



HAL
open science

L'introduction de la loi de Pareto dans la modélisation financière

Christian Walter

► **To cite this version:**

Christian Walter. L'introduction de la loi de Pareto dans la modélisation financière. *L'Année Sociologique*, 2023, Vilfredo Pareto (1848-1923), 73 (1), pp.113-150. 10.3917/anso.231.0113 . halshs-04494659

HAL Id: halshs-04494659

<https://shs.hal.science/halshs-04494659>

Submitted on 8 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

L'introduction de la loi de Pareto dans la modélisation financière

Christian WALTER¹

Résumé

La loi de Pareto entre dans la modélisation financière au début des années 1960 grâce à la découverte de deux propriétés de cette loi, non perçues auparavant : c'est une loi d'invariance par changement d'échelle, c'est une loi limite en probabilité. Ces deux propriétés de la loi de Pareto sont introduites par Benoît Mandelbrot qui cale la loi de Pareto sur les lois alpha-stables de Paul Lévy dans les modélisations des prix du coton puis des actifs financiers. Cette insertion de l'ancienne loi de Pareto dans la modélisation des dynamiques boursières change son statut et lui confère une nouvelle existence dans l'économie financière, en injectant des lois de puissance dans les modèles de prix. Après avoir rappelé les propriétés mathématiques du cadre parétien, en particulier le paradoxe de la valeur probable d'une loi de puissance, on présente le débat sur la modélisation entre une combinaison de modèles (gaussien pour les valeurs moyennes et parétien pour les valeurs extrêmes) et un modèle unique pour l'ensemble des valeurs (alpha-stable), en introduisant la théorie des valeurs extrêmes et la loi de Pareto généralisée. Puis on présente l'arrivée en finance de la loi de Pareto en suivant la démarche de Mandelbrot, ce qui permet d'introduire la notion de hasard parétien. Enfin on propose de considérer les pratiques financières du capital investissement comme des « mathématiques parétiennes naturelles », hypothèse qui ouvre des perspectives sur l'élargissement de la rationalité aux contextes parétiens.

Mots clés : loi de puissance, valeurs extrêmes, hasard parétien, Mandelbrot, lois alpha-stables, processus de Lévy, capital investissement, rationalité

Abstract

Pareto's distribution became part of financial modelling at the beginning of the 1960s due to the identification of two unidentified properties of this distribution: it is a scale invariant distribution and a limit distribution in probability. These two properties of Pareto's distribution were first introduced by Benoît Mandelbrot, who based Pareto's distribution on Paul Lévy's alpha-stable distributions in the modelling of cotton prices and then of financial assets. This incorporation of the ancient Pareto law into the modelling of the dynamics of the stock market changes its status and confers on it a new life in financial economics, by feeding power laws into the models of the stock market. After having recalled the mathematical properties of the Paretian framework, in particular the paradox of the expected value in a power law, we introduce the modelling debate between a mixture of models (Gaussian for average values and Paretian for extreme values) and a single model for all values (alpha-stable), by the introduction of the extreme value theory and the generalized Pareto distribution. Then we describe the arrival of Pareto's distribution in finance by using Mandelbrot's methodology, which enables us to present the concept of Paretian chance. Finally, we suggest considering the financial practices of private equity as "natural Paretian mathematics", a hypothesis that opens perspectives on the enlargement of rationality to Paretian environments.

Keywords: power law, extreme values, Paretian chance, Mandelbrot, alpha-stable distributions, Lévy processes, private equity, rationality

Introduction

La loi de Pareto semble aujourd'hui posséder un caractère universel. Ses applications s'étendent de la linguistique aux réseaux sociaux, de la géographie à l'économie. De la distribution de la taille des villes à celle des entreprises, de la fréquence des mots dans le discours à la consultation

¹ FMSH, ISJPS (univ. Paris 1 Panthéon-Sorbonne, UMR 8103)

des pages Wikipédia, de la répartition des patrimoines à la répartition de la population etc. la loi de Pareto semble omniprésente dans les phénomènes les plus divers et à toutes les échelles (West, 2017). En physique, cette loi a été associée à la notion de criticalité auto-organisée, une théorie qui permet de modéliser des changements brusques résultant de petites perturbations aléatoires (Bak, 1997), comme dans les modèles de tas de sable auxquels on ajoute régulièrement des grains jusqu'à ce que le tas atteigne un état appelé « critique » tel que l'ajout d'un seul grain provoque une avalanche. Les variables qui décrivent le comportement du tas de sable suivent une loi de Pareto.

