A Simple Marshallian Theory of Top Incomes

Early-stage draft, comments welcome

Karl Harmenberg*

May 14, 2020

Abstract

I introduce a simple model which endogenously generates a Pareto distribution in top incomes, consistent with empirics. Workers inhabit different niches, and the income of a worker is determined by the niche-specific supply of labor and a constant-elasticity labor-demand curve. The highest paid workers are the ones that inhabit a niche with few other workers. A Pareto tail in incomes emerges as long as the distribution of workers over niches satisfies a regularity condition from extreme-value theory, satisfied by virtually all continuous distributions in economics.

1 Introduction

The Pareto property of top incomes is one of the most striking regularities in economics. Since Pareto (1896)'s original discovery, it has remained true across time and place that the top of the income distribution closely follows a Pareto distribution, i.e., the share of individuals s with income above a certain income level y is well approximated by the functional form $s \propto y^{-\alpha}$.

In this paper, I introduce a simple theory for why income distributions generically feature this regularity. In the model, the labor market consists of different niches and the pay of workers within a niche depends on the supply of workers in the niche together with a constant-elasticity labor demand curve. The highest paid workers are therefore the ones that inhabit a niche with few other workers. Lacking a detailed theory of how workers distribute over the niches, it is in general difficult to say much about the shape of the income distribution. However, to characterize the shape of the income distribution at the very top, it is only necessary to characterize the asymptotic shape of the distribution of workers across niches. A simple proof shows that the top of the income distribution is asymptotically Pareto, as long as the distribution of workers is regular, in the sense of extreme value theory. This class includes the normal distribution, the

*kha.eco@cbs.dk, Copenhagen Business School.

log-normal distribution, the Pareto distribution, the Fréchet distribution, the beta distribution, the gamma distribution, the exponential distribution, and virtually all continuous distributions considered in economics. We can therefore maintain a considerable degree of agnosticism with respect to the underlying processes that generate the distribution of workers over niches, while maintaining the robust prediction that top incomes are Pareto distributed.

Related literature Gabaix (2009) surveys the literature on power laws in economics and finance, and Sornette (2006) surveys a broader literature on power laws in the natural sciences.

The closest paper to this paper is Geerolf (2017). At a deeper level, although I frame my model in terms of variations in labor supply and Geerolf (2017) frames his model in terms of variations in labor demand stemming from superstar effects, the models share a similar approach to top income inequality. Mathematically, both Geerolf (2017) and this paper employ what Sornette (2006) calls a 'power law change of variable close to the origin' to generate a Pareto tail in income. I show that a simple Marshallian supply-demand framework with a standard constant-elasticity demand curve generates a Pareto tail in labor income and Geerolf (2017) shows that an assignment model with a constant-elasticity production function generates a Pareto tail in firm size and labor income.

2 Model

The labor market consists of a continuum of niches which workers can inhabit and the income of a worker in a niche is determined by the number of workers in the worker's niche. The best paid niches are the ones with very few workers. The model does not take a stand on how the distribution of workers over niches is generated. Some well-paid niches may require skill to reach while access to other well-paid niches may be regulated, and yet other well-paid niches may require a high degree of luck to reach. Regardless of how the distribution of workers over niches arises, once we have a distribution of workers over niches, we can compute the induced income distribution. The main result of the paper is that, for a large class of distributions, the induced income distribution features an endogenous Pareto tail.

2.1 Distribution of workers across niches

Without loss of generality[1], we order the niches from most common to least common on the real non-negative half-line $\mathbb{R}_+$. Let niches be indexed by x and let the distribution of workers across niches be given by f_x. The ordering of niches implies that f_x is monotonically decreasing. To have an unbounded income distribution, we assume that for all $\epsilon > 0$, there exists a niche x such that $0 < f_x(x) < \epsilon$. Let $\overline{x} = \sup_x\{f_x(x) > 0\} \in \mathbb{R}_+ \cup \{\infty\}$ denote the supremum of the support of f_x. Finally, assume that f_x is regular in the following sense:

Assumption 1. *For all $\epsilon > 0$, there exists a niche x such that $0 < f_x(x) < \epsilon$. Furthermore, f_x is regular, i.e., the following limit exists,*

$$\lim_{x \to \overline{x}} \frac{\partial}{\partial x} \frac{\overline{F}_x(x)}{f_x(x)} = \xi,$$

and $\xi > -1$.

