A Simple Theory of Pareto-Distributed Earnings*

Karl Harmenberg†

December 5, 2023

Abstract

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

*I thank editor Eric Young, an anonymous referee, Timo Boppart, Nathan Hancart, Hannes Malmberg, and Rafael Treibich for valuable comments. All errors are of course mine alone. Funding from the Norwegian Research Council, grant number 316301, is gratefully acknowledged.

†karl.harmenberg@econ.uio.no, University of Oslo.

1 Introduction

The Pareto property of top earnings is one of the most striking regularities in economics. Since Pareto (1896)'s original discovery, it has been verified across regions and time periods that the top of the earnings distribution closely follows a Pareto distribution, i.e., the share s of individuals with earnings above a certain earnings level y is well approximated by the functional form $s \propto y^{-\alpha}$. In this paper, I provide a theory of this empirical regularity leveraging the basic economic idea that scarcity leads to high prices, or in the case of labor, high earnings.

The model is simple. Workers are distributed over a continuum of different 'niches' and earnings in a particular niche is determined by the number of workers in the niche together with a downward-sloping labor-demand curve featuring a limit elasticity as the number of workers tend to zero. The top earners are the ones who find themselves in a niche with very few competitors. In this setting, for a large class of distributions of workers across niches, the induced earnings distribution features a Pareto distribution for top earnings.

Related literature The predominant framework for generating a Pareto tail in earnings is through random-growth processes, following Champernowne (1953), Simon (1955) and Kesten (1973).[1] By contrast, this paper provides a *static* theory of top-earnings inequality leveraging that scarcity leads to high prices. Geerolf (2017) also provides a static theory of top-earnings inequality. Although my model leverages variations in scarcity and Geerolf (2017) studies an assignment model generating superstar effects, the models share an approach to top-earnings inequality which mathematically boils down to what Sornette (2006) calls a 'power law change of variable close to the origin' to generate a Pareto tail in earnings.

More broadly, 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.

2 Model

The labor market consists of a continuum of niches which workers can inhabit. The earnings of a worker in a niche is determined by the number of workers in the worker's niche and the best paid niches are therefore the ones with very few workers. The model does not take a stand on how the distribution of workers over niches is generated.

[1] See, e.g., Nirei and Aoki (2016), Toda and Walsh (2015), Gabaix et al. (2016), Jones and Kim (2018), and Beare and Toda (2022).

2.1 Setting

Notation Workers belong to different niches $x \in \mathbb{R}^n$. The number of workers in a particular niche x is given by the density function $f(x)$. All workers in a given niche earn the same amount, $y(x)$.

Earnings in a niche The earnings $y(x)$ of a worker in a niche x with mass of workers $f(x)$ is determined by an inverse labor-demand curve, $y(x) = D^{-1}(f(x))$. That is, the earnings in a niche is determined by the number of workers inhabiting the niche. The inverse labor-demand curve D^{-1} is shared across niches and satisfies the following properties:

Assumption 1. *The inverse labor-demand curve $D^{-1} : \mathbb{R}_+ \to \mathbb{R}_+$ is differentiable and strictly monotonically decreasing. The limit elasticity of income exists when the mass of workers L approaches zero,*

$$\lim_{L\to 0} \frac{-(D^{-1})'(L)L}{D^{-1}(L)} = \frac{1}{\varepsilon}.$$

This is equivalent to assuming that, locally near zero, earnings are determined by a constant-elasticity labor-demand curve, $y \propto L^{-1/\varepsilon}$. The labor-demand curve can for example be derived from a CES production function $Y = \left(\int f(x)^{(\epsilon-1)/\epsilon} dx\right)^{\epsilon/(\epsilon-1)}$ which implies that the marginal product of a worker in niche x is given by $y(x) = \left(\frac{f(x)}{Y}\right)^{-1/\epsilon}$.

