Information Technology and the Rise of the Power Law Economy

Completed Research Paper

Erik Brynjolfsson
Massachusetts Institute of Technology
Sloan School of Management
100 Main Street
Cambridge, MA 02142
erikb@mit.edu

Guillaume Saint-Jacques
Massachusetts Institute of Technology
Sloan School of Management
100 Main Street
Cambridge, MA 02142
gsaintja@mit.edu

Abstract

We show that the dramatically increasing share of income going to top earners is consistent with the rise of the "power law economy" and argue this may reflect networks and increased digitization. Specifically, tax data (1960-2008) show an increasing dispersion of incomes in the portion of the distribution drawn from a power law, as opposed to the long-established log-normal distribution. We present a simple model arguing that the increased role of power laws is consistent with the growth of information technology, because digitization and networks facilitate winner-take-most markets. We generate four testable hypotheses which are supported by the data. (1) Our model fits the data better than a purely lognormal distribution, (2) the increase in the variance of the log-normal portion of the distribution has slowed, consistent with slowing skill-biased technical change, (3) more individuals now select into the power law economy, (4) skewness has increased within that economy.

Keywords: Income Inequality, Digital Economy, Information Economy, Network Effect, Power Laws, Digitization, Economic Impacts

Introduction

In 1970, a person in the top 1% of the wage distribution earned between seven and eight times as much as someone in the bottom 90%. By 2008, this ratio had roughly tripled to about 22. While US has experienced a tremendous increase in income inequality over the past forty years, research on this phenomenon is at a critical crossroads. On one hand, the facts documenting the increase in US inequality are clear. On the other hand, there is not a consensus on the driving forces that caused it. Establishing the role of technology, if any, and specifically the ways that information technologies can affect the income distribution, is a key task for IS researchers.

Most of the literature on the effects of information technologies and inequality has focused on skill-biased technical change, documenting and explaining a broad increase in the relative wages of skilled workers, such as college graduates, over unskilled workers. A less well studied phenomenon, which we focus on in this paper, is the potential of technology to create superstar effects. Digital technologies can amplify the ideas, talents or luck of a small handful of innovators, vastly increasing their income as they reach a broader market. In contrast to earlier production and distribution technologies, goods and services that are digitized can be replicated at nearly zero cost, with perfect fidelity and reach global audiences almost instantaneously. This fundamentally changes the way that value is distributed and allocated, and it would be surprising if it didn't have significant effects on the income distribution.

In this paper, we argue that the shape of the income distribution is changing in specific way: using nearly 50 years of tax data from the United States Internal Revenue Service (IRS) and a new modelling approach, we show that a bigger share of individual incomes are drawn from a power law, or Pareto distribution, as opposed to the long-established log-normal distribution that historically governed incomes. We argue that the increased prevalence of Power Law distributions is consistent with the effects of the diffusion of Information Technology, because digitization and networks facilitate winner-take-most markets. We present a simple theoretical model of income distribution and of the role of information technology, and estimate it using the IRS public use tax files. We find that more and more individuals seem to participate in winner-take-most markets, and that, within these markets, competitiveness has been steadily increasing.

The remainder of this paper is organized as follows: Section 2 highlights some of the relevant literature. Section 3 introduces simple models of the key economics of information technology and its potential to produce highly skewed income distributions, and specifically power laws. Our models imply four hypotheses on the effects of IT the distribution of income and ways to empirically test them. Section 4 presents our basic tax data and shows evidence of fractal effects in income, which are consistent with the presence of power laws. Section 4 presents our empirical strategy and assesses our hypotheses more formally. Section 5 offers some discussion of alternative stories, robustness checks and extensions. Section 6 concludes with a summary and some implications.

Relevant Literature

Recent Trends in Inequality

This paper draws on three main spans of literature. First, it builds on recent research on income inequality and the role of technology. Second, it builds on the more specific span of the economics literature dedicated to understanding functional forms of the income distribution. Finally, it draws on recent research in statistics and actuarial science for tools to estimate specific mixture models of statistical distributions.

The recent increase in income inequality has been widely documented. Some of the most recent research started in France (Piketty 2001) and in the US (Piketty and Saez 2003), and was then extended to a large number of countries (Atkinson and Piketty 2010). The starting point of this literature is the Kuznets curve: the idea that inequality follows an inverted U-shape over time, in the sense that the industrial revolution brought about an increase in inequality which gradually reversed over time (Kuznets 1955). After forty more years of data, Piketty argues that the Kuznets curve has now turned up again: after decreasing around WW2, inequality has been increasing again in the United States since the 1980s. This phenomenon has accelerated in recent years: while the 1990s expansion led to a 10% increase of top 1% incomes and a 2.4% of bottom incomes, the more recent 2001-2007 expansion led to an 11% increase at the top and an increase under 1% at the bottom (Saez 2013). In other words, 75% of recent income growth went to the top 1%.

Why should anyone care about the income distribution? There are several types of arguments present in the literature: First, even if increased inequality is linked with a higher mean income, large groups of the population can be left significantly worse off. This has been the case for the median family between 1990 and 2008 (Bernstein 2010). Second, increased inequality is associated with decreased mobility (Corak 2013). This finding was coined by Krueger (2013) as the "Great Gatsby curve", and may create concern that inequality can be self-perpetuating. Recent evidence shows that household mobility as become very low (Debacker et al. 2013). Third, increased economic inequality may lead to increased political inequality, leading to a vicious circle (Acemoglu and Robinson 2012). Finally, a great deal of happiness seems to be dependent of relative levels of income and wealth rather than absolute levels, and a dislike of inequality may directly enter into the utility function for many people.

Theories of Inequality: Institutions vs Markets

There is no general consensus regarding the driving forces behind inequality. A first class of arguments emphasizes the impact of institutions and historical events. Among the proponents of institutional explanations, one can find arguments relating to recent drops in tax rates (Piketty and Saez 2003) favoring the rich, or to various types of rent-seeking (Bivens and Mishel 2013). It is also argued that past reductions in inequality were mostly accidental rather than the result of market forces. In this literature, emphasis is put on the effects of the two World Wars and on the great depression as the main driver of the large drop in inequality that was observed in the first half of the 20th century: a lot of wealth was destroyed, and wealthy individuals were disproportionately affected.

A second class of arguments emphasizes markets and technology as drivers of inequality. These "market" arguments can roughly be decomposed into two main categories: *skill-biased technical change* and *superstars*. Skill biased technical change is the idea that new technologies emerge that have much higher complementarities with skilled labor than unskilled labor, therefore increasing wages of skilled individuals and depressing the wages of others. SBTC is an important part of the labor economics literature focusing on the college premium. For a comprehensive review of this literature, see Katz and Autor (1999), Katz (2000), and Acemoglu and Autor (2012). It is worth noting that this literature typically focuses on the bottom 99%, i.e. tends to exclude top incomes from the analysis. Generally speaking, the bulk of the income distribution seems to follow a log-normal distribution. This seems to be the case over time, and in a large number of countries (Lopez and Serven 2006).

On the other hand, the literature on superstars is concerned with a much smaller number of individuals: rather than investigating the difference between skilled and unskilled labor (two very large groups), it focuses on a very small of fraction of the population which receives very high incomes. While small in numbers, this group can command a large (and growing) share of aggregate income. The literature on superstars emphasizes the effects of technology through economies of scale that allows the most talented individuals to replicate their talent across larger and larger markets (Rosen 1981). The very top of the income distribution seems to follow a Pareto distribution rather than a lognormal one. Pareto distributions seem to be present in a very large array of phenomena: City Size, frequency of words in languages, casualties in wars, cotton prices, movie profits, publications and citations by researchers, social networks, internet links (Faloutsos et al. 1999), and many more. As analyzed in the literature on the "Long Tail", sales of books and other products online can be very well described by a simple Pareto distribution (Brynjolfsson, Hu and Smith, 2003). For a list of documented phenomena following a Pareto distribution, see Andriani and McKelvey (2009).[1]

Theory: Distributional Effects of Digitization

The dominant explanation of income inequality has historically been skill-biased technical change: new technologies are used that require higher-skilled labor, thereby increasing wages of high-skill workers and

[1] It is worth mentioning that more complex statistical distributions have also been used to describe incomes: Lognormal and Pareto distribution have recently been put together in the double Pareto Lognormal distribution, which is shown to fit income data very well (Hajargasht and Griffiths 2013). However, it is interpreted in terms of proportional random shocks applied to an exponentially growing population rather than in terms of individual productivity or markets (Reed and Jorgensen 2004).

depressing wages of lower-skilled workers. Autor et al. (2008) show an increase in both inter-group (according to age, education and gender) and intra-group (residual) inequality. They document polarization, i.e. middle income workers being most negatively affected relative to those at the high and low end. They offer the explanation that computerization substitutes for routine information processing work more than non-routine cognitive work (at the high end) or non-routine manual work (at the low end). They also emphasize the importance of residual inequality: even within an industry and a job category, inequality is increasing. In an earlier paper, Autor et al. (1998) show the influence of computerization on inequality through skill upgrading: the college premium increases faster in computer-intensive industries. Bresnahan et al. (2002) also show that greater use of information technology within a firm leads to the employment of higher educated workers and greater investments in training.