Remark 1. *The uniform distribution is regular with $\xi = -1$, the Weibull distribution is regular with $\xi < 0$, the Pareto and Fréchet distributions are regular with $\xi > 0$, and the Gaussian, log-normal, Gumbel, exponential, stretched exponential, and loggamma distributions are regular with $\xi = 0$.*

Assumption 1 stems from extreme-value theory and has previously been used in economics by Gabaix and Landier (2008). It is satisfied by virtually all continuous distributions considered in economics and it is therefore a weak restriction on the density function to insist that it is regular in the above sense. By differentiating, Assumption 1 implies that $\lim_{x \to \overline{x}} -\frac{\overline{F}_x(x) f'_x(x)}{f_x(x)^2} = 1 + \xi$. Note also that $\xi > -1$ implies that f_x is eventually strictly monotonically decreasing.

The constant ξ measures the fatness of the tail of the distribution. For the thin-tailed distributions such as the normal distribution, the exponential distribution, and the log-normal distribution, $\xi = 0$. For bounded distributions such as the Weibull distribution, $\xi < 0$. For distributions with fat tails, such as the Pareto and Fréchet distributions, $\xi > 0$. In particular, for a Pareto distribution $\xi = 1/\alpha$.

2.2 Income in a niche

The income of a worker in a niche x with mass of workers $f_x(x)$ is given as a function of the mass of workers, $y = g_y(f_x(x))$. That is, the income in a niche is determined by the amount of workers inhabiting the niche. The income function g_y is shared across niches.

[1] Let the density function be represented by a histogram. Intuitively, we want to argue that, without loss of generality, the histogram can be sorted in descending order. If the density function is over (say) two dimensions, we want to argue that we can take all the columns of the two-dimensional histogram and order them in a row in descending order.

Let the space of niches be $\mathbb{R}^N$ (or a manifold). Let $f : \mathbb{R}^N \to \mathbb{R}_+$ be a sufficiently smooth density function. Define $G(a) = \int [f(x) > a] dx$. By Sard's theorem, G inherits almost everywhere differentiability and is almost everywhere non-singular. Write $\tilde{f}(a) = G^{-1}(a)$. The new density $\tilde{f}(a)$ on $\mathbb{R}_+$ is a reordering of f, equivalent to a reshuffling of the histogram, and it inherits (almost everywhere) differentiability from f.

We assume that the income function g_y satisfy the following properties:

Assumption 2. *The income function* $g_y : \mathbb{R}_+ \to \mathbb{R}_+$ *is differentiable and strictly monotonically decreasing. The limit elasticity of income exists when the mass of workers* L *approaches zero,*

$$\lim_{L\to 0} \frac{-g_y'(L)L}{g_y(L)} = \frac{1}{\epsilon}.$$

This is equivalent to assuming that, locally near zero, income is determined by a constant-elasticity labor demand curve, $y \propto L^{-1/\epsilon}$.

What is a niche? In effect, we have assumed that the demand for labor is constant across niches. This is largely without loss of generality since we are free to redefine the size of a niche in a way that equalizes demand across niches. In general, let labor demand in a particular niche be given by $A_i D(y_i)$. The income in a given niche is then determined by $L_i = A_i D(y_i)$ or $y_i = D^{-1}(L_i/A_i)$. Reparametrize the space so that $d\tilde{x} = dx/A(x)$ or $\tilde{x} = \int_0^x \frac{1}{A(x')} dx'$. For example, let us consider the case where demand is proportional to population density and, rather than considering the continuous case, let each niche be a discrete geographic bin. The reparametrization means that the the grid of bins should be very fine on Manhattan but coarse in northern Sweden so that each niche contains an equivalent-sized population.

