Income Distribution, Econophysics and Piketty

Anwar Shaikh

To cite this article: Anwar Shaikh (2017) Income Distribution, Econophysics and Piketty, Review of Political Economy, 29:1, 18-29, DOI: 10.1080/09538259.2016.1205295

To link to this article: http://dx.doi.org/10.1080/09538259.2016.1205295

Published online: 19 Jul 2016.
Income Distribution, Econophysics and Piketty

Anwar Shaikh

Department of Economics, New School for Social Research, New York, NY, USA

ABSTRACT
Piketty, Atkinson and Saez have put the analysis of income distribution back on center stage. The distinction between property income and labor income plays a central role in this framework. Property income derives from the rate of return on stocks of income-earning wealth and is more unequally distributed than labor income. Piketty argues that, because the rate of return ($r$) is generally greater than the rate of growth of the economy ($g$), property income tends to grow more rapidly than labor income, so that rising income inequality is an intrinsic tendency of capitalism despite interruptions due to world wars and great depressions. This article argues the exact opposite. The rise of unions and the welfare state were the fruits of long-term historical gains made by labor, and the postwar constraints on real and financial capital arose in sensible reaction to the Great Depression. The ‘neoliberal’ era beginning in the 1980s significantly rolled back all of these. The article uses the econophysics two-class argument of Yakovenko to show that we can explain the empirical degree of inequality using two factors alone: the profit share and the degree of financialization of income. The rise of inequality in the neoliberal era then derives from a reduction in the wage share (rise in the profit share) in the face of assaults on labor and the welfare state, and a sharp increase in the financialization of incomes as financial controls are weakened. These are inherently socio-political outcomes, and what was lost can be regained. Hence, there is no inevitable return to Piketty’s ‘patrimonial capitalism’.

1. Introduction

Income distribution is very much back in vogue. Piketty, Atkinson and Saez put 15 years of dedicated effort into their World Top Incomes Database (http://topincomes.g-mond.parischoolofeconomics.eu/), and their path-breaking work has already changed the way we see the world and helped fuel the resurgence of interest in inequality. This type of effort harks back to the tradition of Smith, Ricardo, Marx, Keynes and Kalecki, among others, in which economic theorizing is grounded in the observed patterns of actual economies. The distinction between property income and labor income plays a central role in all of these frameworks. Property income derives from the rate of return on stocks of income-earning wealth. Piketty argues that the rate of return ($r$) is generally greater than the rate of growth of the economy ($g$).
greater than the rate of growth of the economy \( (g) \), so that property income tends to grow more rapidly than labor income. Given that the former is much more unequally distributed than the latter, a relative rise in property income worsens overall income inequality. Piketty claims this is an inherent tendency in capitalism itself, interrupted only by countervailing factors such as wars, revolutions and depressions (2014, pp. 25–27, 242–246).

According to Piketty, one cannot hope to glean much about the inherent tendencies of capitalism through an analysis of the postwar era because the two world wars and the Great Depression that preceded it make this period aberrational (ibid., pp. 216–218). In this article I will argue the exact opposite. The first part of the postwar period, the so-called Golden Age for Labor, reflects the fruits of hard-fought historical gains made by labor in the context of constraints on real and financial capital put into place after the Great Depression.\(^1\) The subsequent 'neoliberal' period beginning in the 1980s significantly rolled back labor’s gains, weakened the welfare state and loosened the controls on capital. I will show that we can explain the corresponding rise of inequality in the neoliberal era using two factors: the reduction in the wage share (hence the rise in the profit share) and the increasing degree to which the dividends, rents, interest and royalties flowing from profit are ‘financialized’ through an ever-mounting pyramid of assets. Both the balance of power between labor and capital and the financialization of incomes are socio-political outcomes. What was lost in the neoliberal era can be regained. Hence I dispute Piketty’s claim that the rise of inequality in this era is an inevitable return to the norms of ‘patrimonial capitalism’.

