

The incorporation of Pareto's Law into financial modelling: the 1962 turn

Christian WALTER

Ethics & Finance Chair, FMSH

ISJPS | UMR 8103 CNRS-Université Paris 1 Panthéon-Sorbonne

50th Annual Meeting of the History of Economics Society, 25 June

Contents

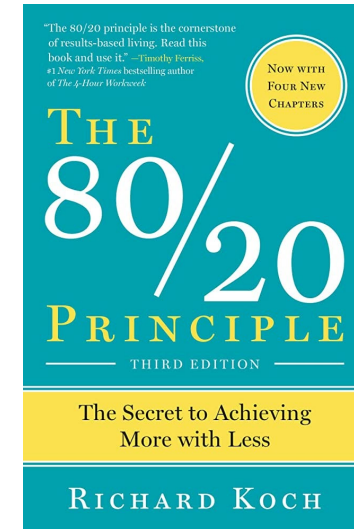
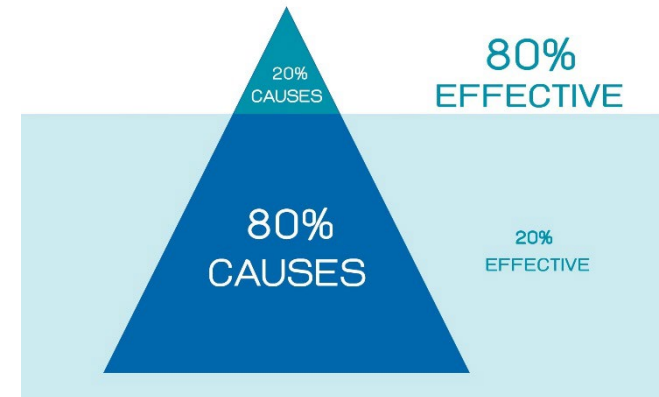
1. The Paretian framework
 - Two innovative ideas of Vilfredo Pareto
2. The Mandelbrot paper of 1959 (1962)
 - Paretian distributions and income maximization
3. The Mandelbrot paper of 1962 (1963)
 - The variation of certain speculative prices

« L'introduction de la loi de Pareto dans la modélisation financière », *L'Année sociologique*, vol. 73, no. 1, 2023, pp. 113-150. DOI : 10.3917/anso.231.0113

The 80-20 rule

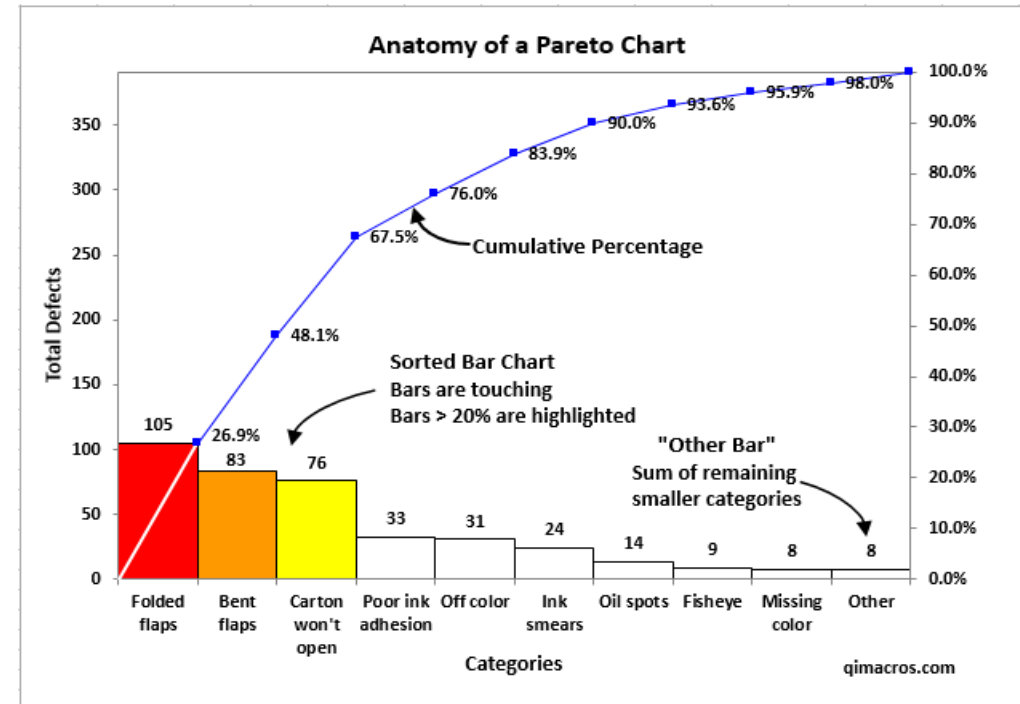
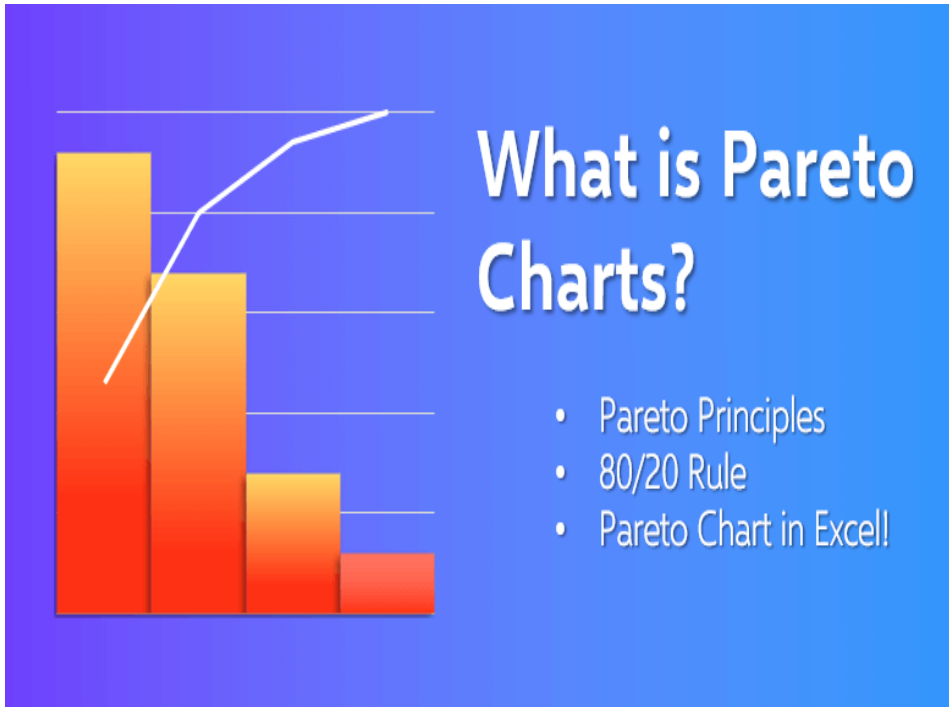
- Pareto originally used this distribution to describe the allocation of wealth among individuals since it seemed to show rather well the way that a larger portion of the wealth of any society is owned by a smaller percentage of the people in that society.
- He also used it to describe distribution of income. This idea is sometimes expressed more simply as the **Pareto principle** or the “80-20 rule” which says that 20% of the population controls 80% of the wealth.
- Different values of the coefficients in a Pareto distribution will produce a “90-10 rule” or a “70-30 rule.”