L'une des caractéristiques mathématiques de la loi de Pareto vient de ce qu'elle possède une queue de distribution en « loi de puissance » (*power law*). Une loi de puissance est une fonction f_a définie par $f_a : x \rightarrow x^{-a}$ où a est l'exposant de la fonction puissance. Les queues de distribution d'une fonction puissance sont ultra-étirées et décroissent très lentement, contrairement aux queues de distributions de la loi normale qui sont courtes et qui décroissent rapidement en exponentielle de forme $\exp(-x^2/2)$. Ce phénomène a été recensé de très nombreuses fois. En 1972, Paul Samuelson écrivait que « de telles distributions ultra-étirées se rencontrent fréquemment en économie » (Samuelson, 1972). En 2000, Daniel Zajdenweber compila un almanach de queues de distribution ultra-étirées (taille des entreprises, chiffres d'affaires annuels, ventes de disques d'or, box-office des films, sinistres en assurance, gisements pétrolifères etc.) l'amenant à considérer que, dans le monde, il y avait « des extrêmes partout » (Zajdenweber, 2000 : 7). En 2006, le nombre de phénomènes à queues de distribution ultra-étirées est tel que Chris Anderson estime ces phénomènes comme la « bande-annonce » de l'économie du XXI^e siècle (Anderson, 2006 : 25).

La modélisation financière est un domaine dans lequel la loi de Pareto fut à l'origine d'un renouvellement en profondeur des modélisations néoclassiques à base de lois normales, impropres à ajuster les queues de distributions empiriques très étirées des variations boursières. Ce fut Benoît Mandelbrot qui introduisit en 1962 cette loi à partir de l'usage des processus alpha-stables. Dans les années 1990, l'abondance de données disponibles combinée au regard nouveau des physiciens sur les variations boursières aboutit à l'émergence du courant de modélisation appelé « éconophysique » (Mantegna & Stanley, 2000 ; Gabaix, 2009 ; Bouchaud, 2021) dont l'une des caractéristiques fut de généraliser l'usage des lois de puissance à la modélisation des phénomènes économiques et financiers. Le développement fulgurant de ce courant au départ hétérodoxe car venant de physiciens et non de financiers, joint à la redécouverte des processus de Lévy non stables dans les années 1990, conduisit à l'omniprésence des lois de puissance dans la modélisation financière au tournant des années 2000. Parallèlement, l'usage des processus de Lévy non stables dans les modélisations de queues de distribution non gaussiennes mais non paréliennes, fit redécouvrir les spécificités des lois de Pareto par comparaison avec des comportements de queue qui n'étaient pas modélisables par des lois de puissance, et l'on chercha alors à expliquer l'émergence de lois de puissance par des mécanismes économiques d'interactions entre agents.

La littérature sur ce thème étant extrêmement importante, il ne saurait être question ici d'en faire ne serait-ce qu'un résumé. Notre intention est de nous concentrer sur l'arrivée de la loi de Pareto dans la modélisation financière au tournant des années 1960, et d'indiquer une piste de recherche qui propose de montrer que la loi de Pareto pourrait éclairer certains aspects de la dynamique des attentes des investisseurs et des prix des sociétés dans le secteur de l'investissement en capital sur la partie non cotée de l'économie. Les professionnels du non coté agissent souvent d'une manière qui semble contredire la rationalité à l'œuvre dans la théorie financière néoclassique. Sous l'angle du risque financier, la façon d'agir des investisseurs en capital semble illogique selon les critères de la théorie financière néoclassique.