The space of niches Although the 'true' space of niches may be high dimensional, without loss of generality, we order the niches from most common to least common on the real non-negative half-line $\mathbb{R}_+$.[2]

The distribution of workers over niches We make some relatively weak assumptions on the distribution of workers over niches, f. The ordering of niches implies that f is monotonically decreasing. To have an unbounded income distribution, we therefore assume that for all $\epsilon > 0$, there exists a niche x such that $0 < f(x) < \epsilon$. Finally, we assume that f is regular in the sense of extreme-value theory.

Assumption 2. *The distribution of workers over niches, f, satisifies the following properties.*

1. *The density of workers $f : \mathbb{R}_+ \to \mathbb{R}_+$ is strictly monotonically decreasing.*
2. *For all $\epsilon > 0$, there exists a niche x such that $0 < f(x) < \epsilon$.*

[2] To be precise, let X be the space of niches and let $f^*(t) = \inf\{s > 0 : \mu(\{x \in X : f(x) > s\}) \le t\}$ be the *decreasing rearrangement* of f as in section 1.4.1 of Grafakos (2014). The densities f and f^* have the same mass of workers in niches with at least y workers, so f and f^* induce the same earnings distribution. Further, f^* inherits continuity from f if X is connected (continuity follows from f^* being monotone and surjective on the range of f, which is an interval because of connectedness).

3. *The density of workers f is regular in the sense that the following limit exists,*

$$\lim_{x \to \overline{x}} \frac{\partial}{\partial x} \frac{1-F(x)}{f(x)} = \xi, \tag{1}$$

where F is the cumulative distribution function corresponding to the density function f and $\overline{x} = \sup_x\{f(x) > 0\} \in \mathbb{R}_+ \cup \{\infty\}$ denotes the supremum of the support of f. The limit ξ satisfies $\xi > -1$.

Remark 1. *The regularity condition given by Equation* (1) *stems from extreme-value theory and corresponds to Equation (9) of Gabaix and Landier (2008). With $\overline{x} = \infty$, it implies the von Mises condition $\lim_{x\to\infty} \frac{1-F(x)}{xf(x)} = \xi$, as in, e.g., Resnick (1987), section 1.4.*

Remark 2. *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 $\xi = 0$.*

The regularity condition is satisfied by virtually all continuous distributions considered in economics and it is therefore a weak restriction on the density function. 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 $\xi = 1/\alpha$.

What is a niche? In effect, we have assumed that the demand for labor is constant across niches. This is without loss of generality since we are free to redefine the size of a niche in a way that equalizes the level of demand across niches. Let earnings be given by $y = D^{-1}(a(x)f(x))$. Reparametrize the space of niches so that $d\tilde{x} = a(x)dx$. The new density function is given by $\tilde{f}(x) = f(x)/a(x)$ (such that $\tilde{f}(x)d\tilde{x} = f(x)dx$) and $y(x) = D^{-1}(\tilde{f}(x))$.

2.2 The earnings distribution

A distribution of workers across niches f together with an inverse labor-demand function D^{-1} generates an earnings distribution with probability density function g and cumulative density function G. The earnings distribution g has a Pareto tail if

$$\lim_{y\to\infty} \frac{yg(y)}{1-G(y)} = \alpha.$$

We now state and prove the main result of the paper.

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

Proof. Write $y = D^{-1}(f(x))$. Since both $D^{-1}(\cdot)$ and $f(\cdot)$ are strictly monotonic, the correspondence is bijective and we have $1-G(y) = 1-F(x)$. Differentiating with respect to x yields $g(y)\cdot(D^{-1})'(f(x))\cdot f'(x) = f(x)$ or $g(y) = \frac{f(x)}{(D^{-1})'(f(x))f'(x)}$.

By differentiating, Assumption 2 implies that $\lim_{x\to\overline{x}} -\frac{(1-F(x))f'(x)}{f(x)^2} = 1+\xi$.

