CAPITAL AND LABOR: THE FACTOR INCOME COMPOSITION OF TOP INCOMES IN THE UNITED STATES, 1962–2006

Anthony B. Atkinson and Christoph Lakner*

Abstract—This paper finds that capital and labor incomes in the United States have become more closely associated since the 1980s. This has contributed to the well-known increase in the top 1% share of total income, exacerbating rising inequality in capital incomes and earnings. We show that the trend in the association is U-shaped as the recent increase contrasts with a tendency toward a weakening association until the 1980s. The paper, using data derived from tax records, studies the asymmetries in the association and tests for robustness to alternative income definitions, including the role of income from closely held businesses at the top.

I. Introduction

RECENT literature has documented the increase in income inequality at the very top of the distribution (Atkinson & Piketty, 2007, 2010). Between the mid-1980s and the mid-2000s, the income share of the top 1% in the United States approximately doubled, while around the same time, their share of income from capital declined, and salaries and incomes from closely held businesses became more important (see figure 1; also see Piketty & Saez, 2007a). These changes in the income composition have been even more pronounced over the long run, with the share of income from capital among the top 1% dropping from close to 50% in the 1920s to less than 20% in the 2000s (Piketty & Saez, 2003, 2007a). Such an increase in the labor share at the top could arise from a change in the association between labor and capital or because the share of tax units in the top who receive only capital income declines while there remain separate classes of laborers and capitalists. As part of the “hypermeritocratic society” (Piketty, 2014), the inequality in the wage distribution has increased substantially (Piketty & Saez, 2007a), providing support for the second channel. At the same time, tax units increasingly have income from both capital and labor (Wolff & Zacharias, 2009), and pure rentiers have virtually disappeared (Atkinson, 2009). Furthermore, income from closely held businesses, in particular partnerships and S-corporations, has become increasingly important at the top (Smith et al., 2019).

While the literature (Piketty & Saez, 2007a) has focused on the distribution of total income, as well as the distributions of capital and labor income separately, the association between the two income sources has received little attention. To extend the existing literature, this paper examines the association between capital and labor incomes at the top of the U.S. income distribution. Using data based on U.S. tax returns, we can directly measure the association. In a classical society, capitalists are at the top of the capital distribution and bottom in the wage distribution, that is, the correlation between capital and labor incomes is negative. On the other hand, this correlation is generally positive (but less than 1) in modern economies (Piketty, 2014). Our paper also links the study of the functional (capital versus labor) and the personal (rich versus poor) income distributions. For classical economists, there was a clear mapping between these two types of distribution, with capitalists being rich and workers poor (Milanovic, 2017). This mapping is more complicated in modern economies, since many people have income from more than one source (Atkinson, 2009), the inequality within capital and labor incomes is increasing (Lydall, 1968), and mixed income is becoming more important at the top (Smith et al., 2019).

An important issue for our study is how to split income from closely held businesses, particularly partnerships and S-corporations, between capital and labor. This source of income has become increasingly important for the top 1% (figure 1). Smith et al. (2019) show that most of this income is a return to human capital and should thus be allocated to labor. Their findings support the view of Piketty and Saez (2003) that the working rich, who are increasingly paid in the form of capital incomes, are at the top of the capital distribution and at the bottom of the labor distribution.

Received for publication December 10, 2017. Revision accepted for publication February 19, 2020. Editor: Brigitte C. Madrian. Anthony Atkinson sadly passed away in January 2017 before the new version of the paper was finalized. This paper would not exist without his contributions. Subsequent revisions were carried out by C. Lakner, who is responsible for any errors. This paper has been developed from one of the chapters of C. Lakner’s DPhil thesis.

Emmanuel Saez participated in the first stages of this project, and we thank him for his assistance with the U.S. data. We also thank David Splinter, Gabriel Zucman, and Eric Zwick for answering questions about their data and papers. We thank the editor, Brigitte Madrian, and four anonymous referees, whose comments helped us improve the paper substantially. We are grateful to seminar participants at the Oxford INET, the World Bank, and the ECINEQ and LIS conferences, and to Rolf Aaberge, François Bourguignon, Francisco Ferreira, La-Bhus Fah Jarasavattakul, Stephan Klasen, Sebastian Königs, Branko Milanovic, Bob Rijkers, Tim Smeeding, Francis Teal, and Adrian Wood for discussions and comments on earlier versions. C.L. gratefully acknowledges financial support from the Economic and Social Research Council (grant ES/G011974/1), the Studienstiftung des deutschen Volkes, the Institute for New Economic Thinking (INET) at the Oxford Martin School, and the Center for Equitable Growth at UC Berkeley. The findings, interpretations, and conclusions expressed in this paper are entirely our own. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations or those of the executive directors of the World Bank or the governments they represent.

A supplemental appendix is available online at https://doi.org/10.1162/rest_a_01063.

1S-corporations are businesses with few shareholders that are taxed at the personal instead of the corporate level. Partnerships are also taxed at the personal level. See Smith et al. (2019) for an overview of the different corporate forms and how they are taxed.

2As explained by Smith et al. (2019), the existing literature uses a dollar-level analysis. This paper, like Smith et al., uses data based on tax returns to study person-level questions. Aaberge, Atkinson, and Königs (2018), which has been developed in parallel to our paper, analyze the association between capital and labor incomes in Norway.
of business incomes, are an important driver for the increase in top income shares. In our baseline results, we use the factors estimated by Smith et al. (2019), but in the appendix, we show that the results are robust to following Piketty, Saez, and Zucman (2018), who apply a smaller labor share to these incomes.

We find that the association between capital and labor incomes increased between 1988 and 2006. Tax units at the top of the distributions of capital and labor incomes are increasingly the same people. There is some evidence of a U-shaped pattern: the association became weaker from 1966 to the mid-1980s and then started increasing. The association is found to be asymmetric in some parts of the distribution, with four-fifths of the top 1% of earners being among the top quintile of capital incomes, compared with only two-thirds of the top 1% capitalists being in the top quintile of earnings. That is, top wage earners are almost guaranteed to also receive considerable capital incomes, while rentiers with little wage income have not disappeared from the top. This asymmetry is not found at the very top of the distribution (top 5% and up), suggesting that tax units at the very top of either capital or labor incomes face a high probability of being at the top of the other distribution. This finding is consistent with an increasing role of income from closely held businesses at the very top. Such income has become increasingly concentrated, as well as increasingly associated with total income, driven primarily by income from partnerships and S-corporations. However, even without business income, labor incomes have become more closely aligned with total incomes, suggesting an important role for the working rich.