PARETO PRINCIPLE



Pareto charts

A Pareto chart is a **bar graph**. The lengths of the bars represent frequency or cost (time or money) and are arranged with longest bars on the left and the shortest to the right. In this way the chart visually depicts which situations are more significant.



Examples of Pareto charts

Figure 3. Frequency distribution of individual performance (O'Boyle and Aguinis, 2012)

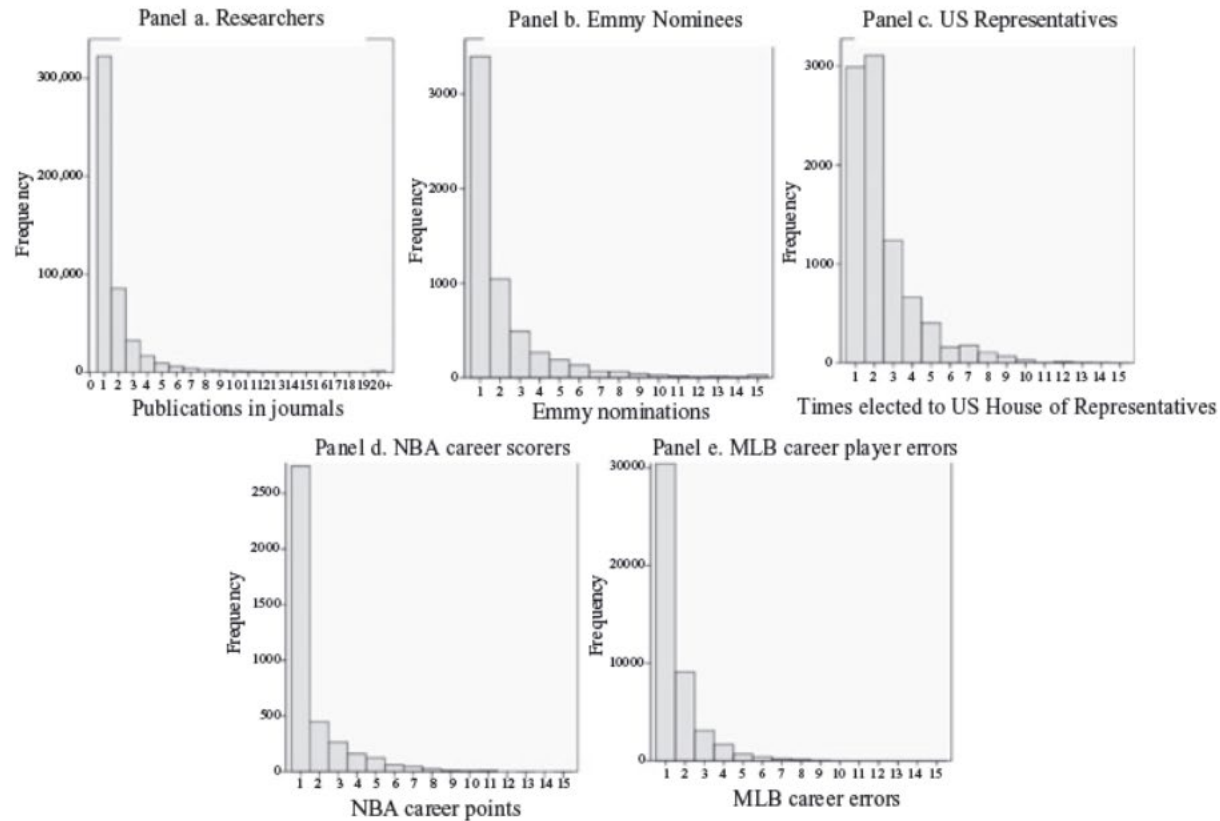


Figure 2: Distribution of Individual Performance for Researchers ($N = 490,185$), Emmy