Therefore,

$$\lim_{y\to\infty} \frac{yg(y)}{1-G(y)} = \lim_{x\to\overline{x}} \frac{D^{-1}(f(x))g(y)}{1-F(x)} = \lim_{x_0\to\overline{x}} \frac{-f(x)^2}{(1-F(x))f'(x)} \frac{-D^{-1}(f(x))}{f(x)(D^{-1})'(f(x))} = \frac{\varepsilon}{1+\xi}.$$

□

3 Discussion

Although Theorem 1 covers both the case where the distribution of workers over niches f has bounded support ($\xi < 0$) and where the distribution of workers over niches features a fat tail ($\xi > 0$), it is instructive to consider the case of a thin-tailed distribution with unbounded support ($\xi = 0$) such as the the normal distribution, the log-normal distribution, and the exponential distribution. For these distributions, Theorem 1 gives the sharp result that the Pareto tail parameter is equal to the elasticity of labor demand, independent of the particular distribution of workers across niches. Figure 1 shows how the earnings distributions generated by an exponential, a normal and a log-normal distribution of workers over niches all generate asymptotic Pareto tails. Although the log-normal distribution has, loosely speaking, many more outliers than the normal distribution, top-earnings inequality as captured by the Pareto coefficient is the same under both distributions of workers across niches. This result captures a *Stiglitz effect*: reducing the number of top earners (e.g., hedge fund managers) increases the earnings of the remaining top earners, leaving equilibrium top inequality intact. de Vries and Toda (2022) document a large variation in labor income Pareto tail exponents across 52 countries. Through the lens of the model, this suggests that countries feature labor-market institutions or technology which generate labor-demand curves of different slopes. de Vries and Toda (2022) instead rationalize their findings with heterogeneous returns to career advancement. My theory and theirs are not mutually exclusive, career advancement can be interpreted as reaching a niche with few other workers and countries with high top inequality are those where such advancement yield high rewards.

The model introduced in this paper makes structural assumptions on the demand side of the economy but remains vague with respect to the supply side of the economy. If labor supply is determined in a frictionless fashion that implies wage equalization, then the density of workers is either uniform or degenerate, violating Assumption 2. It is therefore important that there is some hetereogeneity or friction inducing a non-degenerate non-uniform density of workers over niches. For a concrete structural model, see Heathcote et al. (2017). By assuming an exponential distribution in innate ability, Heathcote et al. (2017) generate an exact Pareto tail. This paper shows that the earnings distribution features a Pareto tail for a generic distribution