In almost all our results, the 1980s mark a turning point, which coincides with the sweeping changes to the U.S. federal income tax introduced by the Tax Reform Act of 1986 (TRA86). An important question is to what extent our results, and especially the split between capital and labor, are affected by retiming or accounting responses to the tax reform (Slemrod, 1992). In particular, S-corporation filing status became more attractive following TRA86, which reduced the top personal tax rate below the corporate tax rate (Slemrod, 1996; Auerbach & Slemrod, 1997). While partnership and S-corporation incomes account for a quarter of the increase in the top 1% share observed around the years of TRA86, an analysis of a panel of tax units shows that three-fifths of this extra income at the top stems from existing partnerships and S-corporations (Auten & Splinter, 2019). Furthermore, when there are passive owners, active owners have an incentive to pay themselves high salaries since any profits are split across

---

3In most specifications, we also find an increase in the association over the entire period (from 1966 to 2006). However, these results are difficult to interpret since they span the Tax Reform Act of 1986 (TRA86). Hence, we focus on the post-TRA86 period (beginning in 1988) when we test for increasing association. Furthermore, the increase in the association between 1966 and 2006 is not confirmed for all of the robustness checks we consider.

4Earnings and labor incomes are used interchangeably. In most of the analysis, capital and labor incomes include a share of income from closely held businesses, as described in more detail in section II.

5By reducing the top marginal personal income tax rate below the corporate rate, TRA86 provided incentives to move income from C-corporations (subject to corporate income tax) to S-corporations (subject to personal income tax). TRA86 also raised tax rates on capital gains by including all realized capital gains in taxable income (Slemrod, 1996).
both active and passive owners (Piketty et al., 2018). This would tend to shift income in the opposite direction, toward labor.\footnote{This suggests that the residual S-corporation profit should be allocated to capital. With lower tax rates at the top, active owners may bargain more aggressively for higher salaries (Alvaredo et al., 2013).}

We address this issue in several ways. First, our analysis tends to focus on long-term trends to abstract from short-run fluctuations resulting from accounting responses (Alvaredo et al., 2013). Second, when we analyze changes between two years, we leave out the years most affected by TRA86 and concentrate on the post-reform period.\footnote{For example, in section IV, we test for an increase in the association between 1966 and 1985, and 1988 and 2006. After TRA86, the incentives to file as an S-corporation or a C-corporation have remained similar (Smith et al., 2019). More generally, the result of an increasing association until the mid-2000s is robust to choosing later years, such as 1990 and 1992, as the starting point for the second interval (see the appendix).} Third, the results are robust to a broad range of alternative income definitions, such as the inclusion of capital gains, an important channel of tax avoidance (Piketty, Saez, & Stantcheva, 2014). The results are also robust to scaling up the tax-data-based capital incomes to match national accounts, to account for some of the shifting of income between corporate and personal taxation. Finally, the results are robust to alternative splits of income from closely held businesses between capital and labor.

We examine the association in two ways. We begin by decomposing the top 1% income share by factor incomes, which is frequently done with standard inequality measures but has not been applied to top income shares. The inequality in total income is decomposed into the labor share in total income, the top 1% share within the distributions of capital and labor incomes, and the alignment coefficient, which captures the association. The alignment coefficient, like the Pearson correlation coefficient, is affected by monotone transformations in the marginal distributions. In the second part of the paper, we thus use a rank-based measure of association that is more general and invariant to such transformations. Specifically, we analyze the association matrices between labor and capital, which are a discrete approximation to the copula density and equivalent to transition matrices in the study of economic mobility. The literature on mobility also provides a test of increasing association based on cumulative association matrices.

The remainder of this paper is structured as follows. Section II describes the data. In section III, we develop and estimate a decomposition of top income shares by factor incomes. In addition to the main results for capital and labor, this section also provides a detailed account of the role of income from closely held businesses by considering it as a separate category. Section IV reports the results from the association matrices. Section V concludes and discusses how the changes in the association may be related to top marginal tax rates. The online appendix provides additional results and several robustness checks. Our results are robust to treating negative incomes and capital gains differently, to including estimates for tax units that do not file a return, to scaling up the tax-data-based capital incomes to match national accounts, and to how income from closely held businesses is split between capital and labor.

\section{The Data}

We use the Statistics of Income public use files (PUF) by the U.S. Internal Revenue Service over the period 1962 to 2006, which are (nearly) annual public-use samples based on federal income tax returns.\footnote{Like other papers using these data (Saez & Zucman, 2016; Piketty et al., 2018), we exclude the microdata for 1960, since they contain fewer tax return variables. There exist no PUFs for 1963 and 1965. More information on these data is available at http://users.nber.org/~taxsim/gdb/.} The data are based on a random sample of tax records filed during a particular calendar year. Importantly for our analysis, high-income returns are oversampled, and we use the sampling weights provided to adjust for this. Following Piketty and Saez (2007a) and the literature using tax records more generally, the unit of analysis is a tax unit as defined under U.S. tax law. We thus include singles and married couples without adjusting for differences in tax unit size.\footnote{Our results are robust to including only tax units that have two adults. Between the 1960s and 2000s, the average tax unit size declined from 2.6 to almost 2 persons. The proportion of married tax units declined by 10 percentage points in the whole population but slower at the top of the distribution (Saez, 2004). Lakner (2014) shows that the trend in top shares is robust to accounting for tax unit size. Piketty et al. (2018) also find very similar trends for tax units and adults. Tax units tend to be smaller than households; Hungerford (2010) estimates 75% of households to have one tax unit and another 17% to have two tax units.} Not every tax unit files a tax return; historically, there were high exemption levels and income taxes applied only to the most affluent taxpayers. In the baseline results, we ignore the nonfiling issue because it requires an arbitrary assumption about the income composition of the nonfilers and is unlikely to affect our results, since the fraction of filers is high and stable over the period of analysis (94% on average, compared with 9% before World War II). The appendix provides a robustness check that includes nonfilers, with very similar results. The PUFs are subject to some adjustments, especially at the top, that try to minimize the risk that individual taxpayers can be identified (Winglee et al., 2002). As a result, an observation in the PUF never contains all the information on a tax return and may include information from other returns. Further information is provided in appendix A.3, where we also show some robustness checks.