However, these explanations focus on the "other 99%" (Autor 2014), and do not address the top 1%, which is a main cause of the increase in aggregate inequality. The chart below shows the evolution of the ratio of reported income of the top 1% of taxpayers to the bottom 90% of taxpayers. The increase since the 1980s is rather dramatic: where the average wage in the top 1% was roughly 10 times higher than the average wage in the bottom 90%, it is now over 20 times higher. This is evidence that understanding inequality likely requires paying special attention to the top 1% of wage earners.

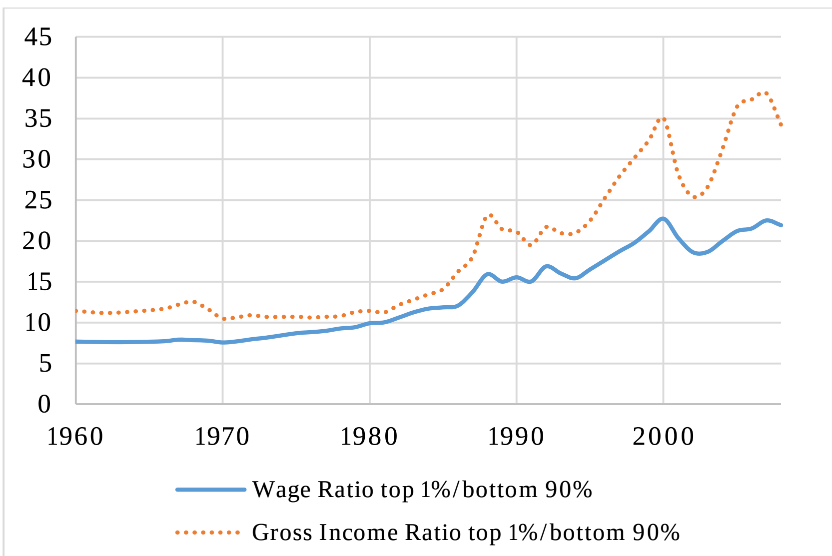


Figure 1: Ratio of average income reported by the top 1% to the bottom 90%

Recent research has argued that some people are earning a disproportionate share of income because they have become much more productive. This seems to be the case for managers (Kaplan and Rauh 2013) and CEOs (Gabaix and Landier 2008). The link between increases in managerial productivity and information technology can be established using the concept of effective size (Kim and Brynjolfsson 2009). As companies become more IT-intensive, managers can control larger and larger spans of activity. A good manager can create a lot of value, and a bad manager can destroy a lot of it.

Investigating the increase of top 1% incomes likely requires focusing on the properties of the digital revolution. The digital economy seems to have two main characteristics: first it is *digital*, allowing for production that is nearly *free*, *perfect*, and *instantaneous*. Digital goods can be replicated at almost zero marginal costs: For example, streaming music to two customers costs virtually the same as streaming to one customer only. Digital goods can be replicated perfectly, as the act of copying data does not degrade it. Further, goods can be delivered nearly instantaneously to anyone connected to the global internet. This alone can lead to large economies of scale and generate winner-take-most economies.

Second, the digital economy is fundamentally *networked*. For example, when choosing a new instant messaging app to use, or a new smartphone game to download, users may be influenced by the number of other people already using these services. The more people use the platform, the more likely they will learn of it and, in many cases, the more valuable it becomes to any individual user. As a consequence, individuals tend to preferentially attach to services that already have a large user base. Or, they may primarily choose their new apps from the top-ranked apps from their favorite app store. Fleder and Hosanagar (2009) have shown that such recommender systems can create biases towards already popular products, creating a rich-get-richer effect.

How the Industrial Economy Generates Log-normal Income Distributions

The main theory of the income-generating process that leads to a lognormal distribution posits that individuals receive a series of random shocks to their income (Gibrat 1931; Mincer 1958). Individual characteristics seem to often be roughly normally distributed. This is the case for height, IQ, gripping strength, and many more. This motivates the question: "How can one reconcile the normal distribution of abilities with a sharply skewed distribution of incomes"? (Pigou 1932)

Lognormal distributions are typically generated as the product of a large number of random variables. For example, assume that the productivity of a car mechanic is the product of a number of her individual abilities, such as attention to detail, information gathering ability, skill in operating vehicles, decision making skill, ability to establish and maintain relationships, strength, fine motor skills, and so on. The higher the number of relevant characteristics, the more the distribution of productivity in the population will resemble a lognormal. This follows straightforwardly from the central limit theorem in log space:

If $X_1, \dots X_n$ are i.i.d. individual abilities with positive finite mean and variance, and $Y = \Pi_{i=1..n} X_i^{1/n}$, then

$\log(Y) = \frac{1}{n} \sum_{i=1..N} \log(X_i)$, and by the C.L.T, log(Y) converges to a normal (i.e. Y is lognormal).

Two main implications follow: first, if a car mechanic becomes 25% stronger, his productivity will increase by 25%. But if he becomes both 25% stronger and has a 25% increase in fine motor skills, one can expect his productivity to increase by more than 50%. In other words, there are complementarities between individual abilities, so that the overall productivity effect of having high abilities is higher than the sum of marginal productivity effects of these abilities taken individually. If compensation is equal to marginal productivity (i.e. if labor markets are competitive), then this process will create a log-normal distribution of income.

Second, we can expect industrial technical change to alter the distribution of skills among workers: for example, the ability to write efficiently using word processing might have higher population variance than the ability to write with a pen. In other words, we might expect industrial technical change to affect the variance of income-producing skills within the lognormal economy, thereby increasing the variance of income in the lognormal economy as a whole. Therefore, an increase in observed variance of the lognormal income distribution can be thought of as roughly capturing "industrial" technical change.

How the Digital Economy Generates Power Laws

Previous theories of Pareto distributions in labor income often rely on matching individuals with economic fundamentals that are already known to be Pareto distributed. For example, Gabaix and Landier (2008) match managers with firms, the size of which empirically follows a Pareto distribution. This can generate Pareto distributions for managerial compensation as well. A key takeaway from this literature is that infinitesimal differences in individual ability may result in very large differences in pay. This often results from matching processes between individuals and jobs that have higher and higher stakes, creating winner-take-most job markets.

We argue that digitization and networks can generate a Pareto distribution of income more directly. Let us go back to the example of competing messaging apps. Let us assume there are K apps, which all start out with N_0 users each. At each time t, L new users arrive and face the choice of which app to use. If there are network effects (i.e. if an app is more valuable when it has more users), then each new user will tend to choose apps that already boast a large user base. For simplicity, let us assume that a new user's probability of choosing app i is simply app i's share of total users.

$$P(\text{"new user chooses } i\text{"}) = \frac{N_i}{\sum_{j=1,\dots K} N_j}$$

At $t=0$, all apps have equal probability of gaining a new users. At $t=1$, the app that ended up gaining the arriving users now has higher probability of gaining future users than other apps. This self-reinforcing mechanism, known as the preferential attachment model, leads to a power law, as was shown in the context of social networks (Barabási and Albert 1999). A simple simulation can illustrate this point.

We start with K=5 apps, having one user each. At each time, three new users arrive and preferentially choose one of the apps based on its current market share:

	Market Shares				
t	App #1	App #2	App #3	App #4	App #5
0	20.0%	20.0%	20.0%	20.0%	20.0%
10	11.4%	11.4%	2.9%	71.4%	2.9%
100	7.2%	4.3%	0.3%	87.9%	0.3%
1,000	6.8%	2.4%	0.1%	90.6%	0.0%
10,000	6.4%	2.8%	0.2%	90.5%	0.1%
100,000	6.3%	2.8%	0.2%	90.7%	0.1%

Log(Users) / Rank plot at t=1M for 100 apps

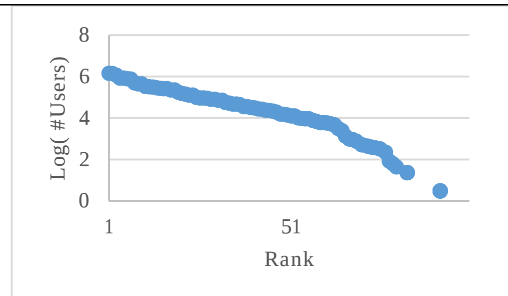


Figure 2A-2B: A simulation of power laws through preferential attachment

A common way of recognizing a power law is to regress the logarithm of the level of the variable on the rank of the observation. Figure 2B illustrates that power laws look like straight lines when represented this way. As can be seen from figure 2A, such mechanisms can converge to a very unequal distribution of users (and therefore of income), which then persists over time. This is a fundamentally different dynamic than what is observed in the industrial economy, and can have a large effect on income distribution[2].