Notre intention est de suggérer que l'usage de la loi de Pareto pourrait permettre de rendre logique ce qui paraît aujourd'hui illogique.

La section 2 présente un exposé introductif au cadre de pensée parétien en faisant apparaître le déplacement conceptuel qu'il opère. On rappelle comment une fonction puissance peut rendre caduque dans certains cas les calculs statistiques à base de moyenne et de variance, ce qui pose la question la moyennisation artificielle des données en présence de lois de puissance. La section 3 s'intéresse à la « pointe de la courbe » et introduit les lois du maximum. La section 4 présente la manière dont Mandelbrot fit entrer la loi de Pareto en finance au moyen de la transformée de Fourier, et introduit la notion de hasard parétien. La section 5 présente la proposition d'utiliser la méthodologie de l'ethnomathématique pour rendre compte de l'activité du capital investissement qui, pour le moment, échappe à la modélisation classique en finance. La conclusion esquisse ce que pourrait être une finance parétienne.

Le cadre de pensée parétien

La courbe des revenus et les lois de puissance

La courbe parétienne des revenus est aujourd'hui un objet extrêmement connu de la sociologie économique et de l'économie politique dont elle est devenue une boîte à outils classique (Bouvier, 1999 ; Steiner, 2013, voir aussi le numéro « Autour de Pareto » de la *Revue européenne des sciences sociales*, 2013). Dans les années 1895, Vilfredo Pareto se pose la question de la distribution des revenus (ou des patrimoines). Pour Pareto, si la distribution des revenus n'est pas gaussienne, cela signifierait que les revenus ne sont pas distribués « au hasard » mais résultent d'une loi mécanique à trouver. On aurait dans ce cas la preuve « qu'il y a une *cause* qui produit [leur] tendance à se disposer suivant une certaine courbe » (Pareto, 1895, cité par Barbut, 1999 : 88).

La courbe des revenus prend généralement l'allure d'une courbe à forte asymétrie : au lieu d'une distribution gaussienne des revenus (les « riches » et les « pauvres » se répartissant symétriquement autour d'un revenu « moyen »), on se trouve en présence d'une queue de distribution très étirée sur la droite qui signifie que les « grandes valeurs » (les revenus élevés) ne sont pas rares, d'une forte proportion de revenus moyens légèrement décentrés, et d'une croissance rapide des revenus faibles sur la gauche (absence de queue de distribution à gauche).

Une première idée « novatrice »

Cherchant à modéliser le mieux possible la partie droite de la courbe de répartition (les revenus élevés), Pareto a une première idée, « novatrice à son époque » (Barbut, 1999 : 89), considérer, non les tranches de revenus selon leur ordre de grandeur, ce qui reviendrait à tracer l'histogramme des fréquences relatives des revenus (une approximation de leur fonction de densité), mais la distribution des fréquences *cumulées* ou fonction de répartition. Écrivons cela en langage mathématique d'aujourd'hui. Si X est une variable aléatoire représentant la valeur du revenu (ou du patrimoine), en choisissant de considérer la fonction de répartition (cumulée) et non la fonction de distribution (relative) Pareto s'intéresse à $F_X(x) = \Pr(X < x)$ où « Pr » signifie « probabilité » (noté en caractères droits).

Comme il s'agit de caractériser et de modéliser les grands revenus (ou les grands patrimoines), il faut aussi définir un « revenu-seuil » x_0 au-delà duquel s'effectueront les observations, et la démarche de Pareto revient à s'intéresser aux valeurs de X supérieures à x_0 . Si x_0 est trop faible, il y aura trop de revenus moyens parmi la sélection des « grands ». Si x_0 est trop grand, il n'y aura pas assez de revenus élevés. Le choix du seuil x_0 exercera donc une influence importante pour la description des queues de distribution et leur ajustement par une fonction mathématique. Savoir quelle fonction décrit au mieux une queue de distribution nécessite de choisir un seuil