Income is defined as (taxable) gross market income, as reported on federal income tax returns.\footnote{We exclude any income that is not taxable (e.g., nontaxable fringe benefits such as health insurance), since it is not reported on the tax return. We also exclude nonmarket or transfer income such as Social Security and unemployment insurance benefits. Since the PUFs do not capture all tax liabilities (e.g., exclusion of state and local taxes), we focus on gross incomes, as do Piketty and Saez (2007a), and most of the literature on top incomes. It is also unclear how to split total federal income taxes between capital and labor without additional imputations. Finally, most of our analysis uses income that is observed from tax records (like Piketty & Saez, 2007a), which does not match national income (like Piketty et al., 2018) or macro-totals in the financial accounts (like Saez & Zucman, 2016) (also see figure A.6). However, the appendix includes robustness checks that scale up} We follow Piketty and Saez (2007b), Piketty et al. (2018), and Saez and...
Zucman (2016) to construct income components from the raw data that are comparable over time. We define labor income as the sum of wages and (taxable) pensions. Capital income is defined as the sum of dividends, (taxable) interest, rents, estate income, and royalties. An important question is how mixed income of noncorporate businesses, which reflects returns to both human and physical capital, is split between labor and capital. We allocate income from sole proprietorships (Schedule C) two-thirds to labor and one-third to capital. While these weights are arbitrary, they are similar to the earlier literature and close to factor shares found in national accounts (Gollin, 2002; Feldstein, 2008; Elsby, Hobijn, & Sahin, 2013; Karabarbounis & Neiman, 2014). Income from partnerships and S-corporations, which has become increasingly important for top incomes (figure 1; see also Cooper et al., 2016), is split three-quarters to labor and one-quarter to capital, using the allocation factors estimated by Smith et al. (2019). Because negative incomes can result in top shares for the distribution of capital or labor income that are greater than 1, we drop observations that are negative in labor, capital, sole proprietorship, or pass-through income (defined as partnerships plus S-corporations). Table 1 presents summary statistics for the baseline income definition. The number of observations is 84,000 per year on average.

In appendix A.5, we show that our results are robust to several alternative specifications. First, instead of dropping negative observations, we set them to 0 (similar to Saez & Stantcheva, 2018). Second, we include capital gains, which are an important income source at the top. The tax data only report realized capital gains, which are lumpy because realizations respond to changes in the tax code or asset prices. An accruals-based approach changes the timing of capital gains in the short run, but the long-run trend remains similar (Larrimore et al., 2021). Third, we include nonfilers. Fourth, we scale up the capital incomes recorded in the tax data to match retained earnings reported in national accounts and distribute the corporate income tax to tax units. In the baseline, we include only income that is taxable at the personal level, so we exclude undistributed corporate profits. Fifth, we use a different rule for splitting income from closely held businesses (similar to Piketty et al., 2018, and CBO, 2012), by including S-corporation profits with capital and allocating two-thirds of partnership and sole-proprietorship income to labor.

### III. Decomposition by Factor Incomes

We begin the analysis with a decomposition of top income shares by factor incomes. This is a formal derivation and the first empirical application of the decomposition by Atkinson (2007), who builds on Meade (1964). It is closely related to factor income decompositions of other inequality measures (Shorrocks, 1982; Lerman & Yitzhaki, 1985; Milanovic, 2017). Like these decompositions, inequality in total income is decomposed into three elements: the share of each factor in total income, the inequality in the distribution of income from each of the factors, and a term capturing the association between the incomes from different factors and total income.

The income share of top quantile $i$ can be written as $S_i = \frac{1}{T}$, where $Y$ is total income in the data and $Y_i$ denotes total income of tax units with income greater or equal to $y_i$, the threshold income (e.g., $y_i$ is the 99th percentile in the case of the top 1% income share). For any individual $j$, total income $y_j$ is derived from $M$ components, such that $y_j = \sum_{m=1}^{M} x_{j,m}$, where $x_{j,m}$ is individual $j$’s income from factor $m$. Therefore, the numerator of $S_i$ can be written as $Y_i = \sum_{j=1}^{N} y_j \times 1 \{y_j \geq y_i\} = \sum_{j=1}^{N} \sum_{m=1}^{M} x_{j,m} \times 1 \{y_j \geq y_i\}$. Defining $\bar{X}_{i,m} = \sum_{j=1}^{N} x_{j,m} \times 1 \{y_j \geq y_i\}$ (i.e.,

| Year | Overall mean | Top 5% mean | Top 1% mean | Top 0.5% mean | Number of observations |
|------|--------------|-------------|-------------|--------------|------------------------|
| 1966 | 36,942       | 145,850     | 282,535     | 372,306      | 71,322                 |
| 1970 | 39,665       | 150,851     | 278,648     | 362,651      | 70,628                 |
| 1980 | 38,200       | 148,347     | 270,494     | 356,229      | 117,113                |
| 1985 | 38,198       | 152,622     | 282,462     | 387,047      | 66,504                 |
| 1990 | 40,352       | 197,711     | 449,369     | 665,475      | 65,599                 |
| 2000 | 48,985       | 294,495     | 750,024     | 1,146,274    | 110,888                |
| 2006 | 49,284       | 308,672     | 802,109     | 1,230,014    | 111,694                |