Nominees ($N = 5,826$), United States Representatives ($N = 8,976$), NBA Career Scorers ($N = 3,932$), and Major League Baseball (MLB) Career Errors ($N = 45,885$).

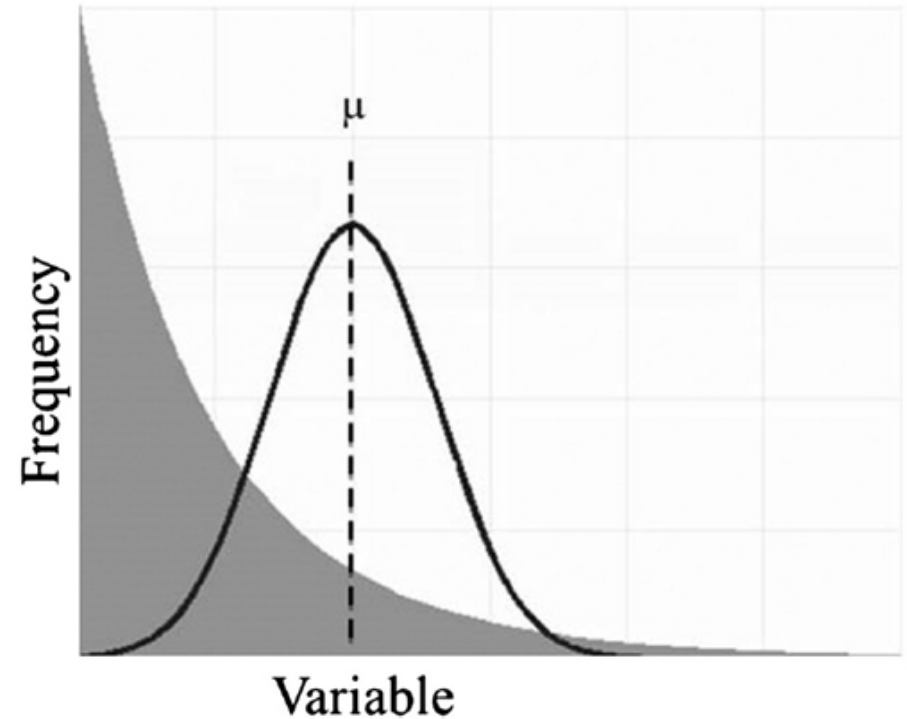
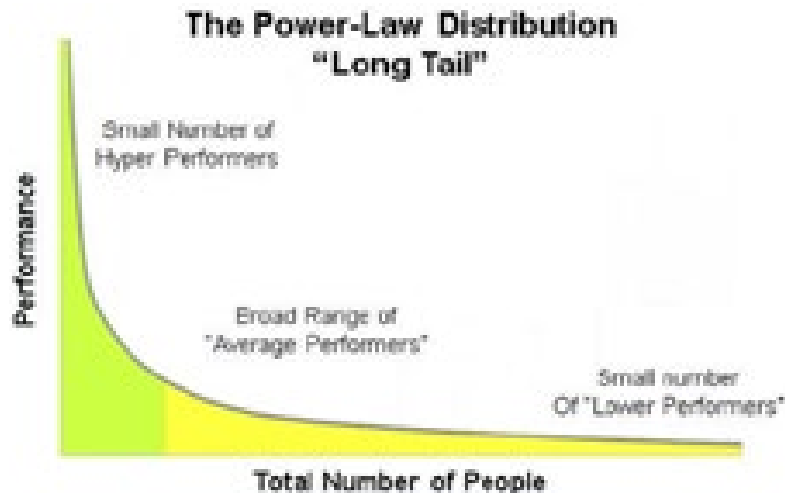
Note. For all Y axes, “Frequency” refers to number of individuals. For clarity, individuals with more than 20 publications (Panel a) and more than 15 Emmy nominations (Panel b) were included in the last bins. For panels c–e, participants were divided into 15 equally spaced bins.