2. On the distribution of labor and property incomes

The overall distribution of income, comprised of labor and property incomes, has recently become an object of investigation in the econophysics literature (Rosser 2006; Yakovenko and Rosser 2009). Yakovenko and his co-authors provide substantial empirical evidence that the bottom 97–99 percent of the distribution of income is exponential while the top 1–3 percent is Pareto. They further argue that the bottom part is essentially labor income while the top is largely composed of property income (Dragulescu and Yakovenko 2002; Silva and Yakovenko 2004; Yakovenko 2007). In an earlier paper, my co-authors and I showed that the distribution of directly measured labor income is indeed exponential, and that this feature even obtains for male, female, white and African-American labor incomes separately (Shaikh, Papanikolaou, and Weiner 2014). And of course Pareto himself showed in 1897 that property incomes follow the power law which now bears his name (Pareto 1964; Rosser 2008).

Yakovenko and his co-authors provide an ingenious method of combining the two classes of distribution from which they derive an approximation to the Gini coefficient of the overall income distribution. They first note that an exponential distribution has a fixed Gini coefficient \( G^* \), equal to 0.50 for individual incomes and 0.375 for two-person family incomes derived from exponentially distributed individual incomes (Dragulescu and Yakovenko 2001). They then point out that, on a Lorenz curve, the small Pareto
section signifying the top 1–3 percent of the total population can be usefully represented by a fixed term that kicks in at the top level in order to account for the Pareto tail (Silva and Yakovenko 2004). This takes advantage of the fact that the population percentage at the higher end is very small but its income fraction \( f \) is nonetheless significant. On this basis, they show that the overall Gini coefficient can be well approximated by a linear relation between the Gini of the exponential portion of the distribution \( G' = 0.50 \) or \( 0.375 \) and a term that depends solely on the top income share in total IRS income \( f \).

The relation displayed in Equation 1 was shown to provide an excellent fit with the actual data.\(^2\) I will call this approach the econophysics two-class (EPTC) theory of income distribution.

\[
G \approx G' + f(1 - G') \tag{1}
\]

3. Empirical evidence

3.1 Exponential and pareto components of income distribution

The EPTC claim is that the overall distribution of income separates into two parts: an exponential distribution for the bottom 97–99 percent of the population and a Pareto distribution for the top 1–3 percent. For incomes \( y \) an exponential distribution function has a probability distribution \( P(y) = (\overline{y}) e^{-(y/\overline{y})} \), where the parameter \( \overline{y} \) represents the mean income of the theoretical exponential distribution. The cumulative probability for incomes \( y \) is \( C(y) = e^{-(y/\overline{y})} \), which is parameter-free in normalized income \( y/\overline{y} \), so that \( \ln C(y) = -(y/\overline{y}) \) is linear in the \emph{level} of income. For parameters \( a \) and \( k \), the Pareto distribution has a cumulative probability from below of \( C(y) = 1 - (k/y)^a \) and hence a cumulative probability from above of \( C(y) = 1 - C(y) = (k/y)^a \) so that \( \ln C(y) = \alpha + \beta \ln y \) is linear in the \emph{log} of income, where \( \alpha = a \ln k \) and \( \beta = -a \).

Figure 1 displays actual 2011 US Internal Revenue Service (IRS) individual income data\(^3\) on a log-linear scale for incomes up to \( \$200,000 \) (the bottom 97 percent of the population), and we see that these are indeed close to the theoretical exponential distribution represented by the dashed line. At the other end, as noted, a Pareto distribution has the property that the natural log of its cumulative probability from above is linear with respect to the natural log of bin size, and Figure 2 shows that incomes above \( \$200,000 \) do indeed conform to a linear path on a log–log scale. The EPTC group shows that the same patterns obtain in \emph{all} years in the US from 1983–2008 (extended to 2013 by José Coronado), and in Japan and the UK (Yakovenko 2007, pp. 13–15; Banerjee and Yakovenko 2010).