All mean incomes refer to sum of income components per tax unit, and are expressed in 2006 USD. Excluding capital gains, negative observations and nonfilers. CPI data are taken from Piketty, Saez, and Zucman (2019).
the total income from factor \( m \) among top quantile \( i \), the top income share can be written as

\[
S_i = \sum_{m=1}^{M} \frac{X_{i,m}}{Y} = \sum_{m=1}^{M} \frac{X_{m}}{Y} \frac{X_{i,m}}{X_{m}} = \sum_{m=1}^{M} \frac{\mu_m}{\mu} S_{i,m} A_{i,m}, \tag{1}
\]

where \( X_m = \sum_{j=1}^{N} x_{j,m} \) is the total income from factor \( m \). \( X_{i,m} = \sum_{j=1}^{N} x_{j,m} \times 1 \{ x_{j,m} \geq x_{i,m} \} \) is the total income from factor \( m \) among top quantile \( i \) of the distribution of component \( m \). The first term in the final expression is the share of total income derived from income source \( m \). \( S_{i,m} = \frac{X_{i,m}}{X_m} \) denotes the share of total income from factor \( m \) that accrues to the top quantile \( i \) of recipients of income from factor \( m \), for example, the share of capital income going to the top 1% of capitalists. It thus captures inequality in the marginal distribution. Atkinson (2007) refers to the final term as the “alignment coefficient,” which captures the extent to which the rankings according to income from factor \( m \) and total income coincide. It is defined as \( A_{i,m} = \frac{S_{i,m}}{S_{i}} \), where \( S_{i,m} = \frac{X_{i,m}}{X_m} \) is the share of total income from factor \( m \) received by the top quantile \( i \) of total income recipients.\(^\text{16}\) The alignment coefficient lies between 0 and 1 since shares are nonnegative and \( S_{i,m} \geq S_{i,m} \).\(^\text{17}\) If top income recipients (according to total income) receive no labor income, \( \dot{X}_{i,l} = 0 \) and \( A_{i,l} = 0 \). On the other hand, if everybody in the top quantile \( i \) of the total income distribution is also found in the top quantile \( i \) of the distribution of labor income, then \( S_{i,l} = S_{i,l} \) and \( A_{i,l} = 1 \).

### A. Results

Figure 2 shows the results of the factor income decomposition for the top 1%. Panel A shows \( S_1 \), the top 1% share of total income, which roughly doubled over this period (as was already shown in figure 1). While we estimate the top share at a somewhat lower level than Piketty and Saez (2007a) do, the two series track each other very closely, as discussed in appendix A.4. The remaining three panels of figure 2 refer to the different components of the decomposition: the labor share (panel B); the share in total labor (capital) income of the top 1% of labor (capital) income recipients (panel C, using separate axes); and the alignment coefficients (panel D). The results for the top 5% and top 0.5% are shown in the appendix (figures A.1 and A.2). In almost all our results, the

\(^{16}\)Following Shorrocks (1982), \( \dot{S}_{i,m} \) may be called the “pseudo share.” It is different from \( S_{i,m} \) because observations are ranked according to total, not factor, income.

\(^{17}\)For example, this can be seen by noting that the labor income of the tax units in the top quantile of labor but not total income is greater than the labor income of the tax units in the top according to total but not labor income.
1980s mark a turning point, which coincides with the introduction of TRA86. As discussed in section I, we focus on long-run trends to abstract from short-run fluctuations that are due to retiming or accounting responses (Slemrod, 1992). Furthermore, our results are robust to a wide range of alternative income specifications (see appendix A.5), such as the inclusion of capital gains, scaling up capital incomes to match retained earnings in national accounts, and alternative splits of partnership and S-corporation income between capital and labor.

The labor share in total income (panel B) fluctuates between 87% and 93%, with no clear trend. At around 75%, Piketty et al. (2018) find a much lower labor share over this period, which is also approximately constant. They estimate a higher capital share because they include capital incomes that are not reported on personal tax returns, such as imputed rents for owner-occupiers, dividends and interest paid to pension funds, and corporate retained earnings (also see appendix A.5, where we include corporate retained earnings).

The inequality in labor incomes, as measured by the top 1% share, increased similar to the inequality in total income, although the top labor share remains at a slightly lower level. The share of labor income going to the top 1% of earners approximately doubled, from 6% in the 1960s to 13% in the 2000s (panel C, left axis). These results mimic the estimates by Piketty and Saez (2007a), who impute for nonfilers and also present independent evidence on executive compensation. Capital incomes are distributed more unequally than either labor or total income, as one would expect (Piketty, 2014). The top 1% share of capital incomes fell until the 1980s, then increased similar to labor and total income but continued to rise in the 2000s. The top 1% of capitalists now account for almost half of capital incomes, compared with less than 30% in the 1980s (panel C, right axis). These results follow a similar trend to the taxable capital income shares reported by Saez and Zucman (2016), who allocate all self-employment income to capital.

The alignment coefficient for labor income declined slightly from 92% to 89% in the early 1980s, before rising to 98% in the 2000s (panel D, left axis). For capital income, the alignment coefficient is lower and follows a U-shaped pattern; it declines from almost 80% in the 1960s to 60% in the 1980s, before rising to 80% by the end of the period. A value of 80% for the capital alignment coefficient means that 80% of total capital income of the top 1% of capitalists goes to tax units that are also in the top 1% of total income.

Given that the top 1% capitalists receive 48% of capital income (panel C), this implies that 38% of all capital income goes to tax units in the richest 1% (this is $\tilde{S}_{1,c}$ above). The same statistic was around 16% of all capital income in the mid-1980s. These estimates suggest that over the past twenty years, capitalists are increasingly also at the top of the income distribution. Labor income has an even stronger association with total income: around 98% of labor income of the top 1% of earners is received by tax units that are also in the richest 1%, compared with 80% for capital. We examine this asymmetry in the association in more detail below.

In the next section, we look more closely at changes in the association between 1988 and 2006. Over this period, the top 1% share of total income increased by just over half (from 10.7% to 16.3%, panel A). Labor incomes matter more for explaining the top 1% share of total income, since labor accounts for a much greater share of total income (panel B) and since the alignment coefficient is higher for labor than for capital. For labor, the factor share, the top 1% share, and the alignment coefficient all increased between 1988 and 2006. For capital, the factor share declined, while the top 1% share and the alignment coefficient increased between 1988 and 2006.

The analysis has so far focused on splitting total income into labor and capital incomes, although the decomposition in equation (1) can be applied more generally to $M$ income components. Figure A.3 decomposes total income among labor, capital, and business income. Because the income from closely held businesses has become more important at the top (see the discussion in section I), we consider it here as a separate component.