A commonly used power law is the Pareto distribution, with survival function:

$$\Pr(X > x) = \left(\frac{c}{x}\right)^{\alpha} \ \ \forall x \in (c, +\infty)$$

When α decreases, the distribution gets a fatter and fatter tail. When α drops below 2, the variance is infinite. When α drops below one, the mean becomes infinite as well.

A Simple Mixture Model of Income Distribution

Considering that software investments have increased over 50% between 2000 and 2012 we now turn to a model of an increasingly digitized economy.

[2] The large concentration in information goods market has been documented in (Jones and Mendelson 2011)

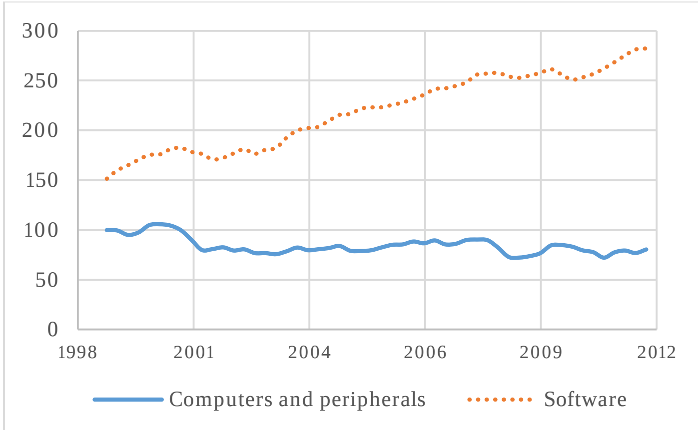


Figure 3: Software and Computer investment in billions of dollars, 1999-2012.

If, as Marc Andreessen put it, "software is eating the world" (Andreessen 2011), then we can expect that a power law will govern an increasing share of the income distribution. This is not only because the IT sector will grow, but mostly because IT tends to permeate all sectors. A striking example of this is the recent phenomenon of some teachers becoming millionaires from sharing their educational material online (Singer 2015). The distributional consequences can be large. For instance, more and more people will be below average income, and a smaller number of individuals will have a larger share of the overall income (Brynjolfsson et al. 2014).

Let us consider what happens when individuals optimally choose the economy in which they want to participate based on their expected income in each economy. An example could be an accountant who programs mobile apps on evenings, or a Hollywood waiter who is considering becoming an actor.

We posit that each person has two potential incomes drawn from two distributions:

$$X_L \sim LogNormal(\mu, \sigma)$$

$$X_P \sim Pareto(c, \alpha)$$

Where X_L is the income the individual would have if she took a job in the traditional economy (i.e. if she became an accountant or waiter), and X_P is the income she would have if she tried her luck in the power law economy (as an actor or app programmer). For simplicity, assume that each individual perfectly observes income offers from each option, and then chooses

$$Y = max(Y_P, Y_L).$$

Throughout the rest of this paper, we will refer to this model as the "Max model". Note that there are other ways to mix a Pareto distribution and a Log-Normal one, such as using a spliced model, where a lognormal distribution is fitted at the bottom and a Pareto at the top under a set of restrictions that guarantee that the resulting distribution is continuous and differentiable. Such approaches introduce scaling factors which complicate interpretation and do not map into a simple model of income generating processes, which is why we do not discuss them in the body of the paper. However, as shown in the appendix, they generate findings that are broadly consistent with the ones we derive using the Max model.

Our theory implies four hypotheses. Using the Max model, we can formally state each of them, and we can use our data to test them:

> *Hypothesis 1: If superstar effects are an important part of the economy, a model of income distributions incorporating both log-normal and Pareto distributions should fit the data significantly better than a traditional log-normal alone.*

If the lognormal variance does capture some dimension of SBTC, then we could expect it to increase over time, but much more sharply in early years than in recent years:

> *Hypothesis 2A: Industrial technical change should lead the variance of the underlying lognormal distribution to significantly increase over time.*
> *Hypothesis 2B: If industrial technical change is slowing, the observed variance of the log-normal part of the income distribution should be stabilizing over time.*

Similarly, “software eating the world” can manifest itself in two ways: we might see more and more people joining the Pareto economy (hypothesis 3), and we might rewards within the Pareto economy becoming more and more unequally distributed (hypothesis 4).

> *Hypothesis 3: If more and more people are joining the power law economy, the number of individuals whose income seems to be drawn from a Power law may be increasing over time.*

> *Hypothesis 4: If the power law economy is becoming more digital (closer to free, perfect, and instant) and more networked, incomes within that economy may become increasingly skewed, resulting in a lower Pareto parameter, α.*

Data and results

Relevant Data Sources

Studying income inequality can usually be achieved using one of two different data sources:

- Survey data, which may contain a wealth of demographic variables, information on sectors of activity and occupation. The main limitation of survey data lies in their poor sampling properties at the top: top incomes represent a very small fraction of the population, and are likely to be missed by survey sampling. However, these individuals account for a very significant fraction of total income. This is why survey data is generally not used to compute income shares. Furthermore, the study of the global shape of the income distribution requires fine-grained data, and would therefore not be achievable on survey data.
- Tax data. Because filing taxes is mandatory above a certain income threshold, federal income tax files are very helpful in studying income shares and the income distribution. However, they are surrounded by a number of strict confidentiality policies, so that demographic information is poor.

For the purposes of this paper, we use the second kind of source, tax data. We do this in particular because we are interested in studying power laws, and because so much of the action in power laws occurs at the very top of the distribution where there are very few observations, tax data provides more suitable information.

We use the IRS Tax Model Files, which are samples of US Federal Individual Income Tax returns between 1960 and 2008 (except for 1961, 1963 and 1965). The data has been anonymized and blurred out of concern for taxpayer's privacy. In particular, the data has very little demographic information, such as job or industry. For example, even state of residence is not available to researchers for all taxpayers whose income exceeds $200,000. However, its very good sampling properties at the top make it the best dataset to estimate income distributions.

In order to establish the validity of our approach, we first share a few summary statistics about the sampling design and argue that it can help us estimate a parametric model of income distribution much better than survey data could. Then, using the "wages" variable contained in the data, we provide some summary statistics about inequality, including top income shares. The following table illustrates the quality of sampling for year 1960:

q0-q50		q50-q80		q80-q90		q90-q95		q95-q99		q99-q99.9		q99.9+	
n	w/n	n	w/n	n	w/n	n	w/n	n	w/n	n	w/n	n	w/n
45,000	676	17,000	1111	6,000	1035	6,000	514	12,000	210	7,000	81	10,000	6

*Note: **n** represents the number of observations in each quantile, and **w/n** the average weight per observation in each quantile. A lower value for **w/n** at the top of the income distribution means better sampling at the top.*

Table 1: Sampling properties of IRS public use tax data

As can be seen from table 1, the sampling design of our data seems to lend itself well to a study of top incomes: relative to the rest of the population, the top incomes are significantly oversampled. Where, on average, a thousand individuals whose labor income falls between the 50th and 80^{th} percentile are grouped into one observation, only 6 individuals from the top 0.1% are grouped into one observation. This level of precision allows us to work on the shape of the income distribution as a whole, rather than conditional averages.

A first look at income densities

Let us now have a first look at the shape of the distribution of income. Using Gaussian kernel smoothing as in (DiNardo et al. 1996) allows us to graph densities of log incomes for year 1960 and 2008 as follows.