2.3 The income distribution

A distribution of workers across niches f_x together with an income function g_y generates an income distribution f_y. We say that the income distribution f_y has a Pareto tail if

$$\lim_{y\to\infty} \frac{-y f_y(y)}{\bar{F}_y(y)} = \alpha.$$

We can now state and prove the main result of this paper:

Theorem 1. *Given Assumption 1 on the distribution of workers across niches and Assumption 2 on the income function, the income distribution features a Pareto tail. The Pareto tail coefficient is given by* $\alpha = \epsilon/(1+\xi)$.

Proof. Write $y_0 = g_y(f_x(x_0))$. Since both $g_y(\cdot)$ and $f_x(\cdot)$ are (eventually) strictly monotonic, the correspondence is bijective and we have

$$\bar{F}_y(y_0) = \bar{F}_x(x_0).$$

Differentiate both the left-hand side and the right-hand side with respect to x_0. We get

$$-f_y(y_0)g_y'(f_x(x_0))f_x'(x_0) = f_x(x_0).$$

or

$$f_y(y_0) = -\frac{f_x(x_0)}{g_y'(f_x(x_0))f_x'(x_0)}.$$

Therefore,

$$\lim_{y_0\to\infty} \frac{-y_0 f_y(y_0)}{\bar{F}_y(y_0)} = \lim_{x_0\to\bar{x}} \frac{-g_y(f_x(x_0))f_y(y_0)}{\bar{F}_x(x_0)} = \lim_{x_0\to\bar{x}} \frac{-f_x(x_0)^2}{\bar{F}_x(x_0)f_x'(x_0)} \frac{-g_y(f_x(x_0))}{f_x(x_0)g_y'(f_x(x_0))} = \frac{\epsilon}{1+\xi}.$$

□

For some functional forms, the income distribution is exactly (globally) Pareto shaped. For example, with the income function $g_y(L) = L^{-1/\epsilon}$, a triangular distribution of workers on $[0,1]$ given by $f_x(x) = 2-2x$ generates a Pareto distribution in income with $\alpha = 2\epsilon$, an exponential distribution of workers on $[0,\infty)$ given by $f_x(x) = \lambda\exp(-\lambda x)$ generates a Pareto distribution in income with $\alpha = \epsilon$, and a Pareto distribution of workers on $[1,\infty)$ given by $f_x(x) = \beta x^{-\beta-1}$ generates a Pareto distribution in income with $\alpha = \frac{\beta}{1+\beta}\epsilon$.

Theorem 1 shows that, even though the income distribution may not be following a Pareto distribution throughout the entire distribution, it will do so at the top of the income distribution for an arbitrary distribution of workers across niches, and the Pareto coefficient is given by $\alpha = \frac{\epsilon}{1+\xi}$. Although the normal distribution, the exponential distribution, and the log-normal distribution are visibly distinct, through the lens of extreme value theory they all have in common that they are thin-tailed distributions with $\xi = 0$. The Pareto coefficient of the endogenous income distribution is therefore unchanged, $\alpha = \epsilon$, regardless of whether the underlying distribution of workers across niches is normal, exponential or log-normal. Figure 2.1 shows graphically how the income distributions generated by an exponential, a normal and a log-normal distribution of workers over niches all generate asymptotic Pareto tails.