Power-law “long tail”

L-curves vs. bell curves

Power law distributions are sometimes called L-curves to contrast with the bell curves associated with normal distributions, as depicted in the frequency distribution

A power law is a function f_a defined by $f_a : x \rightarrow x^a$ where a is the exponent of the power function. The distribution tails of a power law function are extremely long and decrease very slowly, unlike the distribution tails of the normal distribution, which are short and decrease rapidly as an exponential of the form $\exp(-x^2/2)$.



A first Pareto innovative idea

Seeking to model the right-hand side of the distribution curve (high incomes) as well as possible, Pareto's first idea, which was "innovative at the time" (Barbut, 1999: 89), was to consider not the income ranges according to their order of magnitude, which would mean plotting the histogram of relative income frequencies (an approximation of their density function), but the distribution of **cumulative frequencies** (CDF).

If X is a random variable representing the value of income (or wealth), by choosing to consider the cumulative distribution function and not the probability density function, Pareto is interested in $F_X(x) = \Pr(X < x)$ where "Pr" means "probability".

To obtain his "income curve", Pareto ranks income and cumulative frequency in increasing order, creating "rank-frequency" or "rank-size" curve, which is built by associating an income rank with the value of its cumulative frequency.

$$\text{Fr}(X_i \geq X_{(k)}) = k/n$$

Rank-size distribution is the distribution of size by rank, in decreasing order of size

A second Pareto innovative idea

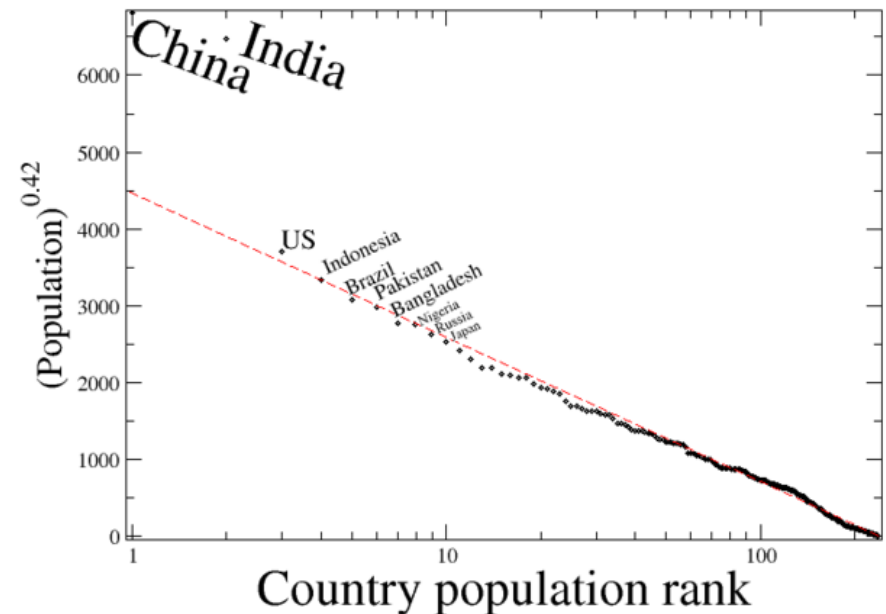
Pareto's second idea was to represent the distribution function on **bilogarithmic plot**, a graph in logarithmic coordinates of the pairs $(X(k), k)$, the set of data $(\ln X(k), \ln k)$: “an engineer's idea” (Barbut, 1999: 89). This bilogarithmic plot shows, beyond a threshold x_0 , a negatively sloped line with coefficient alpha

$$\ln k = -\alpha \ln X_{(k)} + C_+$$

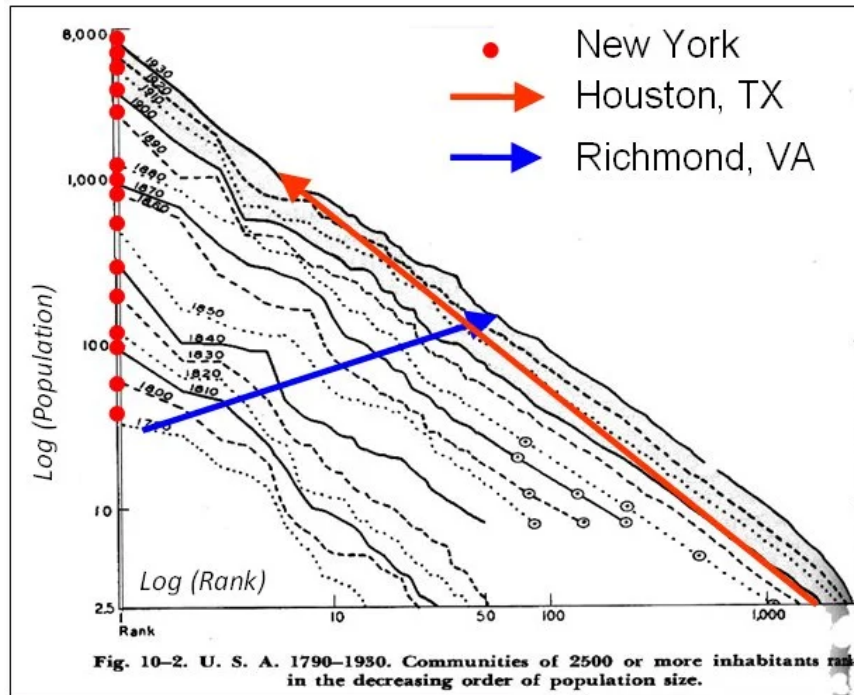
Putting $x_0 = (C_+/\alpha)^{1/\alpha}$, the previous relation is written:

$$\bar{F}(x | X > x_0) = \left(\frac{x_0}{x}\right)^\alpha$$

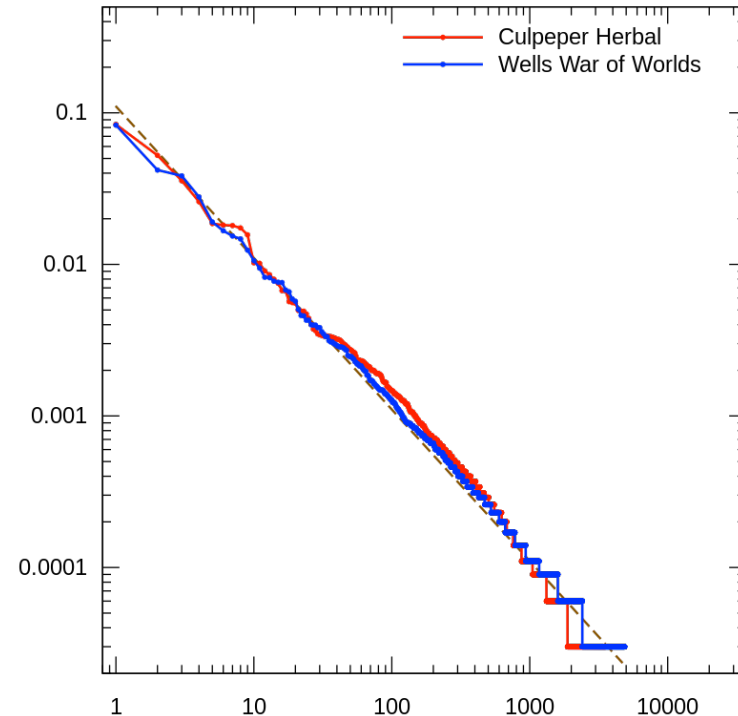
The tail function (denoted \bar{F}) is a **power law**.



Examples of Rank-size plots



From George Kingsley Zipf (1949) *Human Behavior and the Principle of Least Effort* (Addison-Wesley, Cambridge, MA)



A plot of the frequency of each word as a function of its frequency rank for two English language texts: [Culpeper's Complete Herbal](#) (1652) and [H. G. Wells's The War of the Worlds](#) (1898) in a [log-log](#) scale

The **rank-size rule** (or **law**) describes the remarkable regularity in many phenomena, including the distribution of city sizes, the sizes of businesses, the sizes of particles (such as sand), the lengths of rivers, the frequencies of word usage, and wealth among individuals

Mandelbrot's 1959 paper on income

In 1959, Mandelbrot addressed Pareto's issue of income and published two *Comptes Rendus à l'Académie des Sciences* (sessions of 3 August and 23 November 1959) on the modelling of income distributions (Mandelbrot, 1959a, 1959b). These two papers radically changed the way we understand long-tailed distributions.

The starting point of Mandelbrot's thought process is to try to overcome the difficulties arising from the two limitations of Pareto's law.

1. On the one hand, choosing this law means having to split the distribution of incomes into two distinct regimes, high incomes and others.
2. On the other hand, unlike the normal distribution, one of whose main interests stems from its probabilistic status as a limit distribution (existence of the centred limit theorem in its classical form), the Pareto distribution did not appear, in the opinion of the economists of the time, to have a similar theoretical justification in probability theory. Whereas the Gaussian distribution can be obtained by applying the centred limit theorem, the Pareto distribution did not seem to have the probabilistic status of a limit distribution.

Mandelbrot's 1959 breakthrough

On 3 August 1959, Mandelbrot proposed representing the entire distribution of incomes by a totally asymmetric Lévy alpha-stable probability distribution with a finite mean value, which represents a limit in probability for a generalised law of large numbers. A model that removes the two obstacles of Pareto's law in one stroke.

This proposal is revolutionary.

Unlike the normal distribution and the main probability distributions known by their density function, an alpha-stable distribution has no known density function. To describe it, we use a transformation in Fourier space to obtain what is known as the characteristic function, which is the Fourier transform of the density.

Whatever a given probability law, the characteristic function always exists, whereas the density does not always exist.