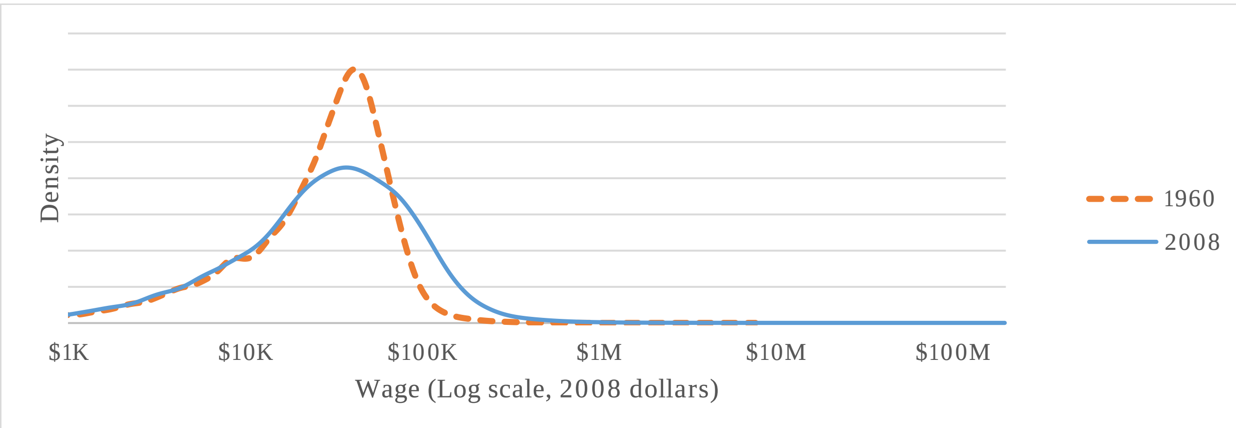


Figure 4: Kernel-Smoothed densities of wage distributions in 1960 and 2008.

One can notice that the bulk of both the 1960 and 2008 distributions look roughly log-normal. However, they show a very long right tail. Furthermore, the tail appears to have very significantly extended between 1960 and 2008. Yet this representation may still understate the importance of changes in the tails, because they are hard to see. Accordingly, we offer an illustration of fractal effects in income shares.

Fractal effects in income shares

Perhaps one of the most compelling ways of illustrating trends in income inequality is to compute "income shares." Income shares tables illustrate the share of all of a specific kind of income that is earned by a given percentile of earners. Here and throughout the rest of the paper, we restrict our analysis to wages (as reported to the IRS). This allows us to exclude capital income, which may be subject to different dynamics and interfere with our focus on individual productivity.

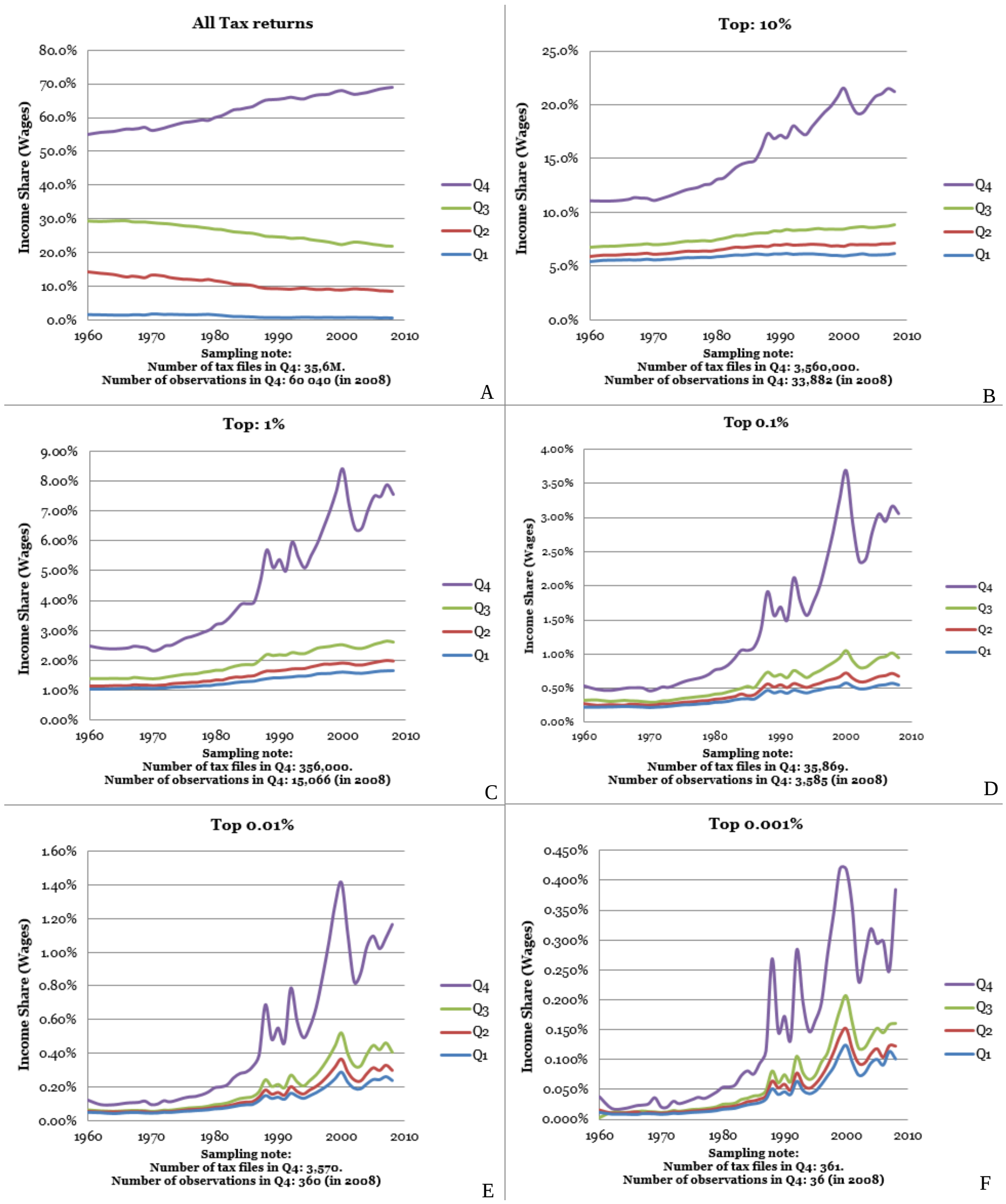
In order to give a sense of the widening of income inequality and in order to underline the influence of power laws, we offer the following visualization:

- First, we sum all the reported wages across all individuals and compute the share of these wages earned by the top 100%, 10%, 1%, 0.1%, 0.01% and 0.001% of the distribution.
- We then break these into quarters and show the evolution of their income shares since 1960.

For example, figure 5B is a decomposition of income shares of the top 10%: Q1 represents the bottom quarter of that group (i.e. individuals between *p90* and *p92.5*), and Q4 represents the top quarter (*p97.5* and above). The other figures represent the same composition, but of different portions of the population (respectively: all of it, and the top 1%, 0.1%, and so on).

This visualization has two main advantages. First, it shows that income increases are concentrated at the very top of the income distribution. More specifically, it seems that a very significant share of income inequality can be traced back to the top 0.1%. Figure 5C shows that the income share of individuals located in the top quarter of the top 1% (i.e. *p99.75* and above) went from about 2.5% in 1960 to almost 8% in 2008. About half of that 5.5 point increase can be traced back to the top quarter of the top 0.1% (i.e. *p99.975* and above): their income share went from .5% to over 3%.

Second, this visualization shows a "fractal effect". No matter how far we zoom in, the global picture is apparently the same. As you go up the income distribution, we see that the 1% have their own 1% and so on, with a similar (if somewhat noisier) pattern of income distribution. This remarkable scale independence is a feature of power laws (also called scale-free distributions).



Note: In 2008, the top 10% starts at an annual wage of $96,000$. The top 1% starts at $260,000. The top 0.1% starts at $860,000. The top 0.01% starts at $3.7M. Finally the top 0.001% starts at $15.4M.

Figures 5A-5F: Income shares fractal patterns, 1960-2008

Estimating the Max model

Having documented the presence and significance of power laws in the U.S. income distribution, let us go back to our model where

$$Y = max(Y_P, Y_L).$$

The underlying potential incomes, Y_P and Y_L are fundamendally unobservable to us as econometricians. We only observe the mixture Y, and recover the original parameters (μ, σ, α) through numerical Maximum Likelihood Estimation and Goodness of fit estimation (using variations of the Anderson-Darling statistic). Once the parameters are recovered, we can simply obtain $Q = P(Y_P > X_L)$, the proportion of people selecting their Pareto draw over their lognormal draw.

Note that the tax data contains many individuals (roughly around 15%) who report an income of zero, as well as many individuals with a very low income. We therefore fit a censored lognormal distribution. For simplicity, we give the Pareto and the Lognormal distribution the same support, by fixing c at the 25[th] percentile and then truncating the lognormal distribution at the same value of c. While this model embodies a number of simplifications, our results do not seem qualitatively sensitive to different specifications (see appendix for the “spliced” model).