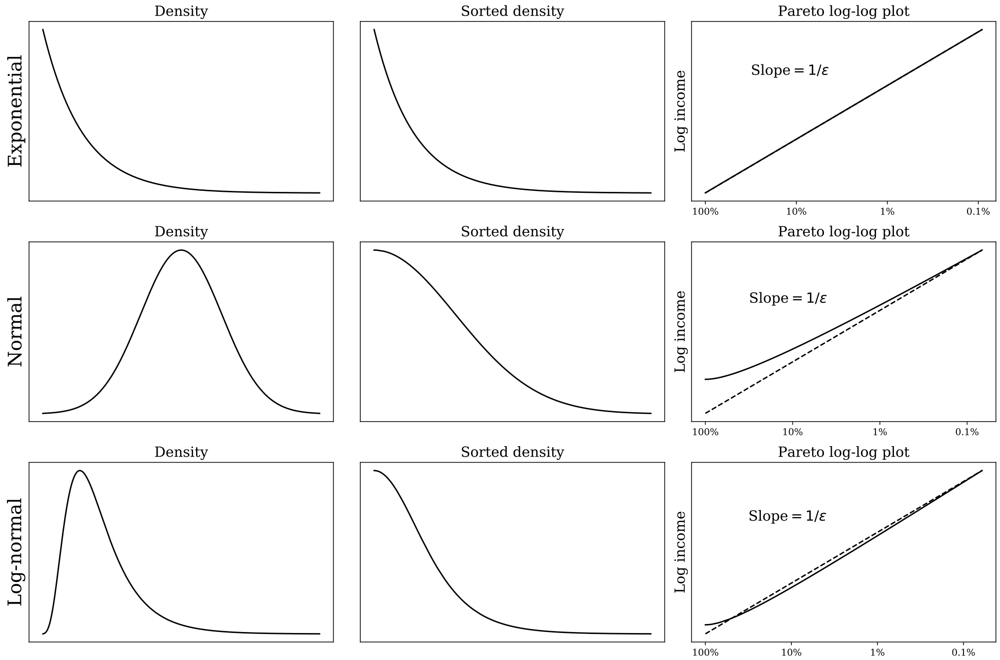

Figure 1: A constant-elasticity demand curve, $y = L^{-1/\varepsilon}$, generates an exact Pareto distribution of income when workers are distributed according to an exponential distribution. For other distributions such as the normal and log-normal distribution, the income distribution asymptotically approaches a Pareto distribution. The sorted density is the *decreasing rearrangement* of the density, see Footnote 2.

of innate ability, not only for the analytically convenient exponential distribution.

Ultimately, any theory of the Pareto property of top earnings needs to rely on some assumptions. The functional-form assumption necessary for Theorem 1 is the assumption that a stable elasticity of the labor-demand curve exists. Mathematically, the earnings distribution inherits its power-law functional form from the power-law functional form of labor demand.

References

Beare, B.K., Toda, A.A., 2022. Determination of Pareto exponents in economic models driven by markov multiplicative processes. Econometrica 90, 1811–1833. doi:10.3982/ecta17984.

Champernowne, D., 1953. A Model of Income Distribution. Economic Journal 63, 318–351. doi:10.2307/2227127.

Gabaix, X., 2009. Power Laws in Economics and Finance. Annual Review of Economics 1, 255–294. doi:10.1146/annurev.economics.050708.142940.

Gabaix, X., Landier, A., 2008. Why has CEO pay increased so much? Quarterly Journal of Economics 123, 49–100. doi:https://doi.org/10.1162/qjec.2008.123.1.49.

Gabaix, X., Lasry, J.M., Lions, P.L., Moll, B., 2016. The Dynamics of Inequality. Econometrica 84, 2071–2111. doi:10.3982/ecta13569.

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

Grafakos, L., 2014. Classical Fourier Analysis. Springer New York. doi:10.1007/978-1-4939-1194-3.

Heathcote, J., Storesletten, K., Violante, G.L., 2017. Optimal tax progressivity: An analytical framework. Quarterly Journal of Economics 132, 1693–1754. doi:10.1093/qje/qjx018.

Jones, C.I., Kim, J., 2018. A Schumpeterian model of top income inequality. Journal of Political Economy 126, 1785–1826. doi:10.1086/699190.

Kesten, H., 1973. Random difference equations and Renewal theory for products of random matrices. Acta Mathematica 131, 207–248. doi:10.1007/BF02392040.

Nirei, M., Aoki, S., 2016. Pareto distribution of income in neoclassical growth models. Review of Economic Dynamics 20, 25–42. doi:10.1016/j.red.2015.11.002.

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

Resnick, S.I., 1987. Extreme Values, Regular Variation and Point Processes. Springer New York. doi:10.1007/978-0-387-75953-1.

Simon, H.A., 1955. On a Class of Skew Distribution Functions. Biometrika 42, 425. doi:10.2307/2333389.

Sornette, D., 2006. Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization and Disorder: Concepts and tools. Springer Science & Business Media. doi:10.1007/3-540-33182-4.

Toda, A.A., Walsh, K., 2015. The double power law in consumption and implications for testing Euler equations. Journal of Political Economy 123, 1177–1200. doi:10.1086/682729.

de Vries, T., Toda, A.A., 2022. Capital and labor income Pareto exponents across time and space. Review of Income and Wealth 68, 1058–1078. doi:10.1111/roiw.12556.