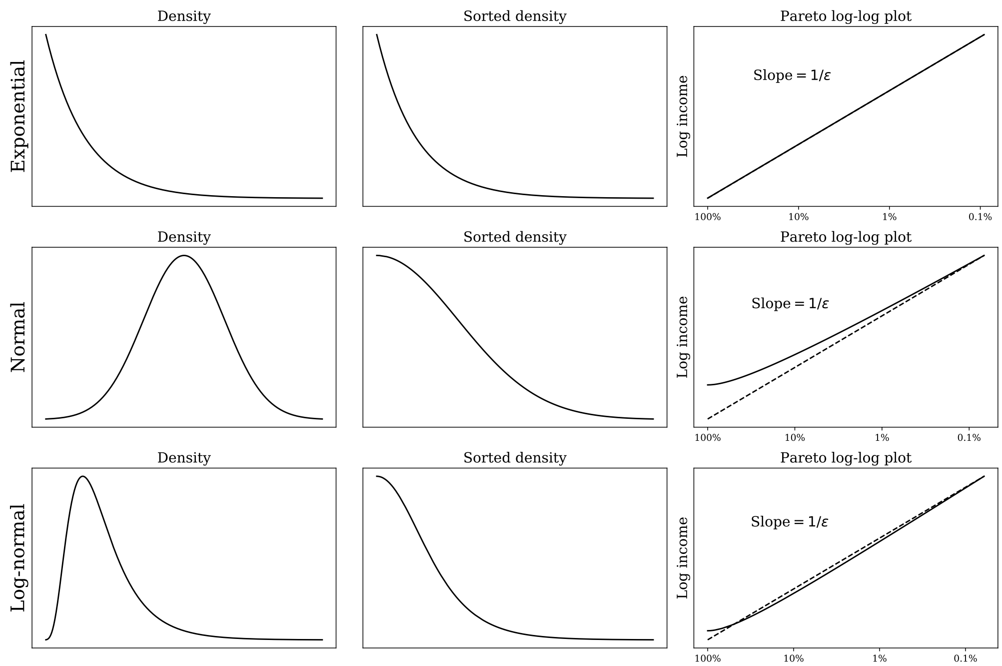

Figure 2.1: A constant-elasticity demand curve, $y = L^{-1/\epsilon}$, generates an exact Pareto distribution of income when workers are distributed according to an exponential distribution. For other distributions such as the normal and log-normal distribution, the income distribution asymptotically approaches a Pareto distribution.

3 Discussion

Theorem 1 shows that a standard supply-and-demand framework, together with the assumption of a constant-elasticity demand curve, generically generates a Pareto distribution in top incomes. Ultimately, any theory of the Pareto property of top incomes needs to rely on some functional-form assumption, and the main takeaway of this paper is that the only functional-form assumption necessary is one which economists have long been comfortable employing.

The theorem also shows that, although the Pareto tail parameter of the income distribution, $\alpha = \epsilon/(1+\xi)$, depends on the distribution of workers across niches, it only depends on the distribution through the fatness of the tail ξ. In particular, top-income inequality is equal to $\alpha = \epsilon$ for all unbounded thin-tailed distributions such as the normal distribution, the exponential distribution, and the log-normal distribution.

Why does the distribution of workers across niches not matter more for the Pareto tail parameter? Imagine you could distribute workers across niches and that your objective was to decrease top income inequality. Would you lower the number of hedge fund managers? In partial equilibrium, reducing the number of hedge fund managers lowers the number of highly paid executives, reducing top-income inequality. However, in general equilibrium, reducing the number of hedge fund managers increases the income of the remaining hedge fund managers, thereby increasing top-income inequality. Theorem 1 shows that, for perturbations that preserve the fatness ξ of the distribution of workers across niches, these two effects offset each other and leave top-income inequality unchanged.

The elasticity ϵ of the labor demand curve is the fundamental determinant of top-income inequality. In models with monopolistic competition, this elasticity also determines the level of "markups"/rents the workers can obtain. A fall in the elasticity ϵ of the labor demand curve for the professional elite can therefore jointly explain both the widening gap between the middle class and the professional elite as well as the recent increase in top income inequality (within the professional elite).

References

Gabaix, X., sep 2009. Power Laws in Economics and Finance. Annual Review of Economics 1 (1), 255–294. URL `http://www.annualreviews.org/doi/abs/10.1146/annurev.economics.050708.142940`

Gabaix, X., Landier, A., 2008. Why has CEO pay increased so much? Quarterly Journal of Economics 123 (1), 49–100.
URL `http://www.nber.org/papers/w12365`

Geerolf, F., 2017. A Theory of Pareto Distributions. Working Paper (May), 1–49.

Pareto, V., 1896. Cours d.économie politique.

Sornette, D., 2006. Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization and Disorder: Concepts and tools. Springer Science & Business Media.