It is straightforward to show that the resulting probability density function of this model is:

$$\frac{e^{-\frac{(\mu-\log(x))^2}{2\sigma^2}}(1-(\frac{x}{c})^{-\alpha})}{\sqrt{2\pi}x\sqrt{\sigma^2}} + \frac{\left(\frac{x}{c}\right)^{-1-\alpha}\alpha(1-\frac{1}{2}\Psi(-\frac{\mu-\log(x)}{\sqrt{2}\sqrt{\sigma^2}}))}{c}$$

Similarly, the resulting cumulative distribution function is:

$$(1-(\frac{x}{c})^{-\alpha})(1-\frac{\Psi(-\frac{\mu-\text{Log}(x)}{\sqrt{2}\sqrt{\sigma^2}})}{\Psi(-\frac{\mu-\text{Log}(c)}{\sqrt{2}\sqrt{\sigma^2}})})$$

Note that by design we have $0 < c < x$. $\Psi(.)$ is the canonical complementary error function, i.e.

$$\Psi(\text{x}) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2}dt$$

Maximum likelihood Estimates (MLE) are obtained by numerical optimization (using the PDF above), as there is no closed form for them because of the need bottom-truncate the distributions. We fit the model using MLE as well as different types of goodness of fit statistics (using the CDF above), including the Anderson-Darling (AD) statistic as well as ADR and AD2R, which are variants of AD that emphasize fit in the right tail (Luceño 2006). These methods yield results that are broadly similar to the ones obtained with MLE, but somewhat more dramatic: they identify a lower Pareto α and a higher number of individuals in the power law economy. However, the trends they identify over time are the same as the ones obtained through MLE. For the sake of clarity, we focus on MLE results for the remainder of this paper. The below figures show MLE estimates of the underlying parameters of the Max model:

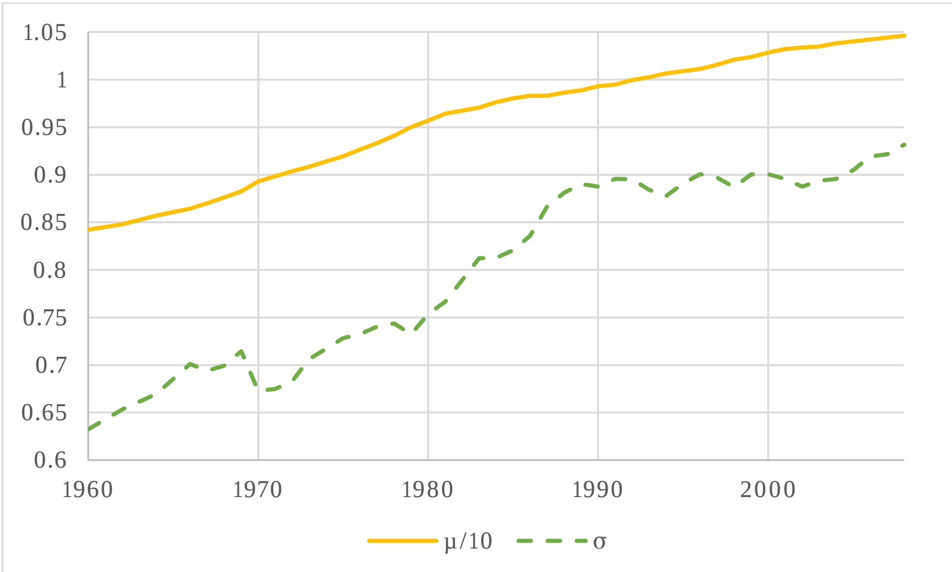


Figure 6: (μ, σ) MLE estimates of the Lognormal/Pareto Mixture based on our model

Note that it is common to report the *inverted Pareto coefficient* $\beta = \frac{\alpha}{\alpha-1}$ instead of reporting α directly. This has two main advantages: first, a higher ß means more inequality, which makes interpretation easier. But more importantly, this representation takes into account the fact that α is an exponent, which means that small changes in α result in large changes in inequality, and that the impact of a linear decrease in α is stronger and stronger as α becomes lower. In particular, as α approaches 2, the variance of a random variable drawn from a Pareto distribution becomes infinite. As α approaches 1, so does its expected value.

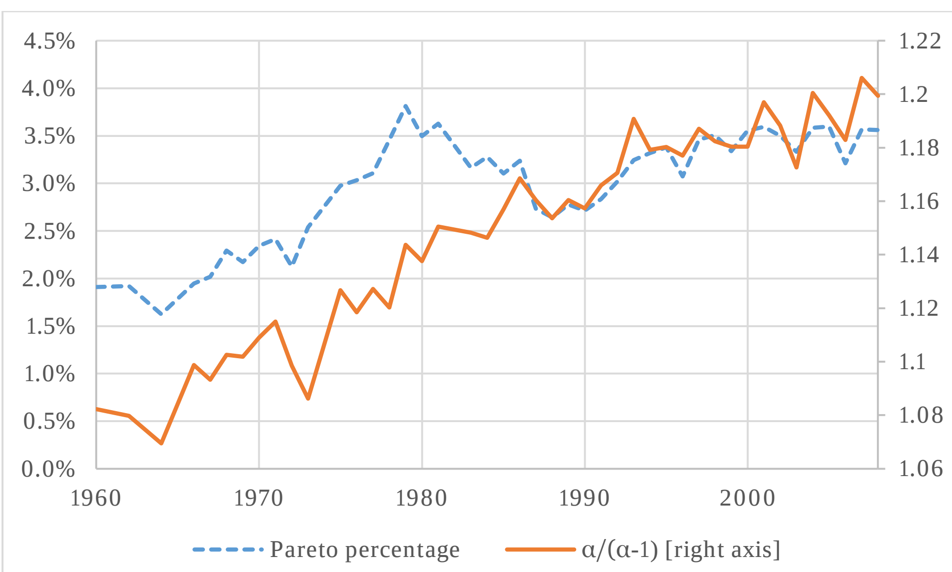


Figure 7: MLE estimates of the Pareto shape (α) and of the percentage of individuals whose income is drawn from the Pareto distribution (Q).

The estimates of this model seem to recover the qualitative facts discussed earlier: more and more people are part of the Pareto distribution, and the Pareto distribution is becoming more and more competitive.

Therefore, we test our four hypotheses using bootstrap: we resample our data 500 times and estimate the model on each sample. This procedure takes about 5 hours using 36 processors. The distribution of the parameters we obtain in this way are then used to build confidence intervals and to test our hypotheses.

Hypothesis 1: The hypothesis that our mixture model fits income data significantly better is tested using differences in Bayesian Information Criteria between the lognormal model (null) and the Max model (alternative), reproduced in the table below. The standard deviations of parameter estimates (in parentheses) are obtained by bootstrap.