As before, there is no clear trend in the labor and capital factor shares (panel A). However, the share of total income from closely held businesses has increased (e.g., from 8.1% in 1988 to 11.6% in 2006). The level of and rise in inequality within labor incomes are lower when business income is excluded but still increases (panel B). Labor income in inequality within business incomes increased, and since the late 1960s, it is more unequally distributed than either labor or capital (panel C; also see Cooper et al., 2016, for a comparison with the concentration of capital incomes). Without business income, the association of labor income with total income is lower, but there is an almost continuous increase over the entire period (panel D). For capital, the association with total income is also lower when business income is excluded, but the time trend is unchanged (panel E). Business incomes at the top are increasingly aligned with total incomes (e.g., the business alignment coefficient increased from 70.3% to 79.4% between 1988 and 2006, panel F), and since the 1980s, they exhibit a stronger association with total income than capital but weaker than labor.

---

18Saez and Veall (2007) find that in Canada, top wages increased similarly to the United States without the same changes in fiscal policy, suggesting that the U.S. increase was real, and not simply an accounting response due to changes in the tax code.

19In their main results for wealth, Saez and Zucman (2016) include other assets that do not generate taxable incomes, such as primary housing. The top 1% share of wealth has increased less than the top 1% share of taxable capital income.

20When retained earnings and the corporate income tax are distributed to tax units, the labor alignment coefficient is much lower initially and increases during most of the period (figure A.7).

21The labor alignment coefficient tends to exceed the capital coefficient also for other income definitions (figure A.7). When capital gains are included and capital incomes are scaled up to national accounts aggregates, the asymmetry only holds from the late 1980s onward.
To understand the role of different types of business incomes, figure A.3 presents two additional series that split out S-corporation income, as well as both partnership and S-corporation income (i.e., only sole proprietor income remains), respectively. S-corporation incomes largely account for the increasing share of business income in total income (panel A). For example, without S-corporation income, the business share increases only by 1 percentage point between 1988 and 2006, while it increases by 3.5 percentage points in the baseline. Focusing on the top, partnership and S-corporation incomes still accounts for most of the changes in business income, but among these incomes, it is now partnership income instead of S-corporation income that matters. For example, the increase in the top 1% share of business income is very similar when S-corporation incomes are excluded (panel C), but when partnership incomes are also dropped, the increase disappears. Similarly, the alignment coefficient for business incomes increases by 8.1 percentage points without S-corporation income (compared with 9.1 percentage points in the baseline), while it falls by −3.6 percentage points when partnership income is also dropped (panel F).

In sum, these results suggest that even when incomes from closely held businesses are excluded, labor incomes have become more closely associated with total incomes at the top, suggesting an important role for the working rich in explaining the rising association. Business incomes themselves have become more concentrated and more closely associated with total income. This has been driven in particular by income from partnerships and S-corporations.

IV. Rank-Based Measure of Association

The association measure that we have used so far, the alignment coefficient, is not independent of monotone transformations in the marginal distributions. For example, a doubling of all labor incomes would tend to change the ranking in the distribution of total income and thus affect the alignment coefficient through $S_{im}$. Only a rank-based measure of association is invariant to all monotone transformations in the marginals (Dardanoni & Lambert, 2001). This means that a rank-based association measure is unaffected by an increase in the inequality in labor incomes that leaves ranks unchanged. In the remainder of the paper, we use an analytical framework based on the copula function, which offers a clean separation of the joint distribution of labor and capital into the marginal distributions and a rank-based measure of the association. Our rank-based association measure is also more general because it considers the entire distribution, while the alignment coefficient, for, say, the top 1% is determined only by whether observations cross the 99th percentile.

Total income is a two-dimensional vector $X = (L, K)$, where $L$ denotes labor income and $K$ refers to capital income. By Sklar’s theorem (Nelsen, 2006), there exists a copula function $C_X$ such that $H_X(l, k)$, the joint distribution function of $X$, can be written as

$$H_X(l, k) = C_X [F(l), G(k)],$$

where $F(l)$ and $G(k)$ are the marginal distribution functions of labor and capital income. The density of the joint distribution is obtained by differentiating with respect to $l$ and $k$,

$$h(l, k) = f(l)g(k)C_{FG}[F(l), G(k)],$$

where $C_{FG}(F(l), G(k))$ is the copula density. The joint density can thus be expressed as the product of the marginal distributions of labor and capital income and the copula density. The copula density is a rank-based measure of the association. A discrete approximation for the copula density is the association matrix between labor and capital income (Bonhomme & Robin, 2009), shown for 2006 in table 2. It shows how the population is distributed across the following eight quantile groups.

$\text{Copula functions have been widely used in actuarial science to describe multidimensional risks. In economics, they have been used to study the joint distribution of income and wealth (Kennickell, 2009; Jäntti, Sierminska, & Van Kemm, 2015), the horizontal equity of the tax system (Dardanoni & Lambert, 2001), income mobility by considering the dependence over time (Bonhomme & Robin, 2009), and multidimensional inequality and poverty (Atkinson, 2011; Ferreira & Lugo, 2013; Decancq, 2014). Aaberge et al. (2018) analyze the association between capital and labor incomes in Norway using a copula-based framework.}$

$\text{While there is considerable overlap between top partnerships and top S-corporations in terms of economic sector, high-skilled services such as legal services and financial investment firms account for a greater share of top partnership profits (Smith et al., 2019).}$

