frac{E\left[Y|Y > y_q\right] q}{E\left[Y\right]} + \frac{E\left[Y|y_p < Y \leq y_q\right] (p - q)}{E\left[Y\right]} \\
&= \left(\frac{E\left[Y|Y > y_q\right] q}{E\left[Y|Y > y_p\right] p} + \frac{E\left[Y|y_p < Y \leq y_q\right]}{E\left[Y|Y > y_p\right]} \frac{p - q}{p} \right) s_p.
\end{aligned}
$$

12

Figure 7: Decomposition of Top Income Shares

(a) Top 1 percent   (b) Top 0.1 percent
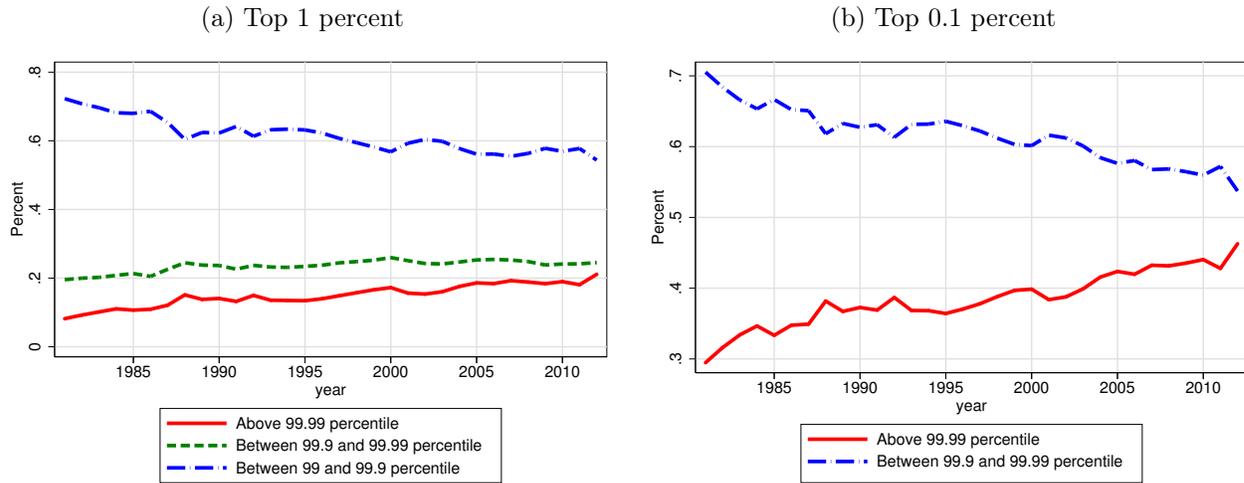
Notes: Proportion of top income shares due to incomes in different categories within the top 1 percent.

The first term in parentheses is the fraction of the top pth income share that is due to incomes above the qth percentile, and the second term in parentheses is the fraction of the top pth income share that is due to incomes between the pth and qth percentile.

The results of this decomposition, displayed in Figure 7, suggest that the overwhelming majority of the increase in top income shares is concentrated in incomes above the 99.99th percentile of the income distribution. In 1981, income earned by the top 0.01 percent accounted for less than 30 percent of the income earned by the top 0.1 percent. In 2012, the top 0.01 percent accounted for over 45 percent of the income earned by the top 0.1 percent. This shift in the composition of income among the top 0.1 percent has meant that although there has been a continued increase in the top 0.01 percent share of total income since 2000, the increase in the share of incomes accruing to the second 0.09 percent has been much smaller over this period.

5   The Bottom 99 Percent

Often overlooked is that the evolution of top income shares reflects changes at both the top of the income distribution and the bottom of the income distribution. An increase in the top 1 percent share could be due to either faster than normal income growth among

the top 1 percent or to slower than normal income growth among the bottom 99 percent. Since each of these possibilities suggests different policy responses and different avenues for theoretical investigations, in this section we suggest a decomposition to quantify the relative importance of average income dynamics in the top and bottom of the distribution. Of course, any decomposition of this type is sensitive to the definition of "normal income growth", so we present results for various alternatives.