x_0 qui définit la partie « queue » de la distribution examinée. Une « queue » ne peut se définir sans avoir au préalable défini le seuil x_0 à partir duquel elle commence. La démarche de Pareto admet deux modélisations probabilistes possibles. Soit considérer les valeurs les plus grandes de X au-delà du seuil x_0 et, dans ce cas, cela revient à introduire la variable aléatoire $\{X \mid X > x_0\}$ (où le symbole « \mid » signifie « tel que ») et chercher à modéliser ses extrêmes, soit considérer les différences de revenus par rapport au revenu-seuil, les excédents $X - x_0$ et, dans ce cas, cela revient à examiner la variable aléatoire $\{X - x_0 \mid X > x_0\}$ et chercher à modéliser les dépassements de seuil. Nous verrons plus loin que ces deux regards conduisent à deux lois limites différentes.

Pour obtenir sa « courbe des revenus », Pareto classe les revenus et la fréquence cumulée par ordre croissant, créant ainsi une courbe appelée « rang-fréquence » qui se construit en associant un rang de classement du revenu et la valeur de sa fréquence cumulée. Ce graphique aussi appelé « rang-taille » est très connu aujourd'hui et a été utilisé par George Kingsley Zipf en 1949 pour la fréquence des mots dans le discours. Pour le construire, on commence par classer les différentes valeurs prises par X selon leur amplitude, de la plus grande à la plus petite. La plus grande a le rang 1 etc. S'il y a n revenus supérieurs à x_0 , l'échantillon ordonné des revenus s'écrit² $X_{(1)} \geq X_{(2)} \geq \dots \geq X_{(n)}$. Pour chaque revenu X_i de valeur x , on considère le nombre de revenus de valeur supérieure ou égale à x , que nous notons $\text{Nb}(X_i \geq x)$. Si $x = X_{(1)}$, il est clair³ que $\text{Nb}(X_i \geq X_{(1)}) = 1$ et que $\text{Nb}(X_i \geq X_{(k)}) = k$. Il y aura donc des points de coordonnées $(X_{(k)}, k)$ dans un plan « amplitude-rang », un graphique dans lequel la valeur du revenu est en abscisse de $X_{(1)}$ à $X_{(n)}$, et le rang de classement du revenu par ordre croissant est en ordonnée de 1 à n . C'est l'inverse de la fonction de répartition des revenus X , en fait son écart à 1. Il s'agit de la fonction de répartition « à l'envers » notée⁴ $\bar{F}(x) = 1 - F(x)$

On a considéré le nombre Nb de revenus mais on peut tout aussi bien considérer la fréquence empirique. Le passage de l'un à l'autre est intuitif. Si l'on considère que l'échantillon des revenus est de taille n , alors la fréquence empirique que l'on note $\text{Fr}(X_i \geq x)$ est simplement égale à $\text{Fr}(X_i \geq x) = \text{Nb}(X_i \geq x) / n$. Ceci nous permet d'écrire que :

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

Que l'on choisisse le nombre Nb ou la fréquence empirique Fr , l'interprétation est identique.

Une deuxième idée « d'ingénieur »

La deuxième idée de Pareto fut de représenter la fonction de répartition sur un papier bilogarithmique, soit un graphe en coordonnées logarithmiques des couples $(X_{(k)}, k)$, soit l'ensemble des points $(\ln X_{(k)}, \ln k)$, « une idée d'ingénieur » (Barbut, 1999 : 89). Cette représentation bilogarithmique fait apparaître, au-delà d'un seuil x_0 , une droite de pente négative de coefficient alpha :

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

Dans la relation précédente, alpha est appelé « index de Pareto » et « C_+ » est une constante positive⁵ qui représente un paramètre d'échelle pour la distribution observée. L'échelle représente intuitivement la « taille » de la queue de distribution (petite ou grande)

² La notation avec une parenthèse sur l'indice k indique le classement des revenus par taille décroissante : $X_{(1)}$ est le revenu le plus grand, $X_{(2)}$ est le deuxième plus grand revenu etc.

