Paretian Distributions and Income Maximizations

Benoit Mandelbrot (1962)

Quarterly Journal of Economics, 76(1): pp. 57–85.

Econ 350, Winter 2019

The Empirical Distribution of Income: Scope of the Weak Law of Pareto

- It is “well-known” that the distribution of income follows the asymptotic (or “weak”) law of Pareto, according to which there exist two constants α and u^*, such that, for sufficiently large values of u,

$$1 - F(u) = \mathrm{Prob}(U > u) \sim (u/u^*)^{-\alpha}.$$

- *We shall suggest one possible reason why several values of the alpha exponent, and several non-Paretian categories, may well coexist in a single strongly interacting economic community*.
- Assume that each individual must choose *one* of N possible "occupations," $P_n (1 \leqslant n \leqslant N)$.
- Model of multinomial choice.

- Within the total population, the distributions of the various U_n are random.
- Assume a certain interdependence between the various offers, and *we shall show that, if the over-all income distribution is Paretian with the exponent α, the offers accepted from each "occupation" taken separately will also be Paretian, but with an exponent of the form $w(n)\alpha$.*
- The quantity $w(n)$ is an *integer* designated as the *"weight"* of P_n; it turns out to be equal to the *number of factors that must be large simultaneously, in order that the offer U_n be accepted and turn out to be large*.
- The concept of weight is central to this analysis.
- If the weight is large, we shall show that the law of Pareto becomes practically useless and should be replaced by the lognormal distribution.

Linear Factor Analysis of the Rental Price of an Undissoluble Bundle of Abilities

- $U_n (1 \leqslant n \leqslant N)$ is the rental price which the occupation P_n is ready to pay for the use of a man's abilities.
- Can be written as a nonhomogeneous *linear* form of F independent *factors* $V_f (1 \leqslant f \leqslant F)$, each of which is randomly distributed in the population, and "measures" one or several "abilities."
- Then one can write:

$$U_n = \sum_{f=1}^{F} a_{nf} V_f + a_{n0} + a_n E_n.$$

The Regions of Acceptance of the Different Offers

- The prospective employee's problem is to determine the n that maximizes the nonlinear function U_n of n.
- Let the random variable W_n designate the incomes *accepted* from the occupation P_n, and
- Let W designate all accepted incomes (observed income distributions).
- The region of acceptance R_n of the offer U_n will be the intersection of the $N-1$ half-hyperspaces defined by the equations:

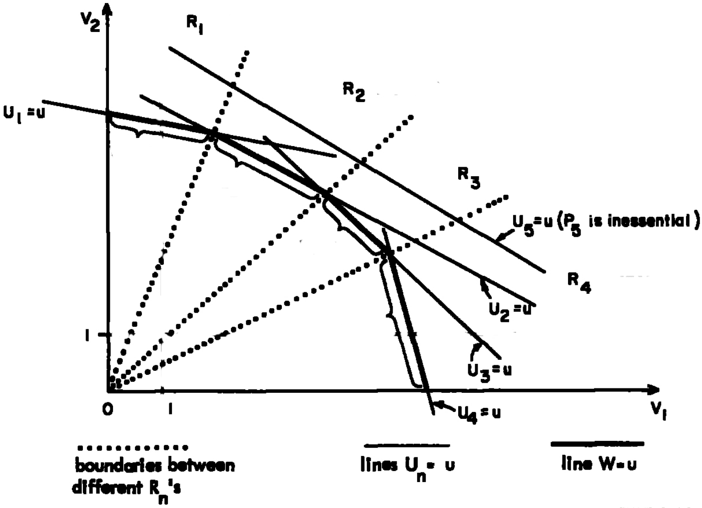
$$U_n - U_i \geqslant 0 \qquad\qquad (i \neq n).$$

Examples

Homogeneous offers.

- The simplest case, from the viewpoint of R_n and of income iso-surfaces, occurs when the offers are homogeneous forms of the V_f.
- Then, the accepted P_n will obviously depend only upon the relative values of the factors, that is upon what we may call the "factor profile" of the individual.
- For the case $F = 2$, the typical shape of the regions R_n and of the surfaces $W = u$, is given in Figure 1.

Figure 1: Example of Regions of Acceptance in the Case of Two Homogeneous Factors



- Figures 2a to 2e are such that, if one factor is present in very large quantity, it fully determines the chosen occupation, so that changes in the values of other factors are irrelevant.
- But in Figure 2f, we see a case where the occupation is not determined by the value of the high factor V_1 alone, and is greatly influenced by V_2 and V_3, which are present in small quantities; small changes in those other factors can totally modify the value of n that will be chosen.