\(^2\)The relation in Equation 1 is a general formula for both individual and family incomes that was provided by Victor Yakovenko in a private communication. If individual incomes follow an exponential distribution \( G' = 0.5 \), the general formula reduces to \( G \approx 0.5 + 0.5f \). For two-person family incomes, if we assume that the underlying distribution of individual incomes is exponential, the resulting Gini coefficient is \( G' = 0.375 \) (Silva and Yakovenko 2004, p. 5), in which case Equation 1 implies \( G \approx 0.375 + 0.625f \). I thank an anonymous referee for the query which led me to contact Yakovenko for the general formula in Equation 1.

\(^3\)The data in Figures 1 and 2 is from the US Internal Revenue Service (IRS), Table 1.4: All Returns: Sources of Income, Adjustments, and Tax Items, by Size of Adjusted Gross Income, Tax Year 2011. The IRS data in thousands of US dollars is based on samples and comes pre-binned with bins of varying width. Each bin was represented by its midpoint, the corresponding number of returns was expressed as frequencies that were cumulated to get the cumulative probability from below \( C' \), and the cumulative probability from above was calculated as \( C = 1 - C' \).
3.2 Bottom incomes as labor incomes

The second aspect of the EPTC argument is the claim that the log-linear part of income distribution is primarily composed of labor income while the log-log partition essentially represents property income. Yakovenko shows that the portion of income in the top distribution that is in excess of an exponential component swells and shrinks with the

**Figure 1.** Individual income distribution below $200,000, cumulative probability from above (US 2011 IRS data: log-linear scale).

**Figure 2.** Individual income distribution above $200,000, cumulative probability from above (US 2011 IRS data: Log–log scale).
movements of the stock market (Silva and Yakovenko 2004, pp. 1, 3; Banerjee and Yakovenko 2010, p. 12). But the claim that the exponential distribution itself is dominated by labor income cannot be directly addressed within IRS income data due to lack of information on distribution by source of income. The March supplement of the US Census Bureau Current Population Survey (CPS) providing individual wage and salary income data does provide enough information to ascertain the distribution of labor income from 1996 to 2008, but not for high-end property income because the CPS data is top-coded.4 Even within labor (wage and salary) incomes in CPS data, salaries of top executives include an unidentified amount of property-type income in the form of performance bonuses (disbursements from net operating surplus) and exercised stock options (capital gains), both of which raise incomes in the top bins and hence raise the observed points above the theoretically expected ones at the tail end of the labor income distribution (Seskin and Parker 1998, p. M-8). This is just what we find in Figure 3, which compares the actual distributions in US CPS wage and salary incomes on log-linear scales over two sets of years, 1996–2002 and 2003–2008, respectively5 (with different symbols for each year within a given set). The corresponding theoretical distribution is depicted by a solid line. Inset in each chart is the estimated exponential mean income (\(\overline{y}\)) in each year. The correspondence between the data and the theory is quite good, except for a smattering of points at the upper end of the spectrum. The latter effect is partly explained by the fact that the degree of deviation from the predicted exponential distribution seems to be multiplicative, so that it is absolutely larger at higher values of income. In addition, some of the deviation may be due to random sample fluctuations arising from the low probabilities associated with the upper points (Shapiro 1995, p. 207), but more likely due to the greater role of property-type income near the top-code limit. Despite this, actual Gini coefficients are close to the predicted value of 0.50 for individual labor incomes characterized by an exponential distribution (Shaikh, Papanikolaou, and Weiner 2014, p. 55 and Table 51, p. 58).

4In an effort to protect the identities of individuals, top-coding replaces the actual earnings of individuals in higher brackets (above $200,000 in 2003–2008) ‘with the mean earnings for individuals with similar characteristics’ (https://cps.ipums.org/cps/topcodes_tables.shtml#1996top). This distorts the actual distribution in the top-coded regions and reduces inequality by assigning the same income to several people in a given range.