³ Barbut (1999) note $N(x_i)$ le nombre des titulaires de revenus supérieurs ou égaux à x_i . Nous choisissons ici une notation qui distingue la variable aléatoire X_i de sa réalisation x_i et de la variable aléatoire classée $X_{(i)}$, notation qui nous permettra d'effectuer ensuite le lien avec les lois alpha-stables et la théorie des valeurs extrêmes.

⁴ L'expression \bar{F} se prononce « F barre ».

⁵ Notée A chez Barbut (1999) et Pareto (1895).

indépendamment de sa « forme » (gaussienne ou non), un peu comme les poupées russes gigognes (les « Matriochka ») ont une taille et une forme. Cette échelle n'est pas indépendante du seuil x_0 choisi. En posant $x_0 = (C+\alpha)^{1/\alpha}$, la relation précédente s'écrit :

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

La fonction de queue (notée \bar{F}) est une *fonction puissance*.

Il est usuel de considérer que la queue de distribution précédente en fonction puissance suit une loi dite « loi de Pareto au sens strict » ou « distribution de Pareto simple » ou « distribution de Pareto de type I ». Mais comme le rappelle Barbut (1999 : 90), « il fallait faire [une] généralisation car une grandeur mesurable n'est définie qu'à un changement d'unité près et à un changement d'origine près ». Pareto généralisa cette première expression en :

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

Cette seconde fonction de queue est une distribution de Pareto dite « de type II ».

L'adéquation de la distribution de Pareto aux grandes valeurs des revenus (ou des patrimoines), pose cependant deux problèmes d'ordres pratique et théorique, obstacles qui ont représenté un frein à l'utilisation de cette loi. Le premier problème provient de ce que la distribution de Pareto ne permet pas de représenter toute la répartition des revenus mais seulement celle des revenus élevés, sans les revenus moyens ou faibles. Pour le dire autrement, si l'on utilise une distribution de Pareto pour ajuster les courbes empiriques de revenus, l'ajustement n'est réalisable que sur les queues, ce qui pose la question de la définition de la queue (donc du choix du seuil x_0 comme vu précédemment). La même fonction ne peut donc pas modéliser les queues de distribution et les parties centrales. Si l'on utilise des distributions de Pareto pour la description des revenus élevés, il sera nécessaire de choisir une autre fonction pour les autres revenus. Il y aura alors une « loi » pour les grandes valeurs des revenus, une autre pour les autres, à l'instar de ce qui se pratique par exemple en assurance dommages pour les sinistres, où l'on distingue la loi des « grands sinistres » de la loi des autres sinistres. Où commencent les « grandes valeurs » devient alors la question à résoudre.

En d'autres termes, choisir d'utiliser une loi de Pareto pour caractériser les grandes valeurs d'un phénomène quelconque donné, c'est se contraindre à scinder le phénomène en deux sous-phénomènes, relevant chacun d'un régime particulier : les grandes valeurs d'une part, les petites de l'autre. Il faudra alors combiner les deux régimes, donc les deux lois.

La perte de pertinence des valeurs typiques à base de moyennes

Le deuxième problème relève de la capacité à utiliser ce qu'on appelle les *valeurs typiques* des populations statistiques, des chiffres qui servent à « résumer » la distribution de ces populations à partir de moyennes ou de transformations de moyennes⁶. Pour toute série statistique, les valeurs typiques d'une distribution sont ses paramètres de valeur centrale (position), de dispersion et de forme (asymétrie et courbure). Ces paramètres sont la moyenne (position), l'écart type (dispersion), l'obliquité (asymétrie) et l'aplatissement (courbure). Ces paramètres sont calculés à l'aide de fonctions appelées les « moments » de la variable aléatoire. Un moment est une moyenne de puissances de cette variable⁷. Connaître les moments permet de comprendre

⁶ Pour une introduction très claire en langue française sur cette question, on pourra consulter Barbut (1991a, 1991b).