We start by defining r_p as the following monotonic transformation of the share of income accruing to the top pth percentile, s_p,

$$q_p \equiv \log\left(s_p\right) - \log\left(1 - s_p\right) - \log\frac{p}{1-p}$$
$$= \log E\left[Y|Y > Y^p\right] - \log E\left[Y|Y < Y^p\right].$$

These transformed top income shares follow a trend similar to the top income shares in Figure 2. We can decompose the change in the transformed top income shares as

$$\Delta q_p = \underbrace{\left(\Delta \log E\left[Y|Y > Y^p\right] - g\right)}_{\text{fast growth above } Y^p} + \underbrace{\left(g - \Delta \log E\left[Y|Y < Y^p\right]\right)}_{\text{slow growth below } Y^p}.$$

The first term is the component of the change in the top income share that is due to faster than normal income growth among the top pth percent, and the second term is the component of the change in the top income share that is due to slower than normal income growth among the bottom $(1 - p)$th percent. Normal income growth is denoted by g.

Table 1 reports the results of this decomposition for the period 1981 to 2012 under alternative assumptions about normal income growth g. With assumed normal income growth over this period of 2 percent, the decomposition suggests that around half of the increase in the top 1 percent income share is due to faster than normal income growth among the top 1 percent, and the other half is due to slower than normal growth among the bottom 99 percent. For the top 0.1 percent income share, approximately 70 percent of the increase is due to faster than normal income growth among the top 0.1 percent, and the remaining 30 percent is due to slower than normal income growth among the bottom 99.9 percent.

Assuming faster normal annual income growth than 2 percent delivers a decomposition in which stagnation in the bottom percentiles contributes more to the increase in top income shares, whereas assuming slower normal income growth leads to a decomposition in which acceleration in the top percentiles contributes more. Annual growth rates in average total incomes (excluding capital gains) have fluctuated a lot over the last century. Over the 30

Table 1: Decomposition of Growth in Top Income Shares, 1981-2012

Normal Income Growth (p.a.)	Growth in Top 1% Share		Growth in Top 0.01% Share	
	Slow Growth at Bottom (%)	Fast Growth at Top (%)	Slow Growth at Bottom (%)	Fast Growth at Top (%)
1%	17	83	8	92
1.5%	32	68	18	82
2%	47	53	29	71
2.5%	77	23	39	61
3%	92	8	50	50

years immediately prior to 1981, average annual income growth was around 1.7 percent, from 1933 to 1981 average growth was 2.8 percent and from 1913 to 1981 average growth was 1.5 percent. With assumed normal growth rates anywhere in this range, stagnation of incomes in the bottom percentiles plays an important role in explaining the growth in top income shares.

6 Taking Stock

Since 2000, different measures of top income inequality have exhibited very different trends. Top income shares based on measures of total income show a continued rise, whereas top income shares based on wage and salary income show no increase in inequality post-2000. The most important difference between these two measures of income is the income that accrues to S-corporations. Moreover, the majority of the recent increase in top income shares is due to an increase in average earnings at very high income levels, much higher than that assumed in typical discussions of top income inequality. Once incomes above the 99.99th percentile are excluded (around $7 million in 2012), we see that little continued growth in top income shares has taken place in the last 20 years. Put simply, so far in the 21st century, all the action in top income shares has been S-corporation income at very, very high income levels.

What does this all mean for researchers and policy makers who are interested in the causes and consequences of changes in inequality?

First, when we discuss inequality, even inequality at the top, we must be clear about exactly which part of the distribution we are talking about. Different data sets and different income concepts are appropriate for different parts of the distribution. For example, when

discussing individuals with the very highest incomes, it is crucial to use data that include S-corporation and partnership income; these are available in IRS data but not in SSA data. At the same time, it is important to not lump all individuals in the top 1 percent or top decile in with these very, very high earners. Individuals above the 90th or 99th percentile, but below the 99.99th or 99.9th percentile, have seen trends in income growth that are very different from those of individuals at the very top, and their incomes tend to be driven mostly by wage, salary, and Schedule SE income.