	Max(P,LN) model				Lognormal Only		
year	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{Q}$	$\hat{\mu}$	$\hat{\sigma}$	ΔBIC
1960	8.42	0.63 (0.005)	13.28 (1.28)	1.79% (0.48%)	8.40	0.65	2028***
1962	8.48	0.65 (0.005)	13.74 (1.5)	1.8% (0.51%)	8.46	0.67	2285***
1964	8.57	0.67 (0.003)	15.29 (1)	1.62% (0.27%)	8.55	0.68	2583***
1966	8.64	0.7 (0.004)	11.15 (0.69)	1.92% (0.44%)	8.62	0.72	2190***
1967	8.70	0.69 (0.004)	11.75 (0.72)	2% (0.43%)	8.68	0.71	2451***
1968	8.76	0.7 (0.005)	10.77 (0.64)	2.25% (0.46%)	8.74	0.72	2247***
1969	8.82	0.71 (0.005)	10.82 (0.74)	2.14% (0.51%)	8.80	0.73	2368***
1970	8.93	0.67 (0.005)	10.1 (0.83)	2.32% (0.47%)	8.91	0.69	1735***
1971	8.99	0.67 (0.006)	9.77 (0.88)	2.34% (0.55%)	8.96	0.69	1610***
1972	9.04	0.68 (0.006)	11.17 (1.42)	2.01% (0.56%)	9.01	0.70	1553***
1973	9.08	0.71 (0.008)	12.74 (1.74)	2.37% (0.77%)	9.05	0.73	2445***
1975	9.20	0.73 (0.004)	8.9 (0.45)	2.95% (0.41%)	9.16	0.75	1773***
1976	9.27	0.73 (0.003)	9.43 (0.34)	3.03% (0.28%)	9.23	0.76	1915***
1977	9.33	0.74 (0.003)	8.85 (0.35)	3.1% (0.32%)	9.29	0.77	1796***
1978	9.41	0.74 (0.003)	9.27 (0.49)	3.46% (0.34%)	9.36	0.78	1881***
1979	9.50	0.73 (0.006)	8.01 (0.71)	3.68% (0.62%)	9.46	0.76	1305***
1980	9.57	0.75 (0.002)	8.25 (0.28)	3.5% (0.23%)	9.52	0.78	1470***
1981	9.64	0.77 (0.004)	7.63 (0.43)	3.58% (0.44%)	9.60	0.80	1364***
1983	9.71	0.81 (0.003)	7.74 (0.3)	3.15% (0.31%)	9.66	0.84	1449***
1984	9.76	0.81 (0.003)	7.84 (0.36)	3.27% (0.35%)	9.72	0.85	1464***
1985	9.80	0.82 (0.003)	7.35 (0.3)	3.1% (0.29%)	9.76	0.85	1218***
1986	9.83	0.84 (0.005)	6.92 (0.44)	3.22% (0.47%)	9.79	0.86	1212***
1987	9.83	0.87 (0.003)	7.22 (0.23)	2.73% (0.26%)	9.79	0.89	1337***
1988	9.86	0.88 (0.003)	7.48 (0.3)	2.65% (0.32%)	9.82	0.91	1402***
1989	9.89	0.89 (0.005)	7.27 (0.49)	2.72% (0.49%)	9.85	0.92	1415***
1990	9.93	0.89 (0.004)	7.34 (0.36)	2.71% (0.36%)	9.89	0.92	1345***
1991	9.95	0.9 (0.005)	7.04 (0.52)	2.8% (0.51%)	9.91	0.93	1272***
1992	9.99	0.9 (0.004)	6.83 (0.32)	3.02% (0.38%)	9.95	0.93	1237***
1993	10.03	0.88 (0.004)	6.24 (0.36)	3.23% (0.43%)	9.98	0.91	959***
1994	10.07	0.88 (0.003)	6.56 (0.31)	3.32% (0.32%)	10.02	0.91	1047***
1995	10.09	0.89 (0.004)	6.54 (0.35)	3.36% (0.42%)	10.04	0.93	1073***
1996	10.11	0.9 (0.004)	6.63 (0.29)	3.07% (0.37%)	10.07	0.93	1064***
1997	10.16	0.9 (0.003)	6.34 (0.26)	3.45% (0.35%)	10.11	0.93	986***
1998	10.21	0.89 (0.004)	6.51 (0.38)	3.47% (0.4%)	10.16	0.92	893***
1999	10.24	0.9 (0.004)	6.6 (0.34)	3.3% (0.38%)	10.19	0.94	905***
2000	10.28	0.9 (0.006)	6.59 (0.59)	3.46% (0.61%)	10.23	0.94	888***
2001	10.32	0.9 (0.003)	6.07 (0.23)	3.59% (0.3%)	10.27	0.93	764***
2002	10.34	0.89 (0.003)	6.28 (0.24)	3.5% (0.29%)	10.29	0.92	838***
2003	10.35	0.89 (0.003)	6.78 (0.28)	3.33% (0.3%)	10.30	0.93	982***
2004	10.38	0.9 (0.003)	6.01 (0.22)	3.58% (0.3%)	10.33	0.93	729***
2005	10.40	0.9 (0.005)	6.27 (0.46)	3.52% (0.52%)	10.35	0.94	791***
2006	10.42	0.92 (0.004)	6.5 (0.37)	3.17% (0.44%)	10.37	0.96	861***
2007	10.44	0.92 (0.003)	5.84 (0.24)	3.56% (0.29%)	10.39	0.96	663***
2008	10.46	0.93 (0.003)	6.03 (0.23)	3.55% (0.31%)	10.41	0.97	768***

Table 2: MLE estimates of the mixture model and Likelihood Ratio Statistic when compared to the simple lognormal model

Using a bootstrap approach, we determine the 0.1% critical value for the ΔBIC statistic for each year. We do not use likelihood ratio tests because of the presence of sample weights in the data, which may lead us to reject the null too often. We obtain BIC critical values by simulating 10,000 datasets drawn from the null distribution, and computing the difference in the BIC of the two models. For year 2008, it is equal to 310. The null model is rejected for all years at the 0.1% level, leading us to conclude that the Max model fits the data significantly better than a simple lognormal. In other words, focusing on the industrial economy only would mean missing an important part of the dynamics of income[3].

Hypothesis 2A (σ increasing)**:** Increased σ (standard deviation of the lognormal distribution underlying the industrial economy) is a prediction of skill-biased technical change. Out of 500 bootstrap estimations, the intersection of the interval of estimated σ for 1960 and the interval of estimated σ for 2008 is empty. This gives us an estimated p-value under 0.002, and we can conclude with confidence that σ has significantly increased between 1960 and 2008.

Hypothesis 2B (regime change in σ)**:** Figure 6 above illustrates the evolution of the log-normal σ over time. Visually, the data estimates seem to exhibit a clear break in trends in the 1980s, with the slope of the evolution of σ declining sharply. More formally, we can test for a structural beak by regressing our estimated σ (from our bootstrap sample of 500 estimates of σ per year) on time and performing a Chow test. Testing for a structural break in year 1988 yields a very large F-statistic of 1464.38, so we can reject the null that there is no trend change in sigma at the 0.1% level.

Hypothesis 3 (Q increasing)**:** Q represents the number of individuals who select into the power law economy. Out of 500 bootstrap estimations, the intersection of the interval of estimated q for 1960 and the interval of estimated Q for 2008 is empty. This gives us an estimated p-value under 0.002, and we can conclude with confidence that Q has significantly increased between 1960 and 2008.

Hypothesis 4 (α decreasing)**:** α measures how fat the tail of the Pareto distribution is. A lower α means fatter tails and more skewed rewards. Out of 500 bootstrap estimations, the intersection of the interval of estimated alphas for 1960 and the interval of estimated alphas for 2008 is empty. This gives us an estimated p-value of under 0.002, and we can conclude with confidence that α has significantly decreased between 1960 and 2008.

One can note that the increase in Q has slowed in recent years compared to the beginning of the period we study. However, rewards are getting more and more skewed within the power-law economy.

Discussion and extensions

Alternative Theories of Inequality*:* Using only available data on income distribution, telling different stories apart has been a long-standing challenge for the literature. Given the scarcity of data, especially regarding our ability to measure IT use at a low level of granularity, identifying the effect of IT using a more traditional regression approach remains a challenge. Let us however note that is that institutional and SBTC-based explanations of inequality do not predict an increased prevalence of power laws, whereas superstar explanations do. We argue that this fact, when combined with fine grained income data such as the IRS public use tax files, can be used to assess various theories using income data and show that observed changes are consistent with an increased role of IT.

Robustness checks and sensitivity: When fitting the Max model, we set the lower bound of the truncated lognormal distribution and of the Pareto distribution be the 25th percentile of the wage distribution. This allows us to avoid starting the Power Law distribution at or close to zero, as the percentage of tax files reporting zero wage income reaches 16% on some years. We used different specifications, such as using an

[3] It is worth mentioning that the distribution of income may be even more closely approximated by more complex and generalized distributions, such as the double-Pareto-lognormal distribution, or by involving Levy or Weibull distributions. We do not do so for the sake of interpretability of the results.

absolute value rather than a percentile, or using different levels. We also used more generalized versions of the Pareto distribution, such as the Lomax. Such variations alter estimated α and Q coefficients, however the direction of changes in coefficients remains the same, and the estimated significance of these changes remains very high. Therefore, the main takeaway from this paper is that more and more people are joining the power law economy and that this economy is more and more dispersed, keeping in mind that the magnitude of these effects is sensitive to specification. Finally, running models on absolute values of income versus well as de-meaned and inflation adjusted income has no impact on the direction of effects, and very little effect on magnitudes. Tables reflecting results under these alternative specifications are available from the authors upon request.

Extending the analysis to capital income: An in-depth study of capital income is beyond the scope of this paper. We present some results using gross income, which includes both wages and capital income in the appendix. The results are similar to the ones obtained on wages, although the mechanisms which drive the Pareto economy in capital income might be different from the ones which drive the Pareto economy in wages. We are showing our Max analysis on gross income as reported to the IRS in Figure 9. The results are broadly similar to the ones we obtained for wages only: a more competitive Pareto economy, and more people joining it.

Conclusion

Our results suggest that as the economy becomes more and more digitized and networked, the income distribution is changing dramatically. While the idea that increased productivity may benefit everyone in the very long term is certainly plausible, in this paper we show that this need not be the case in the short or medium term. The key to understanding inequality seems to lie in understanding the underlying market mechanisms yielding highly skewed income distributions.

We find that (1) the distribution of income is better approximated by a model where individuals choose optimally between draws from a Pareto distribution (power law economy) and Log-normal distribution (industrial economy) than by a Log-normal distribution alone. Using this model, we find that the evolution of the shape of the income distribution is consistent with (2) a leveling off of the variance of the underlying lognormal distribution in the industrial economy, which is consistent with the hypothesis that skill-biased technical change is no longer accelerating as it did earlier in the period; (3) an increased number of individuals selecting into the power law economy which generates a power law distribution of income, and (4) a power law economy delivering increasingly skewed rewards: the winners are winning bigger than ever. We find that in recent years, the number of individuals joining the power law economy has stabilized, but that competitiveness within that economy has increased.