⁷ Le moment d'ordre k d'une variable aléatoire X est défini comme l'espérance mathématique de la puissance k -ième de cette variable soit $E[X^k]$. L'espérance mathématique est le moment d'ordre 1 noté μ , ou « moyenne ». Le

l'allure de la distribution. Par exemple, dans une distribution symétrique comme la gaussienne, les moments d'ordre impair (troisième, cinquième puissance etc. de la variable) sont nuls. La moyenne est le moment d'ordre 1, ce que Laplace appelait « le milieu qu'il faut prendre entre les résultats de plusieurs observations » (Barbut, 1991b : 135).

Statistiquement, on peut toujours calculer une moyenne empirique. Mais la moyenne empirique peut ne pas être significative. Imaginons que l'on calcule la moyenne empirique du volume d'eau (en mètres cubes) des étendues d'eau sur Terre. La moyenne empirique sera à peu près constante pour les lacs mais au moment où l'on ajoutera un océan, elle augmentera de plusieurs ordres de grandeurs. La moyenne empirique ne converge pas mais diverge en raison de la présence de grands océans, qui rend les données hétérogènes en ordre de grandeur. On ne sait jamais quelle est la « bonne » moyenne puisque chaque nouvelle grande valeur peut venir déplacer la moyenne empirique. Les résultats des calculs de moyenne empirique fluctueront sans cesse. C'est un comportement statistique très déroutant. Les moyennes successives restent longtemps voisines d'une même valeur et tout à coup font un bond dans un sens ou dans l'autre. De même les variances empiriques successives pourront rester longtemps stables avant de sauter de plusieurs ordres de grandeur dans un sens ou dans l'autre. Et ainsi pour tous les moments empiriques.

Mathématiquement, cette question est celle du *domaine d'attraction*⁸ d'une distribution qui assure aux moyennes empiriques (aux moments empiriques) de converger vers la moyenne théorique (vers le moment théorique) si celle-ci (celui-ci) existe quand la taille de l'échantillon augmente indéfiniment. Elle renvoie au « théorème central limite » et aux lois des grands nombres. Le problème de la convergence des moyennes a été résolu par Paul Lévy en 1924 (Lévy, 1924). Ce qu'a montré Paul Lévy, c'est que, si l'espérance et la variance sont finies, la seule loi limite possible est la loi normale. Dans tous les autres cas, on obtient des lois limite généralisées qu'il nomme lois « alpha-stables ».

Si la moyenne empirique ne converge pas vers une limite finie, cela signifie que l'espérance mathématique n'existe pas, soit que l'espérance mathématique n'est pas définie (oscillation indéfinie), soit que l'espérance mathématique est infinie. L'inexistence de l'espérance mathématique se traduira par l'impossibilité de parvenir à trouver une valeur fiable pour la moyenne empirique. Et plus généralement pour les moments d'ordre supérieur comme la variance. On ne peut savoir si l'espérance mathématique et les moments existent qu'à partir de la forme mathématique de la loi de distribution qui sert à calculer l'espérance mathématique⁹. La divergence de la moyenne empirique implique une difficulté de l'usage de la loi des grands nombres dans sa forme classique, celle du domaine d'attraction de la distribution normale.

Dans le cas de la gaussienne, tous les moments existent. Dans le cas de la loi de Pareto, le calcul montre que les moments d'ordre k de la variable aléatoire X n'existent pas pour des valeurs de k inférieures à α car dans ce cas l'intégrale qui les calcule diverge. Par exemple, quand α est inférieur à 2, l'intégrale qui calcule le moment d'ordre 2 diverge et la variance théorique devient infinie, même si on peut toujours calculer des variances empiriques. Quand α est inférieur à 1, l'intégrale qui calcule l'espérance mathématique diverge et la moyenne théorique devient infinie, même si on peut toujours calculer des moyennes empiriques. Cela veut dire que, avec une loi de Pareto, les valeurs typiques (moyenne, variance) sont infinies

moment centré d'ordre k sur la moyenne est noté $m_k = E[(X - \mu)^k]$. Le moment centré d'ordre 1 est nul. Le moment centré d'ordre 2 est la variance, dont la racine carrée est l'écart type.