Figure 2: Six Miscellaneous Examples of the Shape of the Surface $W = u$, in the Case of Three Homogeneous Factors

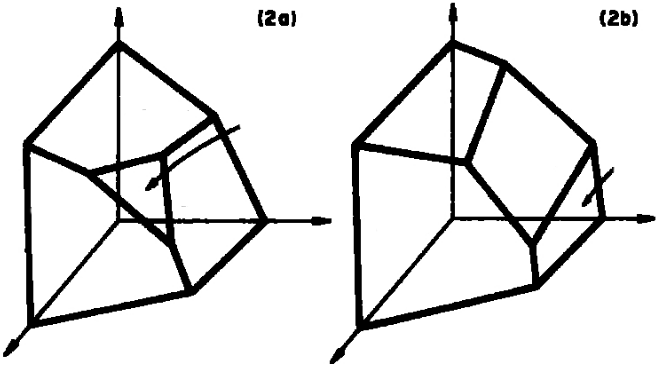
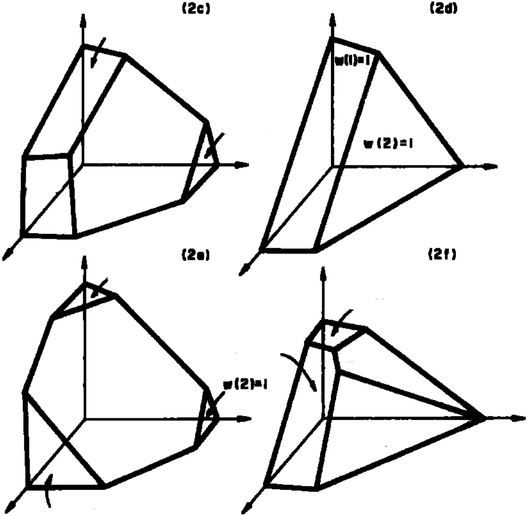


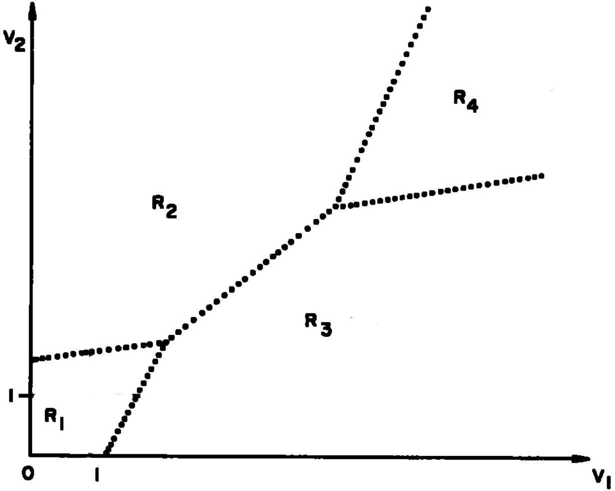
Figure 2: Six Miscellaneous Examples of the Shape of the Surface $W = u$, in the Case of Three Homogeneous Factors



Linear and nonhomogeneous offers.

- In this case, the regions of acceptance R_n need no longer be open cones.
- Some of these regions may be closed, leading to bounded incomes.
- The corresponding occupations will clearly be those which are not very sensitive to the values of the factors.
- On the contrary, fanning-out regions R_n are quite sensitive to the values of the factors.
- An example is given in Figure 3.

Figure 3: Example of Regions of Acceptance in the Case of Two Inhomogeneous Factors



The Assumption That the Factors Themselves are Paretian

- The fundamental – because least obvious – "input" of the theory is the assumption that the probability distribution of each factor in the population is **weakly Paretian**, in the sense that there exist constants u_f^* and $\alpha > 1$ such that, when u increases to infinity,

$$1 - F_f^0(u) = \mathrm{Prob}(V_f > u) \sim (u/u_f^*)^{-\alpha}.$$

- It is possible to prove the following:
- **Assume that offers can be factor-analyzed linearly, that an occupation is chosen by income maximization and that the $1 - F_f^0(u\, u_f^*)$ have the same asymptotic behavior for all the factors.**
- *Then*, in order that the over-all income distribution be Paretian, the same must be true of every factor taken separately.
- The Paretian character of every factor will be treated as one of the "axioms" and need not be discussed any further.