Second, in discussions of income inequality, we should always be aware of the unit of analysis. The unit of analysis in the IRS data is fundamentally a tax unit, which is not necessarily a unit that we care about for welfare and policy purposes. Tax units are a legal construct that respond strongly to changes in incentives, individual incomes, and aggregate economic conditions, and many individuals in the United States are not part of any tax unit since they do not file tax returns. On the other hand, the unit of observation in the SSA data is an individual, which is essentially exogenous to tax policy and economic conditions. This distinction is particularly important for the bottom of the distribution. Hence, for discussions of inequality that compare individuals with high (but not super-high) incomes to individuals in the bottom half of the distribution, the SSA data may be more appropriate than the IRS data.

Third, we need to be careful about the impact that tax incentives have on the allocation of incomes to different sources. We have seen that the most important component of income that is included in the IRS total income measure but missing from the SSA labor income measure is S-corporation income. But interpreting trends in the S-corporation component is extremely difficult. Feenberg and Poterba (1993), Gordon and Slemrod (2002), and Cooper et al. (2016) warn that much of the recent increase in S-corporation income is income that previously accrued to C-corporations. Such income is not "new" income earned by top earners but is simply income that was previously labeled as corporate income rather than household income. A large amount of this relabeling occurred in the wake of the reductions in personal income tax rates that were part of the TRA86. This income shifting potentially accounts for all of the increase in S-corporation income at the top of the distribution prior to 2000.

Focusing on W-2 income largely avoids the issue of income moving between the corporate and personal sectors but misses the income sources where all the action at the top of the distribution has been so far in the 21st century. Moreover, W-2 income is not completely immune from income-shifting activities. When an entity switches from a C-corporation to an S-corporation, some of the income that was previously paid to owners as W-2 income

will be reclassified as S-corporation income. This could lead to a fall in W-2 income that may partly account for the slight downward trend (or lack of an upward trend) post-2000 in top wage and salary income shares. We think that further investigation of these issues – for example the ongoing work by Smith et al. (2017) – is tremendously important for understanding the trends in top income shares in the 21st century.

References

Auerbach, Alan J. and James M. Poterba, "Why Have Corporate Tax Revenues Declined?," NBER Working Papers 2118, National Bureau of Economic Research, Inc September 1987.

Auerbach, Alan J and Joel Slemrod, "The economic effects of the Tax Reform Act of 1986," *Journal of Economic Literature*, 1997, *35* (2), 589–632.

Bakija, John, Adam Cole, and Bradley Heim, "Jobs and Income Growth of Top Earners and the Causes of Changing Income Inequality: Evidence from U.S. Tax Return Data," Working Paper, Williams College 2012.

Bricker, Jesse, Alice Henriques, Jacob Krimmel, and John Sabelhaus, "Measuring income and wealth at the top using administrative and survey data," *Brookings Papers on Economic Activity*, 2016, *2016* (1), 261–331.

Cooper, Michael, John McClelland, James Pearce, Richard Prisinzano, Joseph Sullivan, Danny Yagan, Owen Zidar, and Eric Zwick, "Business in the United States: Who Owns It, and How Much Tax Do They Pay?," *Tax Policy and the Economy*, 2016, *30* (1), 91–128.

Feenberg, Daniel R and James M Poterba, "Income inequality and the incomes of very high-income taxpayers: evidence from tax returns," *Tax policy and the economy*, 1993, *7*, 145–177.

Gordon, Roger H and Joel Slemrod, "Are real responses to taxes simply income shifting between corporate and personal tax bases?," in Joel Slemrod, ed., *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*, Cambridge: HARVARD University Press, 2002.

Guvenen, Fatih, Greg Kaplan, and Jae Song, "The Glass Ceiling and The Paper Floor: Gender Differences Among Top Earners, 1981–2012," Working Paper, University of Minnesota 2014.

Saez, Emmanuel, "Striking it Richer: The Evolution of Top Incomes in the United States," Working Paper, University of California at Berkeley March 2012.

Smith, Matthew, Danny Yagan, Owen Zidar, and Eric Zwick, "Capitalists in the 21st Century," Technical Report, University of Chicago 2017.