### Table 2.—Association Matrix in 2006 (Frequencies in %)

| Capital | P50 | P50–P60 | P60–P80 | P80–P90 | P90–P95 | P95–P99 | P99–P99.5 | >P99.5 |
|---------|-----|---------|---------|---------|---------|---------|-----------|--------|
| P50     | 30.582 | 2.948   | 7.553   | 5.054   | 2.336   | 1.320   | 0.122     | 0.900  |
| P50–P60 | 6.233 | 0.874   | 1.537   | 0.651   | 0.435   | 0.245   | 0.016     | 0.012  |
| P60–P80 | 9.390 | 2.729   | 4.453   | 1.696   | 0.867   | 0.785   | 0.040     | 0.030  |
| P80–P90 | 2.937 | 1.808   | 3.125   | 1.101   | 0.484   | 0.508   | 0.026     | 0.021  |
| P90–P95 | 0.829 | 0.876   | 1.820   | 0.651   | 0.359   | 0.396   | 0.039     | 0.020  |
| P95–P99 | 0.374 | 0.399   | 1.372   | 0.698   | 0.392   | 0.542   | 0.157     | 0.065  |
| P99–P99.5 | 0.010 | 0.017   | 0.093   | 0.091   | 0.071   | 0.100   | 0.056     | 0.063  |
| >P99.5  | 0.004 | 0.003   | 0.036   | 0.054   | 0.055   | 0.106   | 0.042     | 0.200  |
| Total   | 50   | 10      | 20      | 10      | 5       | 4       | 0.5       | 0.5    | 100    |

30.582% of tax units are in the bottom half of both the labor and capital income distribution. In other words, 61% of tax units that are in the bottom half of labor incomes are also in the bottom half of capital incomes.
of labor and capital incomes: The bottom 50% (≤ P50), the next 10% (P50–P60), the next 20% (P60–P80), the next 10% (P80–P90), the next 5% (P90–P95), the next 4% (P95–P99), the next 0.5% (P99–P99.5), and the top 0.5% (> P99.5). For example, table 2 shows that 0.2% of observations were both in the top 0.5% of the labor and the top 0.5% of the capital income distribution. This is greater than 0.0025%, which would be the frequency if the two variables were independent but less than 0.5%, the frequency with perfect association.

The association matrix is equivalent to transition matrices used to study economic mobility. A transition matrix shows the probability of observations moving from, say, the first quintile in the initial period to the second quintile in the next period. For the association matrices considered in this paper, the quantiles are defined according to the capital and labor distributions instead of the income distribution at two points in time. Following Atkinson (1981), who examines transition matrices, we can test whether the degree of association between labor and capital has increased. Consider the following two association matrices $A$ and $A^*$.

$$
A = \begin{pmatrix}
    j - 1 & j \\
    \vdots & \vdots & \vdots & \vdots & \vdots \\
    \cdots & p_{i-1,j-1} & p_{i-1,j} & \cdots & \cdots \\
    \cdots & p_{i,j-1} & p_{i,j} & \cdots & \cdots \\
    \cdots & \cdots & \cdots & \cdots & \cdots \\
    i - 1 & i & i & i & i
\end{pmatrix}
$$

$$
A^* = \begin{pmatrix}
    j - 1 & j \\
    \vdots & \vdots & \vdots & \vdots & \vdots \\
    \cdots & p_{i-1,j-1} + \gamma & p_{i-1,j} - \gamma & \cdots & \cdots \\
    \cdots & p_{i,j-1} - \gamma & p_{i,j} + \gamma & \cdots & \cdots \\
    \cdots & \cdots & \cdots & \cdots & \cdots \\
    i - 1 & i & i & i & i
\end{pmatrix}
$$

where $i$ and $j$ are particular quantile groups (of labor and capital), $p_{i,j}$ is the frequency in the association matrix, and $\gamma > 0$. $A^*$ is obtained from $A$ by a correlation-increasing (or “diagonalizing”) switch, which adds $\gamma$ to the diagonal elements and subtracts it from the off-diagonal elements. This switch increases the weight on the diagonal, such that $A^*$ exhibits a stronger association between labor and capital but leaves the marginal distributions unchanged (also see Atkinson & Bourguignon, 1982).

Let $\alpha$ and $\alpha^*$ be the survival association matrices of $A$ and $A^*$, which are obtained by cumulating the association matrices from above. These are the survival copulas for a discrete distribution (Dardanoni & Lambert, 2001). Table 3 shows the survival association matrix observed for 2006, which is obtained by adding up table 2 from above. For example, 2.34% of tax units are in both the top 20% of earnings and the top 5% of capital incomes. That is, 47% of tax units in the top 5% of capitalists are also in the top 20% of earnings.

Taking the difference between $\alpha^*$ and $\alpha$ yields the following result (see appendix A.2 for intermediate steps and an illustration using the 2006 data):

$$
\alpha^* - \alpha = \begin{pmatrix}
    \cdots & \cdots & \cdots & \cdots & \cdots \\
    \cdots & 0 & 0 & 0 & \cdots \\
    0 & \gamma & 0 & \cdots & \cdots \\
    \cdots & 0 & 0 & 0 & \cdots \\
\end{pmatrix} \geq i - 1 \geq i \geq i + 1
$$

Therefore, if the difference between the survival association matrices in years $t + 1$ and $t$ is everywhere positive, labor and capital incomes have become more closely associated between those years, thus moving away from a class model, where one class is at the top of the labor distribution and the other at the top of the capital distribution.

26In the text, we show only parts of the association matrix to illustrate the effect of the correlation-increasing switch. In table A.2, we apply such a switch to the association matrix that is observed for 2006, with the corresponding survival association matrix in table A.3.

27Given our interest in the top tail of the distribution, it makes sense to consider the survival copula. Similar to the expression above, the joint survival function can be written as $H_X(l, k) = C_X(F(l), G(k))$, where $C_X$ is the survival copula, and $F(l) = 1 - F(l)$ and $G(k) = 1 - G(k)$ are the survival distributions (or complementary cumulative distribution functions) (Nelsen, 2006).

28Decancq (2014) derives dominance criteria for continuous copula functions that are equivalent to the discrete case considered here.
first-order dominance, which will be sufficient for this paper. To go beyond first-order dominance, one would need to place additional restrictions on the social welfare functions, effectively giving a different weight to the association in different parts of the distribution (Atkinson, 1981; Aaberge, 2009; Aaberge et al., 2018).

A. Results

We begin by examining the long-run evolution of some statistics from the association matrix before testing for first-order dominance for selected years. Figure 3 shows several conditional probabilities that are obtained from the survival association matrix. The figure exhibits a distinct U-shape over this forty-year period, with a decline during the initial twenty years and a rise in the twenty years leading up to 2006. For instance, in the 1960s, tax units that were among the top 1% earners were among the top quintile of capitalists with an 80% probability. This number fell to less than 60% by the early 1980s, before again rising to 80% in the 2000s. In other words, the first twenty years showed a declining association between labor and capital incomes, which has largely been reversed now. These estimates also suggest a high degree of association: if the top 1% of earners had randomly been assigned capital incomes, 20% of them would be among the richest quintile of capitalists compared to the observed 80%.