- Moreover, two special weakly Paretian distributions will be needed.
- One is the *strong law of Pareto*, which will considerably simplify the mathematical arguments, although it is unquestionably not a very good representation of the facts.
- This law states that $1 - F(u) = (u/u^*)^{-\alpha}$ for $u > u^*$ and $F(u) = 0$ for $u < u^*$.

THE UNIVERSITY OF CHICAGO

- The other special case is our "law of Pareto and Levy."
- Its density cannot be expressed in closed analytic form and had to be computed by numerical methods from the bilateral generating function of its density $p_\alpha(u)$, which is

$$G(s) = \int_{-\infty}^{\infty} \exp(-su) p_\alpha(u) = \exp\left[-Ms + (ss^*)^\alpha\right],$$

- Scalar M: the mean value of U and α is contained between 1 and 2.

Surfaces of constant probability.

- Examine first the strong Pareto case.
- Since it is considered only for the sake of mathematical simplicity, and since the parameters a_{nf} take care of the scale of the factors, we can suppose that the u^* of all the factors are equal to one.
- The joint density function of the F factors then takes the form

$$\alpha^F (\Pi v_f)^{-(\alpha+1)} \Pi dv_f,$$

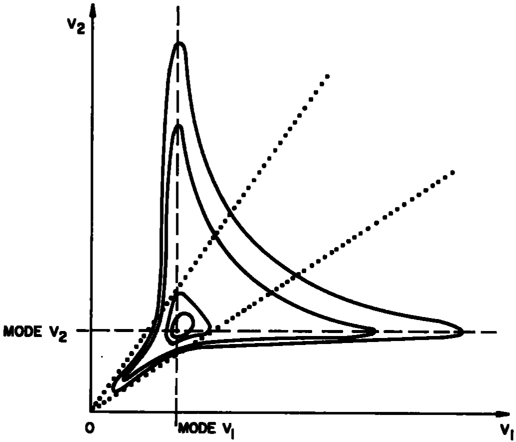
- and *the probability iso-surfaces are the hyperboloids*, of equations

$$\Pi v_f = \text{ constant},$$

- *truncated to the region in which all the v_f are greater than one*.

- Consider then the surface $\sum_{f=1}^{F} v_f a_{nf} = u_n$.
- If the restrictions $v_f \geqslant 1$ are neglected, the point of coordinates u_n/Fa_{nf} is the "least probable" one on that surface, and it happens to be precisely the one where the amounts $v_f a_{nf}$, paid for the different factors, are equal to each other and to u_n/F.
- Examine now the case of weakly Paretian factors, for which the "tail" is preceded by some kind of a "bell."
- Then, the probability iso-surfaces take the form exemplified in Figure 4 for the case $F = 2$.

Figure 4: Example of Probability Iso-Lines in the Case of Two Factors, When the Probability Density Curve of Each Factor is Asymptotically Paretian But Starts by a "Bell"



- If u_n is sufficiently large, there is still a "least probable" point on the surface $\sum_{f=1}^{F} v_f a_{nf} = u_n$.

The Conclusion: The Offers Accepted from Different Occupations Will Be Paretian, with an Alpha Exponent Equal to Integral Multiple of That of the Factors

- As previously stated, whichever the factor loadings, the variables W_n, which represent the *offers accepted* from the professions P_n, will either be bounded or will be such that

$$\text{Prob}(W_n > u) \sim (u/u_n)^{w(n)\alpha}.$$

- The weight $w(n)$ will be shown in the following sections to be equal to the *smallest number of different factors that must be large simultaneously, if W_n is to be large*.
- Could be one definition of task complexity.

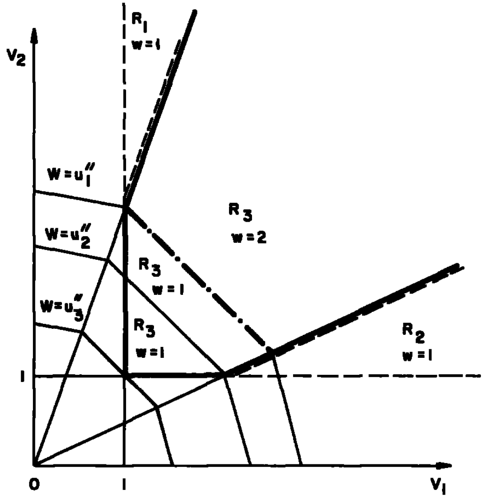
The Evaluation of the Asymptotic Paretian Weight of an Occupation

The Behavior of the Densities Before the Paretian Asymptotic Range is Reached; The Lognormal Approximation

- In this case, no income less than u_3'' is *ever* accepted.
- Let now $u_3'' < u < u_2''$; one sees on the figure that all such incomes come from P_3 and that they are in no way influenced even by the existence of the other two occupations.
- Since U_3 is a weighted sum of two strong Paretian variables, it is weakly Paretian and it may very well happen that its asymptotic behavior begins to apply before the influence of P_1 and P_2 begins to be felt.