Mandelbrot describes the distribution of incomes by the Fourier transform of their density, specifying the form of its characteristic function

Lévy distribution with Paretian tails

The characteristic function used by Mandelbrot is that of the characteristic exponent of an alpha-stable distribution. The Fourier transform of the density of an alpha-stable distribution is generally represented by its characteristic exponent, the expression for which is known. This is the so-called “Lévy Khintchine” representation.

With an alpha-stable distribution, the Lévy measure $\nu(x)$ is written as a Pareto distribution of type I :

$$\nu(x) = \frac{C_-}{x^{1+\alpha}} \quad \text{if } x < 0$$
$$\nu(x) = \frac{C_+}{x^{1+\alpha}} \quad \text{if } x > 0$$

The two different values C_- and C_+ reflect the “size” of the distribution tails on the left and right. The alpha parameter gives the “shape” of the distribution. When alpha = 2, the distribution is normal. When alpha = 1, it is a Cauchy distribution. The parameter alpha is the Pareto exponent. Since the Lévy measure is Paretian, and the Lévy measure defines the “long-term” behaviour of the random process, alpha-stable distributions are asymptotically Paretian with the tail index alpha.

Moving to finance: paretian chance

An alpha-stable distribution is also a limit law, and, in this regard, alpha-stable laws solve the problem of the domains of attraction of laws such as Pareto's law. In this respect, **Pareto's law acquires the status of a limit law**, and Pareto can describe his law as a new "law of errors" (Barbut, 1991a: 42). But a law of errors governed by mechanisms other than the normal law.

In the same way that Gauss' law of errors defines a "Gaussian randomness" (Gaussian random draws), it is possible to define with Pareto's law a "Paretian randomness" (Paretian random draws) that Mandelbrot introduced in 1962.

Finally, if we consider the sequence of successive incomes over time, the different distribution laws, stable by addition, represent an alpha-stable Lévy process (Lévy, 1937) or a continuous "random walk". Starting with the analysis of income distribution, Mandelbrot introduced the probabilistic tool that would enable him to tackle the question of stock price variation. With the transition from the alpha-stable laws of income to stock prices, Pareto's law entered financial modelling.

Random walk model of stock prices

Mandelbrot's 1962 paper

The first evidence of the paretian nature of fluctuations in stock market prices appeared in the now-classic study of variations in the price of cotton between 1880 and 1958. This study, part of an internal IBM research report, was published in reduced form in the *Comptes rendus à l'Académie des Sciences*, session of 4 June 1962 (Mandelbrot, 1962). It is the first explicit consideration of the existence of paretian phenomena in finance, based on the study of large price variations.

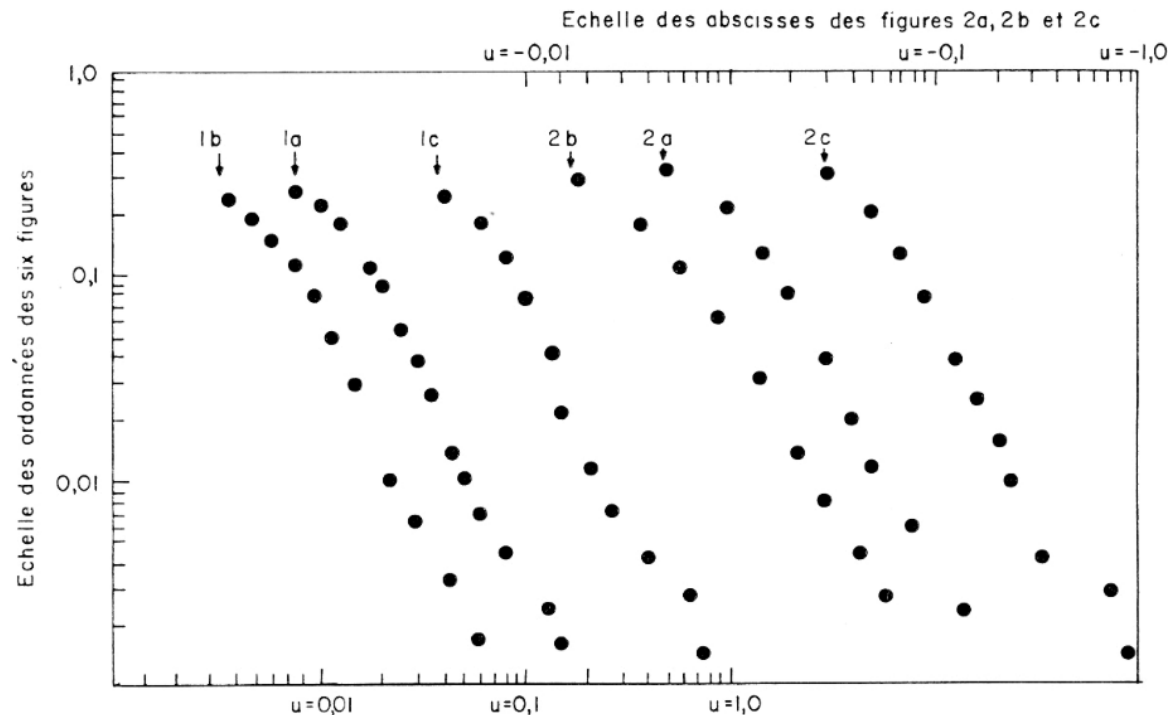
In his 1962 paper, Mandelbrot focused on variations in the first differences of the logarithms of the cotton closing prices over τ days. Given $S(t)$, the quoted price of cotton (a financial asset) at time t , the natural logarithm $X(t) = \ln S(t)$ represents the continuous rate of return on the asset quoted between 0 and t (with the convention $X(0)=1$). From this point of view, assuming that asset variations can be represented by a random variable that depends on time, then $X(t)$ is the stochastic process of the asset price and τ the “time scale” of the price variation analysis. Mandelbrot considers the formula :