5The data in a given CPS publication refers to the incomes in the prior year. We retain the publication year as the reference.
3.3 Gini coefficient as a function of top incomes

The third aspect of the EPTC argument is the claim that the overall Gini coefficient is solely a function of top income as in Equation 1. Figure 4 displays the excellent correlation between the overall Gini coefficient and the income share of the Pareto section \((f)\) in the IRS individual income data for 1996–2013 (Banerjee and Yakovenko 2010, Table 1, p. 12, updated to 2008–2013 by Jose Coronado). As previously noted, here the top incomes are hypothesized to be composed largely of property income since the lack of pertinent information in IRS and the top-coding of CPS data prevents us from directly connecting the two. Notice how closely the fitted relation \(G \approx 0.486 + 0.555f\) shown in the figure approximates the theoretical one of \(G \approx 0.5 + 0.5f\) previously developed in footnote 2. It should be noted that power law applies to the top 1–3 percent of the population—consistent with the general finding that it is the rise in the income of this share of the population which accounts for most of the rise in inequality.

3.4 Gini coefficient and property income, 1947–2011

The balance between labor and capital in the US economy changed dramatically between the first half of the postwar period and the neoliberal era which began in the 1980s. In order to assess the impact of this shift, we need data on profit and property income shares over the whole postwar period. Neither IRS nor CPS income distribution data go back far enough. However, we do have census data on the Gini coefficient and aggregated IRS and Department of Labor data on income by source over the whole postwar period. It should be said that census personal income data is not directly comparable to IRS data on individual income for three reasons: first, the appropriate theoretical Gini coefficient for personal labor income is \(G' = 0.375\), rather than \(G' = 0.50\) as in IRS individual income data; second, unlike IRS individual income, Census personal income is self-reported and...
also excludes capital gains which in itself would create a downward bias in the level of the census Gini coefficient; third, we now explicitly test the hypothesis that the upper part of the income distribution is largely composed of property income by substituting the actual property income share ($\sigma_{PP}$) in place of the top income share ($f$) that the EPTC approach uses as a proxy for the former. With this, the relation developed in Equation 1 becomes

$$G \approx 0.375 + 0.625\sigma_{PP}$$

(2)

Figure 5 displays the census Gini coefficients for 1947–2011. The sharp rise in income inequality in the neoliberal era is immediately visible. The IRS provides aggregated data on income by source in tax returns of individuals for 1921–2012. The Property Income Share with Capital Gains was calculated from IRS Adjusted Gross Income (AGI) as (AGI—Labor Income)/AGI. Figure 6 compares this to Piketty’s capital share (Capital Income/Value Added). Both measures drift downward in the first half of the postwar period from 1947–
1979, and then rise substantially, most of all the IRS measure with capital gains. As we can see, capital gains in the neoliberal era increase both the level and volatility of property incomes.

3.5 Explaining the rise in US income inequality

We can now complete a classical explanation of income inequality. It is useful at this point to decompose the property income share into three parts: the ratio of net operating surplus to value added ($\sigma_P$), which reflects the struggle between labor and capital over wages and profits; the financialization ratio of property income to net surplus ($F$), which reflects the degree to which incomes from wages and salaries, rent, dividends, interest, profits and royalties are amplified through the pyramid of financial assets$^{11}$; and the ratio of total individual income to total value added reflecting the degree to which aggregate domestic income is translated into aggregate personal income ($\varphi$)$^{12}$. See Equation 3:

$$\sigma_P = \frac{\text{Property Income}}{\text{Personal Income}}$$

$$= \left( \frac{\text{Net Surplus/Value Added}}{\text{Personal Income/Value Added}} \right) \left( \frac{\text{Property Income}}{\text{Net Surplus}} \right) = \frac{\sigma_P F}{\varphi} \quad (3)$$

Figure 7 shows that the ratio of personal incomes to value added is stable at about 67 percent over the whole postwar period. Thus the rise in the property income share in the neoliberal era can be attributed to two factors: a 20 percent relative rise in net surplus share

---

$^{11}$Various forms of income give rise to borrowing and lending, and these financial assets in turn give rise to derivative flows and assets, and derivatives of these, pyramiding upward.