The U-shaped pattern is also found for the top 5% and the top 0.5% (figures A.4 and A.5), as well as the alternative income definitions (figure A.8).

For some parts of the conditional distribution, the rank-based measure also confirms the asymmetry that we have found in earlier results. Among the top quintile of either distribution, top labor earners are more likely to also be among the top capital incomes, compared with top capital income recipients being at the top of the labor distribution (figure 3). Of the top 1% of capitalists, only around two-thirds are within the top quintile of earners, compared with more than 80% of the top 1% earners being in the top quintile of capitalists. From looking at this part of the distribution, it thus seems that the share of the working rich with some capital income is increasing, while rentiers are less common but have not disappeared. However, when moving further up the distribution (essentially within the top 5%), the asymmetry is no longer present. In figure 3, the results for the top 5% (as opposed to the top quintile) are more similar across labor and capital and cross in recent years. Therefore, a tax unit that is at the top of either capital or labor incomes faces a high probability of being at the top of the other distribution. The increasing
Table 4.—Difference in Survival Association Matrices (in Percentage Points)

(a) 1985 Compared with 1966

|        | Capital |        |        |        |        |        |        |
|--------|---------|--------|--------|--------|--------|--------|--------|
| Labor  | Top 50% | 0.767  | −1.601 | −1.148 | −0.987 | −1.057 | −0.250 | −0.100 |
|        | Top 40% | 1.155  | −1.188 | −1.078 | −0.911 | −1.020 | −0.267 | −0.105 |
|        | Top 20% | 0.731  | −0.789 | −1.027 | −0.816 | −0.828 | −0.297 | −0.123 |
|        | Top 10% | 0.353  | −0.422 | −0.874 | −0.766 | −0.717 | −0.292 | −0.124 |
|        | Top 5%  | 0.090  | −0.300 | −0.719 | −0.639 | −0.592 | −0.271 | −0.113 |
| Top 1% | −0.013  | −0.062 | −0.265 | −0.279 | −0.268 | −0.165 | −0.094 |
| Top 0.5% | −0.007 | −0.020 | −0.124 | −0.157 | −0.156 | −0.100 | −0.066 |

(b) 2006 Compared with 1988

|        | Capital |        |        |        |        |        |        |
|--------|---------|--------|--------|--------|--------|--------|--------|
| Labor  | Top 50% | 1.201  | 1.548  | 1.905  | 1.628  | 1.126  | 0.235  | 0.115  |
|        | Top 40% | 1.328  | 1.659  | 1.958  | 1.642  | 1.084  | 0.247  | 0.115  |
|        | Top 20% | 0.816  | 1.182  | 1.601  | 1.355  | 0.861  | 0.246  | 0.119  |
|        | Top 10% | 0.404  | 0.702  | 1.216  | 1.066  | 0.700  | 0.248  | 0.118  |
|        | Top 5%  | 0.078  | 0.319  | 0.750  | 0.725  | 0.515  | 0.226  | 0.112  |
| Top 1% | 0.019   | 0.040  | 0.155  | 0.200  | 0.169  | 0.116  | 0.084  |
| Top 0.5% | 0.004 | 0.010  | 0.057  | 0.091  | 0.084  | 0.064  | 0.057  |

(c) 2006 Compared with 1966

|        | Capital |        |        |        |        |        |        |
|--------|---------|--------|--------|--------|--------|--------|--------|
| Labor  | Top 50% | 2.445  | 0.565  | 1.513  | 1.203  | 0.506  | 0.132  | 0.088  |
|        | Top 40% | 2.979  | 1.063  | 1.652  | 1.288  | 0.506  | 0.131  | 0.091  |
|        | Top 20% | 2.068  | 0.975  | 1.185  | 0.929  | 0.413  | 0.116  | 0.085  |
|        | Top 10% | 0.941  | 0.580  | 0.701  | 0.614  | 0.300  | 0.117  | 0.086  |
|        | Top 5%  | 0.331  | 0.263  | 0.370  | 0.331  | 0.181  | 0.108  | 0.092  |
| Top 1% | 0.019   | 0.016  | 0.040  | 0.048  | 0.027  | 0.051  | 0.071  |
| Top 0.5% | 0.006 | 0.011  | 0.014  | 0.012  | 0.011  | 0.031  | 0.051  |

Panel A: Bold cells: Value of survival association matrix is lower in final year (e.g., 1985) than initial year (e.g., 1966). Panel B: Share of tax units in the top 0.5% of both labor and capital increased by 0.057 percentage points between 1988 and 2006.

We test for first-order dominance between 1966 and 1985, as well as between 1988 and 2006, capturing the two twenty-year periods that we have just described but leaving out the years when TRA86 was introduced to avoid conflating changes in tax reporting with real changes in the association between capital and labor incomes. 30 In the appendix, we show that the association result is robust to choosing 1990 to 2006 (table A.1a) or 1992 to 2006 (table A.1b).31

The differences between the survival association matrices are shown in table 4, where we have highlighted the negative cells. We do not find dominance between 1985 and 1966, since there are both positive and negative differences between the two survival association matrices (panel a). At the top of the distribution, 1966 appears to dominate 1985, suggesting a fall in the association, consistent with what we observed above. The period between 1988 and 2006 presents a stark contrast of increasing association (panel b); the survival association matrix in 2006 lies everywhere above the 1988 matrix. In other words, over the same period, as the top 1% income share doubled, tax units increasingly occupied similar positions in terms of earnings and capital income. These dominance results are robust to using alternative income definitions (see appendix A.5).

30 We have chosen 1985 instead of 1986 because 1986 may be affected by the anticipation of TRA86, which was announced in 1986 and became effective in 1987. The second period starts in 1988 after the implementation of TRA86. When analyzing the effect of TRA86, Auten and Splinter (2019) document a strong increase in the top 1% share between 1986 and 1988, followed by a flattening for a few years, suggesting that a large part of the impacts of TRA86 may have been realized by that date.