$$X(t) - X(t - \tau) \stackrel{\text{def}}{=} Z(t, \tau)$$

Following in Pareto's footsteps

Mandelbrot chose two values for the scale τ : 1 month and 1 day. He then considers the tails of empirical distributions and the cumulative frequency of large empirical variations observed, noted Fr , by calculating the expressions $Fr(Z(t, \tau) > u)$ for large positive variations (hence the cumulative frequency of large positive variations) and $Fr(Z(t, \tau) < -u)$ for large negative variations (hence the cumulative frequency of large negative variations), exactly as Pareto did for incomes.

Cumulative frequencies can be adjusted by lines with slopes close to $\alpha = 1,7$



Stability-under-addition property

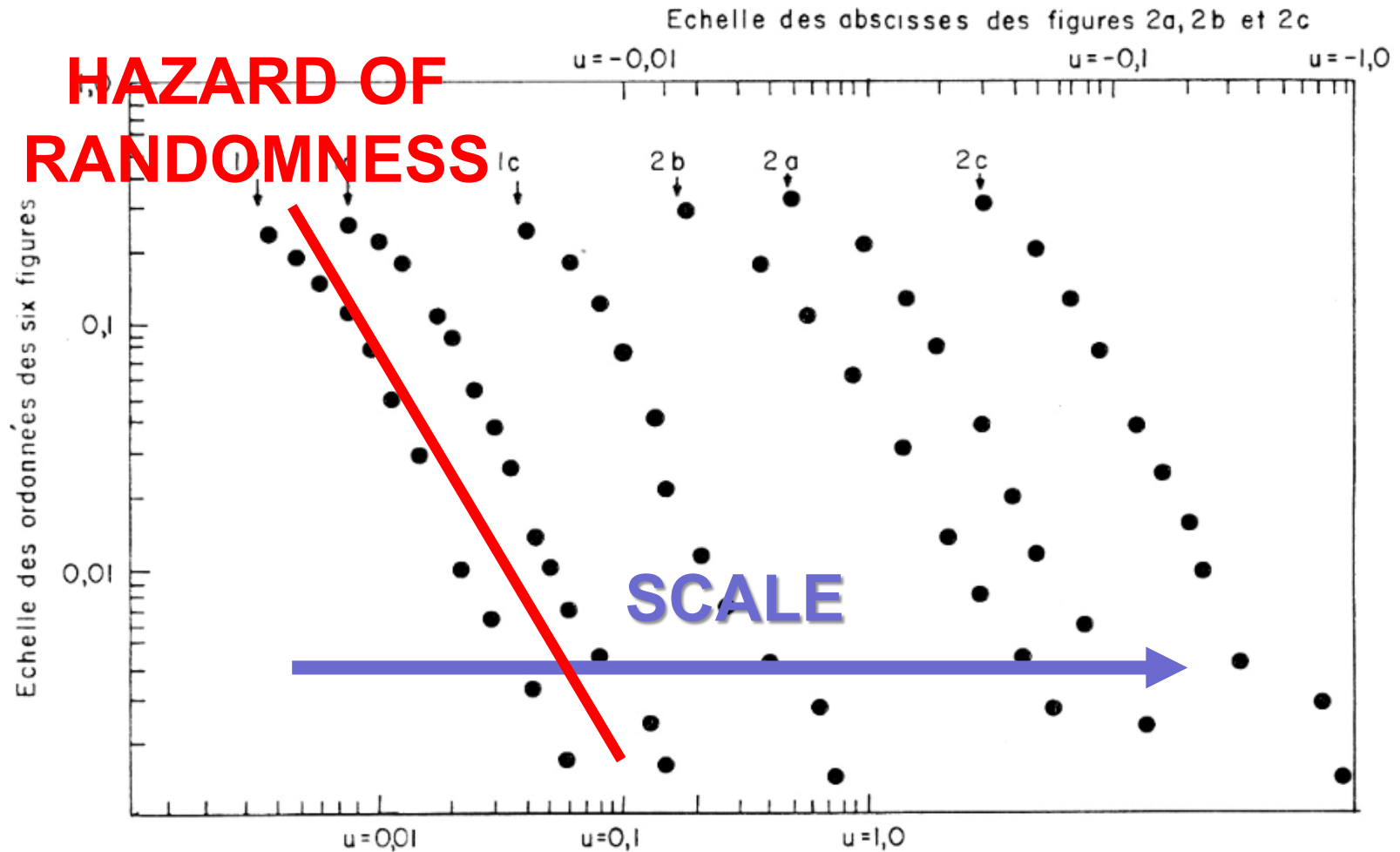
Mandelbrot also observed that the straight lines corresponding to the distributions of the daily and monthly time steps were parallel. From this he deduced that the distributions of the daily and monthly variations differ only by a change of scale.