More and more people moving into the power law economy does not necessarily mean more and more people being better off in absolute terms because of digitization. Indeed, our model accommodates both people moving to the power law economy because of increasing opportunity there and people moving because of decreasing opportunity in the industrial economy. Furthermore, we show that the power law economy is producing increasingly unequal rewards.

If the trends we document in this paper continue, the income distribution will become increasingly skewed toward not just the top 1%, but the top 1% of the top 1%. Future research is needed in order to more causally distinguish between different types of technologies and their effects on inequality. Furthermore, the effects of digitization on worldwide inequality patterns seems like a worthwhile extension of our research.

Appendix

Using a spliced Lognormal-Pareto Specification

As an alternative to our Max model we also use a "spliced" model of the Pareto/Log-Normal mixture. Specifically, we model the income distribution with the following PDF:

$$f(x) = \begin{cases} \beta\dfrac{f_1(x|\mu,\sigma)}{F_1(\theta|\mu,\sigma)} \forall\, x \leq \theta \\ (1-\beta)\dfrac{f_2(x|c,\alpha)}{1-F_2(\theta|c,\alpha)}, \forall\; x > 0 \end{cases}$$

Were $f_1(x|\mu,\sigma)$ is the probability distribution function of a lognormal distribution with parameters μ and σ, and $f_2(x|c,\alpha)$is the p.d.f of a simple, type I Pareto distribution with cutoff c and shape α. β is a weighting coefficient. Note that this distribution is not in general continuous or differentiable, which may be undesirable for modelling purposes, as the empirical income distribution function does not seem to exhibit any major discontinuity. However, ensuring continuity is possible by placing restrictions on μ and β (Nadarajah and Bakar 2012). This requires imposing the following:

$$\mu = \ln(\theta) + \sigma^2 + \frac{\theta\sigma^2 f_2'(\theta|c,\alpha)}{f_2(\theta|c,\alpha)}$$

$$\phi = \frac{f_1(\theta|\mu,\alpha)(1-F_2(\theta))}{f_2(\theta)F_1(\theta)}$$

(with notation change $\beta = \frac{1}{1+\phi}$).

This setup gives us a continuous and differentiable p.d.f. Note that the likelihood function is not in general continuous at θ, so that MLE is consistent but not necessarily efficient (Bee 2012). MLE estimation (fixing c at the 25th percentile) gives the following results:

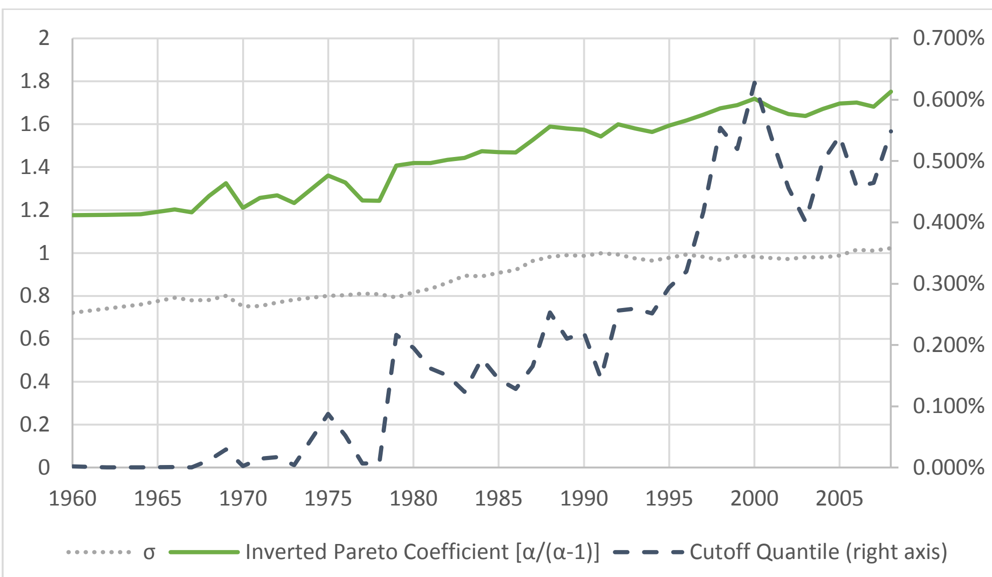


Figure 8: Estimates of the Spliced Model

The change in the number of individuals is in the Pareto economy is lower than predicted by the Max model. This is expected, since this model equates "being in the Pareto" and "being at the top", whereas the Max model allows for individuals with low income to self-select into the power law economy (because their industrial income would be lower still).

Including Capital Income

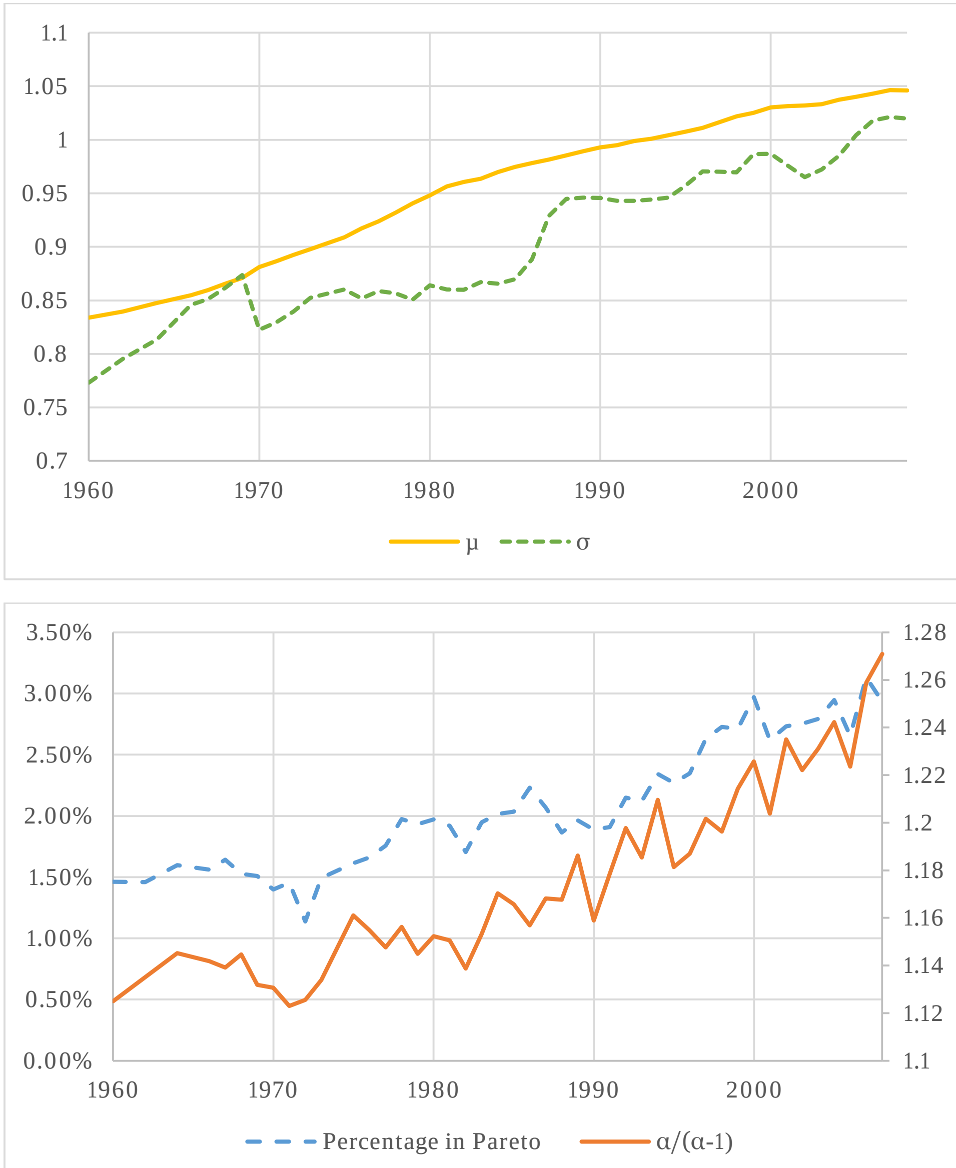


Figure 9: Estimates of the Max Model on Gross Income (including wages and capital income)

Note: in this chart, the truncation cutoffs of the lognormal and the lower cutoff of the Pareto distribution are set at 8%. The tax files we study contain much fewer individuals reporting zero gross income than zero wages, which allows us to lower the truncation cutoff.