31 More generally, the results are robust to choosing any year between 1985 and 1995 as the starting point for the second interval. For years between 1996 and 2004, we still find dominance at the top (the top 40% or top 20%) but not the entire distribution.
Viewed over the entire period from 1966 to 2006 (panel c), the results also point toward increasing association, although the differences are smaller. Given that TRA86 may have affected the association (see the discussion in section I), the results over the entire period are difficult to interpret and we focus on the post-TRA86 period (panel b). Furthermore, when using the alternative split for business income, we do not find dominance between 1966 and 2006 (table A.13).  

V. Conclusion

This paper has studied the association between capital and labor incomes at the top of the U.S. distribution using tax return data. This helps to understand the driving forces behind the rise in the top 1% income share that has been documented (Piketty & Saez, 2003). We find that capital and labor incomes have become more closely associated between 1988 and 2006, such that top capitalists and top earners are increasingly the same people. This rising association has contributed to the well-known increase in the top 1% income share, exacerbating the effects of rising inequality within capital incomes and earnings. In contrast, the twenty years leading up to 1985 saw a tendency toward a declining association by some measures, thus resulting in a U-shaped pattern. Our conclusions are robust to alternative treatments of negative incomes and capital gains, including estimates for tax units that do not file a return, scaling-up capital incomes to match corporate retained earnings, and how profits from closely held businesses are split between capital and labor.

The association is asymmetric in some parts of the distribution, as a top earner is almost guaranteed to also be among the richest fifth of capitalists, while a sizable share of top capitalists fall into the bottom four-fifths of earnings. The association is more symmetric at the very top of the distribution, such as the top 5%, which is consistent with mixed income from closely held businesses becoming increasingly important at the very top. This type of business income has become more closely associated with total income, driven by the increasing role of partnerships and S-corporations. We find that even earnings without business incomes are increasingly aligned with total income, confirming that the working rich are becoming increasingly important.

The reversal from declining to increasing association coincided with a strong fall in the top marginal income tax rate in the United States (dashed line in figure 3). The top marginal rate declined from 91% in the early 1960s to 28% in 1986 and remained below 40% for the rest of the period. Lower taxes at the top raise the reward to bargaining more aggressively for higher pay and therefore may explain the rapid rise in (gross) salaries at the top, which account for a large share of the increase in top income shares (Bakija, Cole, & Heim, 2012; Alvaredo et al., 2013; Piketty et al., 2014).  

32When negatives are replaced as 0, we also find a small negative value at the bottom (table A.8).

33The decline in top marginal tax rates is not the only possible explanation for these patterns (also see Alvaredo et al., 2013). Other explanations include a superstar theory together with a globalized economy (Atkinson, 2008), the spread of performance-based pay (Lemieux, MacLeod, & Parent, 2009), and the role of the financial industry (Kaplan & Rauh, 2010; Philippep & Reshef, 2012).

34Saez and Zucman (2016) discuss the effect of increasing top incomes and high savings rates for wealth inequality, which is closely related to the distribution of capital incomes. Kaymak and Poschke (2016) present a formal model where a decline in income tax progressivity leads to an increase in wealth inequality. The increasing association could also be explained by richer tax units achieving higher returns on their capital income. Saez and Zucman (2016) find no evidence of returns increasing with wealth using a range of data sources.

35Saez and Zucman (2016) find high and increasing saving rates for the top 1%. Estimates for the persistence of the top 1% differ somewhat in the literature. Using data from Social Security records that capture only earnings, Kopczuk, Saez, and Song (2010) show that around two-thirds of the top 1% remain after three years, with little change since the late 1970s. Using tax record data, a broader definition of income, and a focus on the 2000s, Auten, Gee, and Turner (2013) find that only around half of the top 1% remain after three years. However, when considering mobility over a broader range of the distribution, the results point toward considerable persistence at the top. Of taxpayers who were in the top 1% in 1987, three-quarters were in the top 10% in 2007 (Auten et al., 2013). Furthermore, the trend in the top 1% share based on annual data is very similar to the top 1% share that uses earnings averaged over five years (Kopczuk et al., 2010).

36The average age at the top of the income distribution has increased since the 2000s, after having fallen for twenty years (Piketty et al., 2018).

References

Aaberge, R., “Ranking Intersecting Lorenz curves,” Social Choice and Welfare 33 (2009), 235–259. 10.1007/s00355-008-0354-4
Aaberge, R., A. B. Atkinson, and S. König, “From Classes to Copulas: Wages, Capital and Top Incomes,” Journal of Economic Inequality 16 (2018), 295–320. 10.1007/s10888-018-9368-x
Alvaredo, F., A. B. Atkinson, T. Piketty, and E. Saez, “The Top 1 Percent in International and Historical Perspective,” Journal of Economic Perspectives 27 (2013), 3–20. 10.1257/jep.27.3.3
Atkinson, A. B., “The Measurement of Economic Mobility,” in P. J. Egelshoven and L. J. van Gemerden, eds., Inkomensverdeling en openbare financiën (Utrecht/Antwerpen: Het Spectrum, 1981).

— “Measuring Top Incomes: Methodological Issues,” in A. B. Atkinson and T. Piketty, eds., Top Incomes over the Twentieth Century:
—— “High-Income Families and the Tax Changes of the 1980s: The Anatomy of Behavioral Response,” in M. S. Feldstein and J. M. Poterba, eds., *Empirical Foundations of Household Taxation* (Chicago: University of Chicago Press, 1996). 10.3386/w5218

Smith M., D. Yagan, O. Zidar, and E. Zwick, “Capitalists in the Twenty-First Century,” *Quarterly Journal of Economics* 134 (2019), 1675–1745. 10.3386/w25442

Winglee M., R. Valliant, J. Clark, Y. Lim, M. Weber, and M. Strudler, “Assessing Disclosure Protection for a SOI Public Use File,” *Proceedings of the American Statistical Association* (2002).

Wolff, E. N. and A. Zacharias, “Household Wealth and the Measurement of Economic Well-Being in the United States,” *Journal of Economic Inequality* 7 (2009), 83–115. 10.1007/s10888-007-9068-6