In other words, the phenomenon is stable through a change of scale, or “**scale invariant**” (Mandelbrot, 1962: 3968).

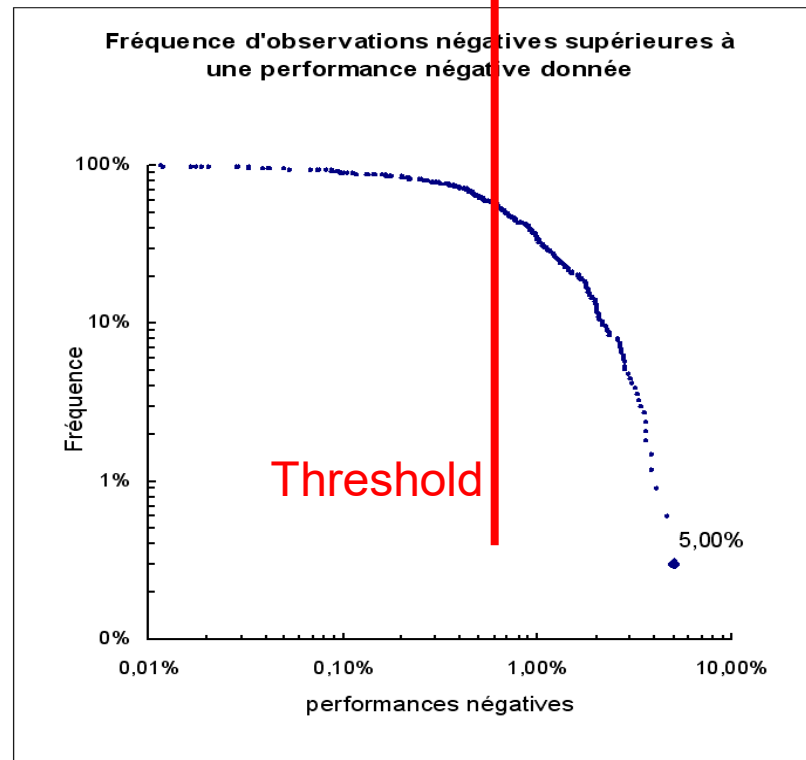
Mandelbrot then proposes his interpretation: the stability of the phenomenon between 1 day and 1 month can be understood as a trace of stability under addition in Lévy’s sense, and in this case the characteristic function should necessarily have the form of the characteristic function of an alpha-stable Lévy distribution.

In other words, the random price process $X(t)$ could be analysed as an alpha-stable Lévy random process

The Paretian chance



An example of Paretian chance



The empirical distribution function $\text{Fr}(Z(t, \tau) < -u)$ of the tail of the distribution of negative variations in the performance $Z(t, \tau)$ of a portfolio of equities in bilogarithmic plot. Above a certain threshold of negative performance (around -2%), the points are aligned on a straight line with a slope of $(-\alpha)$, indicating a power law for large negative values. This is like Mandelbrot's graph (1962). Source: (Walter, 2013).

Two kinds of extremes: beyond Pareto

Since the purpose is to describe and model high incomes (or high wealth), we also need to define a “threshold income” x_0 above which observations will be made, and Pareto’s approach involves looking at values of X above x_0 . If x_0 is too small, there will be too many average incomes among the selection of the “big ones”. If x_0 is too large, there will not be enough high incomes. The choice of threshold x_0 will therefore have an important influence on the description of the tails of the distribution and their adjustment by a mathematical function. Knowing which function best describes a distribution tail requires choosing a threshold x_0 that defines the “tail” part of the distribution under examination. A “tail” cannot be defined without first defining the threshold x_0 at which it starts.

There are two possible models for the Pareto approach. Either consider the highest values of X above the threshold x_0 and, in this case, introduce the random variable $\{X \mid X > x_0\}$ (where the symbol “|” means “such as”) and try to model its extremes, or consider the differences in income in relation to the threshold income, the surpluses $X - x_0$ and, in this case, examine the random variable $\{X - x_0 \mid X > x_0\}$ and try to model the threshold surpluses. These two approaches lead to two different limit laws.

Extreme value theory

Conclusion

After Mandelbrot's 1962 paper, Pareto's law became :

- A limit law in probability (Lévy)
- A new “law of errors” : paretian randomness
“wild chance”