References

Acemoglu, D., and Autor, D. 2012. "What Does Human Capital Do? A Review of Goldin and Katz's The Race between Education and Technology," *Journal of Economic Literature* (50:2), pp. 426–463 (doi: 10.1257/jel.50.2.426).

Acemoglu, D., and Robinson, J. 2012. "Why Nations Fail: The Origins of Power," *Prosperity, and Poverty, Crown Business, New York*.

Andreessen, M. 2011. "Why Software Is Eating the World," *The Wall Street Journal* (available at http://www.wsj.com/articles/SB10001424053111903480904576512250915629460).

Andriani, P., and McKelvey, B. 2009. "Perspective—From Gaussian to Paretian Thinking: Causes and Implications of Power Laws in Organizations," *Organization Science* (20:6), INFORMS, pp. 1053–1071 (doi: 10.1287/orsc.1090.0481).

Atkinson, A. B., and Piketty, T. 2010. *Top incomes: A global perspective*, Oxford University Press.

Autor, D. H. 2014. "Skills, education, and the rise of earnings inequality among the 'other 99 percent,'" *Science (New York, N.Y.)* (344:6186), pp. 843–51 (doi: 10.1126/science.1251868).

Autor, D. H., Katz, L. F., and Kearney, M. S. 2008. "Trends in U.S. Wage Inequality: Revising the Revisionists," *Review of Economics and Statistics* (90:2), pp. 300–323 (doi: 10.1162/rest.90.2.300).

Autor, D. H., Katz, L. F., and Krueger, A. B. 1998. "Computing Inequality: Have Computers Changed the Labor Market?," *Quarterly Journal of Economics* (113:4), pp. 1169–1213 (doi: 10.1162/003355398555874).

Barabási, A.-L., and Albert, R. 1999. "Emergence of scaling in random networks," *Science* (286:5439), American Association for the Advancement of Science, pp. 509–512.

Bee, M. 2012. "Statistical analysis of the Lognormal-Pareto distribution using Probability Weighted Moments and Maximum Likelihood," *Department of Economics Working Papers*, Department of Economics, University of Trento, Italia (available at http://ideas.repec.org/p/trn/utwpde/1208.html).

Bernstein, J. 2010. "Three Questions About Consumer Spending and the Middle Class," *Bureau of Labor Statistics* (available at http://www.bls.gov/cex/duf2010bernstein1.pdf; retrieved May 3, 2015).

Bivens, J., and Mishel, L. 2013. "The Pay of Corporate Executives and Financial Professionals as Evidence of Rents in Top 1 Percent Incomes," *The Journal of Economic Perspectives* (27:3), pp. 57–78 (doi: 10.2307/41955545).

Bresnahan, T. F., Brynjolfsson, E., and Hitt, L. M. 2002. "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence," *The Quarterly Journal of Economics* (117:1), Oxford University Press, pp. 339–376 (doi: 10.2307/2696490).

Brynjolfsson, E., Hu, Y. (Jeffrey), and Smith, M. D. 2003. "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers," *Management Science* (49:11), INFORMS, pp. 1580–1596 (doi: 10.1287/mnsc.49.11.1580.20580).

Brynjolfsson, E., McAffee, A., and Spence, M. 2014. "New World Order: Labor, Capital, and Ideas in the Power Law Economy," *Foreign Affairs* (doi: 10.3318/ISIA.2004.15.1.3).

Corak, M. 2013. "Income Inequality, Equality of Opportunity, and Intergenerational Mobility," *The Journal of Economic Perspectives* (27:3), American Economic Association, pp. 79–102 (doi: 10.2307/41955546).

Debacker, J., Heim, B., Panousi, V., Ramnath, S., and Vidangos, I. 2013. "Rising Inequality: Transitory or Persistent? New Evidence from a Panel of U.S. Tax Returns," *Brookings Papers on Economic Activity*, Brookings Institution Press, pp. 67–122 (doi: 10.2307/23594863).

DiNardo, J., Fortin, N. M., and Lemieux, T. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica* (64:5), The Econometric Society, pp. 1001–1044 (doi: 10.2307/2171954).

Faloutsos, M., Faloutsos, P., and Faloutsos, C. 1999. "On power-law relationships of the internet topology," in *ACM SIGCOMM computer communication review* (Vol. 29), pp. 251–262.

Fleder, D., and Hosanagar, K. 2009. "Blockbuster Culture's Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity," *Management Science* (55:5), INFORMS, pp. 697–712 (doi: 10.1287/mnsc.1080.0974).

Gabaix, X., and Landier, A. 2008. "Why Has CEO Pay Increased so Much?," *The Quarterly Journal of Economics* (123:1), Oxford University Press, pp. 49–100 (doi: 10.2307/25098894).

Gibrat, R. 1931. *Les inégalités économiques*, Recueil Sirey.

Hajargasht, G., and Griffiths, W. E. 2013. "Pareto--lognormal distributions: Inequality, poverty, and estimation from grouped income data," *Economic Modelling* (33), Elsevier, pp. 593–604.

Jones, R., and Mendelson, H. 2011. "Information Goods vs. Industrial Goods: Cost Structure and Competition," *Management Science*, INFORMS (available at http://pubsonline.informs.org/doi/abs/10.1287/mnsc.1100.1262).

Kaplan, S. N., and Rauh, J. 2013. "It's the Market: The Broad-Based Rise in the Return to Top Talent," *Journal of Economic Perspectives* (27:3), pp. 35–56 (doi: 10.1257/jep.27.3.35).

Katz, L., and Autor, D. 1999. "Changes in the Wage Structure and Earnings Inequality," *Handbook of Labor Economics* (3), pp. 1463–1555 (doi: 10.1016/S1573-4463(99)03007-2).

Katz, L. F. 2000. "Technological change, computerization, and the wage structure," in *Understanding the Digital Economy: Data, Tools, and Research*E. Brynjolfsson and B. Kahin (eds.), MIT Press, Cambridge MA.

Kim, H., and Brynjolfsson, E. 2009. "CEO Compensation and Information Technology," in *ICIS 2009 Proceedings* (available at http://aisel.aisnet.org/icis2009/38).

Krueger, A. B. 2013. "Land of Hope and Dreams: Rock and Roll, Economics and Rebuilding the Middle Class," *Speech at the White House*.

Kuznets, S. 1955. "Economic growth and income inequality," *The American Economic Review* (45:1), American Economic Association, pp. 1–28 (doi: 10.2307/1811581).

Lopez, H., and Serven, L. 2006. "A normal relationship? Poverty, growth, and inequality," *Poverty, Growth, and Inequality (January 2006). World Bank Policy Research Working Paper* (3814).

Luceño, A. 2006. "Fitting the generalized Pareto distribution to data using maximum goodness-of-fit estimators," *Computational Statistics & Data Analysis* (51:2), pp. 904–917 (doi: 10.1016/j.csda.2005.09.011).

Mincer, J. 1958. "Investment in Human Capital and Personal Income Distribution," *Journal of Political Economy* (66:4), The University of Chicago Press, pp. 281–302 CR – Copyright © 1958 The University (doi: 10.2307/1827422).

Nadarajah, S., and Bakar, S. A. A. 2012. "New composite models for the Danish fire insurance data," *Scandinavian Actuarial Journal* (2014:2), Taylor & Francis, pp. 180–187 (doi: 10.1080/03461238.2012.695748).

Pigou, A. C. 1932. *The economics of welfare*, Macmillan.

Piketty, T. 2001. *Les hauts revenus en France au 20ème siècle*.

Piketty, T., and Saez, E. 2003. "Income Inequality in the United States, 1913--1998," *The Quarterly Journal of Economics* (118:1), pp. 1–39 (doi: 10.1162/00335530360535135).

Reed, W. J., and Jorgensen, M. 2004. "The double Pareto-lognormal distribution -- a new parametric model for size distributions," *Communications in Statistics-Theory and Methods* (33:8), Taylor & Francis, pp. 1733–1753.

Rosen, S. 1981. "The Economics of Superstars," *The American Economic Review* (71:5), American Economic Association, pp. 845–858 (doi: 10.2307/1803469).

Saez, E. 2013. "Striking it Richer: The Evolution of Top Incomes in the United States," *University of California at Berkley working paper* (available at http://eml.berkeley.edu/~saez/saez-UStopincomes-2012.pdf).

Singer, N. 2015. "A Sharing Economy Where Teachers Win," *The New York Times* (available at http://www.nytimes.com/2015/09/06/technology/a-sharing-economy-where-teachers-win.html?_r=0; retrieved September 7, 2015).