$^{12}$US Net Domestic Income (Value Added) is Gross Domestic Income (line 1) minus Consumption of Fixed Capital (line 21) in National Income and Product Accounts (NIPA) Table 1.10. Gross Domestic Income by Type of Income (http://www.bea.gov/iTable/iTable.cfm?ReqID=9&step=1#reqid=9&step=3&isuri=1&903=51) and Net Operating Surplus is line 10 = Net Domestic Income – Compensation of Employees Paid (line 2) – Net Taxes on Production and Imports (line 7 minus line 8).
from 1979–2012 (from 25 to 30 percent) and a whopping 90 percent relative rise from 1979 to the onset of the 2007 global crisis in the financialization ratio (from 45 to 85 percent). Taken together, by the onset of the crisis these two factors almost double the share of property income in total personal income (from 17 to 32 percent).

The remaining step is to test the EPTC hypothesis of Equation 2; that is, the overall Gini coefficient is essentially a function of the property income share. This is demonstrated in Figure 8. Despite the fact that census data itself excludes capital gains so that the latter had to be estimated from IRS data, the fitted relation $G = 0.249 + 0.646\sigma_{PP}$ shown in the figure is quite close to the theoretical relation $G \approx 0.375 + 0.625\sigma_{PP}$ in Equation 2.
4. Piketty versus classical economics

Piketty’s book has three logical parts. First, the presentation of its empirical findings on the distribution of income and wealth, in which a central claim is that because the profit rate \( r \) tends to be greater than the growth rate \( g \) capitalism has an inherent tendency to produce increasing wealth inequality, interrupted only by countervailing shocks such as world wars, revolutions and depressions (Piketty 2014, pp. 25–27, 242–246). When Piketty looks for an explanation of the profit rate and of wages, he tends to rely quite a bit on neoclassical economic theory, including the notion of an aggregate production function and its associated marginal productivity theory. In reference to the rate of profit, he says that ‘too much capital kills the return on capital [because] the marginal productivity of capital decreases as the stock of capital increases’ (p. 215). In addition, the parameters of the aggregate production function in conjunction with technical change are said to determine the elasticity of substitution between capital and labor.\(^{13}\) Felipe and McCombie (2013) provide a definitive critique of this and other uses of so-called aggregate production functions.

Piketty’s central focus is on the final distribution of personal income achieved. This is a combination of labor income, including wages and salaries, as well as unemployment benefits and transfers, bonuses and stock options, etc. (Piketty 2014, pp. 477, note 479 p. 602); and property income, what Piketty calls income from capital, which includes rent, interest, profits, capital gains, royalties and other income from ownership of land, real estate, financial instruments, etc. (p. 18). All major economic traditions recognize that the distinction between labor and property income is premised on three distinct sets of forces. The direct relation between workers and their employers gives rise to the primary division of the aggregate net product into wages (and salaries) and business net operating surplus. After taxes, the further division of the income flows into dividends, rents, interest and royalties creates a flow of property incomes. Upon this rests a tertiary circulation consisting of deductions due to taxes on labor and property incomes, and additions due to state payments of unemployment insurance and transfer payments to recipients of labor and property income. The overall distribution of personal incomes is the combined effect of primary, secondary and tertiary flows.

As noted, Piketty tends to rely on marginal productivity theory for an explanation of the primary division between wages/salaries and operating surplus. But when he comes to the shape of the final distribution of income, the sum total of the effects of the primary, secondary and tertiary distributions, he refers only to the finding of Vilfredo Pareto that top income and wealth distributions tend to follow a power law (the ‘Pareto Law’). Piketty observes that in practice the Pareto distribution only works for the upper tails of distributions, and that in any case the fitted Pareto coefficients vary over time and space. On this basis, he rejects the notion of any particular overall inequality

\(^{13}\)In contradictory fashion, while he associates the rate of profit with the peacefully attained marginal product of capital, Piketty says the opposite with regard to labor and capital shares: ‘the question of what share of output should go to wages and what share should go to profits … has always been at the heart of distributional conflict’ (2014, p. 39, emphasis added). In a similar vein, he evinces some skepticism regarding the notion that the marginal product of labor determines the wage, at least in the upper reaches of the salary income: ‘Is it really true that the explosion of compensation at the very top is related to objective differences in productivity rather than corporate culture in the US and UK vs much less in Europe and Japan?’ (pp. 330–333).
distribution (Piketty 2014, pp. 364–368). Indeed, Piketty’s goal is to show that changes in the distribution of income can be explained by his general hypothesis, \( r > g \).