Figure 5: Example of Successive Appearance of Occupations, as u Increases

(a)



- Suppose now that $u_2'' < u < u_1''$.
- Then, the accepted income will come from either P_3 or P_2; in this case, there are two occupations and two factors and both accepted income distributions W_3 and W_2 will have weight 1.
- Finally, if $u > u_1''$, one may accept any of the three offers, and the occupation P_3 is thereby finally relegated to a zone that does not contain either of the co-ordinate axes, so that its weight reaches its asymptotic value 2.
- In the more realistic case where the Paretian "tail" is preceded by some kind of a "bell," the above results are slightly modified: the sharp corners characteristic of the strong law of Pareto having disappeared, the above transitions occur slowly and gradually as is shown on Figures 5b and 5c.

THE UNIVERSITY OF CHICAGO

Figure 5: Example of Density Curves for Different Occupations in Bilogarithmic Co-ordinates

(b)

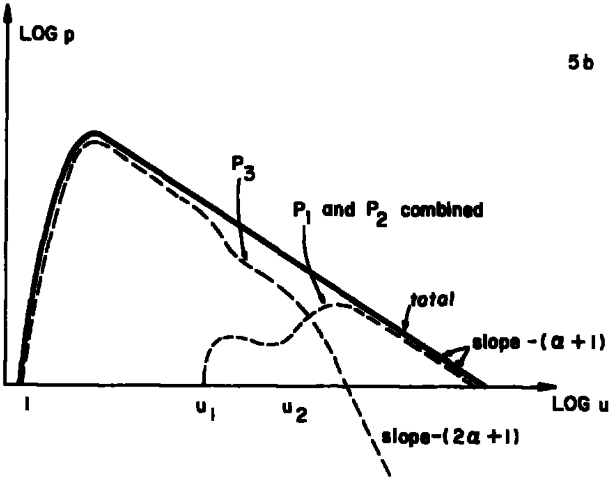


Figure 5: Example of Density Curves in Ordinary Co-ordinates

(c)

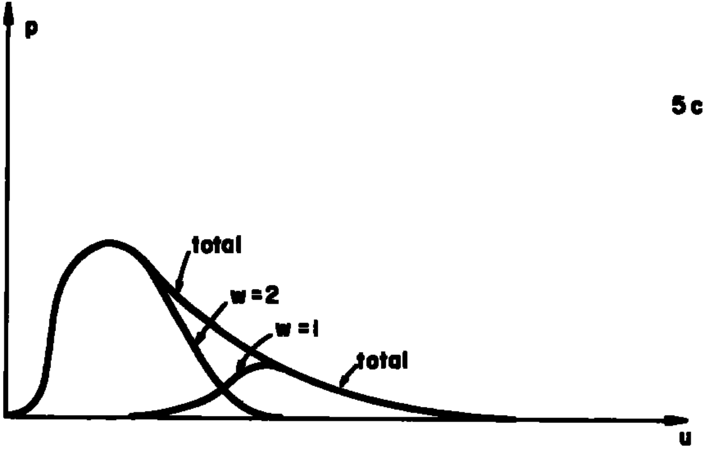
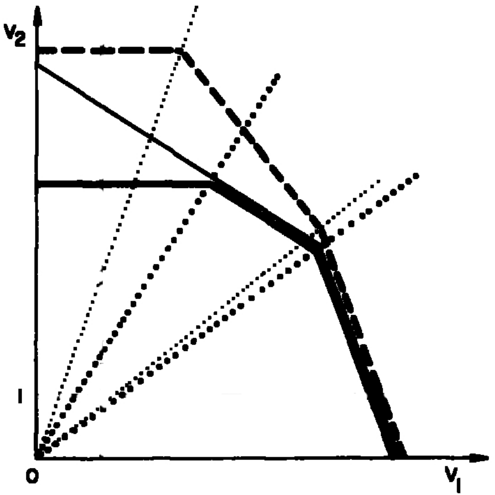


Figure 6: Example of the Change in the Form of the Curve $W = u$, as Factor Loadings Are Modified



solid bold line: situation before the changes in the factor loadings

solid thin line: situation after a change in the factor loadings of U_1 (which thereby becomes inessential)

bold dashed line: situation after the changes in the factor loadings of U_1 and U_2