Given his conclusion that capitalist inequality is structural, and is likely to revert to its earlier, much more unequal, levels, he is forced to turn to the possibility that ‘democracy can regain control of capitalism’ with appropriate policies, principally a ‘progressive global tax on capital’ that can be used to fund a social state for the twenty-first century so as to ‘regulate today’s global patrimonial capitalism justly as well as efficiently’ (pp. 1, 471). He admits that this is a ‘utopian ideal’ but hopes that it could spark regional or continental moves in that direction (p. 471).

From a classical perspective, three further points about Piketty’s argument require consideration. First, in the classical tradition the rate of profit is generally taken to be greater than the rate of growth. At the most abstract level, aggregate value added \( Y = W + \text{NOS} = C + I \), where \( W \) = the wage bill, \( \text{NOS} \) = net operating surplus, \( C \) = consumption and \( I \) = investment. If workers consume all of their wage bill, total consumption of workers and capitalists together will be greater than total workers’ consumption and hence greater than the wage bill (\( C > W \)). It follows that \( \text{NOS} > I \), so that \( r = \text{NOS}/K > g_K = I/K \), where \( g_K \) = the growth rate of capital. But since output will fluctuate around capacity in the long run, if the capital/capacity ratio is constant the rate of growth of output (\( g \)) will satisfy \( r > g \). And if the capital/capacity ratio is rising, the rate of growth of capital will exceed the rate of growth of output so \( r > g_K > g \). It follows that \( r < g \) only if the capital/capacity ratio is falling sufficiently rapidly—i.e. the ‘productivity of capital’ is rising sufficiently rapidly. This is a technological hypothesis that few would argue is intrinsic to capitalism (Shaikh 2016, Chapter 6, Section VIII).

Second, Piketty’s own measure of capital is an inconsistent measure of real and financial assets. He defines ‘capital’ as including the value of land and the values of financial assets because he is concerned with all the sources of property income. But land is valued by the share of the surplus that its owners receive as rent, and financial assets are valued by the portions of dividends, interests, royalties and even rents to which they lay claim. With only the net surplus in the numerator of the rate of profits it would be double-counting to include the values of land and financial assets derived from the surplus in the denominator (Harcourt 2014). Conversely, if land and financial assets are included in the denominator, one must also include the property income flows from these assets in the numerator. But then if profit rates are equalized across real and financial assets, the general rate of profits will be the same as the traditional rate of profits. Piketty’s hybrid definition of the rate of profits is invalid.

Third, since Piketty’s hybrid rate of profits has fixed capital plus land and financial assets in the denominator, it will tend to fall during booms because these inflate the values of land and financial assets. Only a careless reading of both Piketty and the classical texts could take this as evidence of a ‘falling rate of profits’ in the traditional sense. Conversely, his measure of profitability will tend to rise during recessions and depressions.

Like Piketty, the classical tradition is also grounded in a structural analysis of actual capitalism. As in Piketty, wages and property incomes have distinctly different distributions which combine to produce the overall distribution of pre-tax income, and a larger property income share will increase the degree of pre-tax income inequality. But in the classical argument, the property income share is itself determined by the ongoing struggle between capital and labor over the determination of wages and productivity,
and by the extent to which financial capital is given room to exercise its intrinsic tendency for bubbles and troubles. Hence the directions of shifts are not ineluctable. They are in fact political outcomes in which ground gained may be lost but may also be gained back and even extended—within limits shaped by the profitability of capital. Then both pre- and post-tax distributions are in contention.

**Disclosure statement**

No potential conflict of interest was reported by the author.

**References**