⁸ Expression introduite par Paul Lévy (Barbut, 1991a : 36).

⁹ Ce calcul est une intégrale sur la fonction de densité. L'intégrale peut exister ou peut diverger. La limite des moyennes empiriques converge ou non vers l'espérance mathématique.

dans certains cas. Cela pose le problème de l'interprétation d'une moyenne empirique lorsque la moyenne théorique diverge. Et donc des erreurs d'estimation dans le cas de lois de Pareto.

On montre que l'erreur d'estimation croît en fonction de la décroissance d' α . Par exemple si $\alpha = 1,5$, l'erreur d'estimation d'une moyenne de 100 observations est de 22%. Si $\alpha = 1,1$, cette erreur d'estimation augmente à 66%. L'usage d'une moyenne artificielle devient alors totalement injustifié. La perte de pertinence des valeurs typiques empiriques pose un problème d'utilisation pratique des données dans le cas de la présence de lois de puissance.

La concentration et la « loi des 80/20 »

Le grand nombre de valeurs élevées conduit au phénomène de la « concentration » des revenus ou des patrimoines, illustré par l'adage selon lequel « beaucoup ont très peu, et très peu ont beaucoup » exprimé par ladite « loi des 80/20 » selon laquelle 80% des éléments d'un phénomène contribuent à 20% du total observé et réciproquement, ou « principe des 80/20 » (Koch, 1997). La concentration peut se mettre en évidence par des indicateurs globaux de l'inégalité, soit par un calcul direct comme les mesures de concentration ou l'indice de Gini¹⁰, soit par un graphique, la courbe de concentration de Lorenz (Barbut 2007, chapitres 1 et 3).

La courbe de concentration de Lorenz représente le pourcentage du revenu total d'une population donnée en fonction du pourcentage de la population totale, classée du plus pauvre au plus riche. Chaque point de la courbe a pour abscisse une valeur de la fréquence cumulée de la population, et pour ordonnée une valeur de la fréquence cumulée du revenu total. La première bissectrice représente l'égalité absolue entre les individus : les $n\%$ inférieurs de la population reçoivent exactement $n\%$ du revenu global. C'est la ligne de concentration nulle, ou équirépartition parfaite. À l'opposé, tout écart de la courbe de Lorenz à cette bissectrice traduit une inégalité de répartition de revenus. La surface comprise entre la courbe de Lorenz et la bissectrice est l'aire de concentration : elle augmente avec le degré d'inégalité entre les individus. Plus la courbe de Lorenz s'écarte de la bissectrice, et plus cette disparité est grande, plus la concentration est forte, plus la disparité entre les revenus apparaît. À l'extrême, quelques individus possèdent l'essentiel de la richesse totale. Un certain nombre de problèmes ou de paradoxes peuvent résulter de l'usage de telles fonctions pour la mesure de l'inégalité, dont on trouvera un exposé synthétique dans la recension de l'ouvrage de Marc Barbut par Gianluca Manzo (2009).

La pointe de la courbe des revenus et la loi de Pareto généralisée

Considérons la hiérarchisation des grandes valeurs des distributions, par exemple les écarts entre les revenus ou les patrimoines les plus élevés, la « pointe de la courbe des revenus » (Steiner, 2013 : 201). C'est une façon d'aborder la question de Sorel (1897) sur la progression du patrimoine des « riches ».

La valeur typique du maximum

Une particularité importante des distributions à queue étirée en fonction puissance est la très forte hiérarchisation de ces valeurs. Nous allons préciser cette idée en introduisant la notion de rapport de hiérarchisation. Conservons les notations introduites précédemment. La plus grande valeur a le rang 1 etc. S'il y a n revenus supérieurs à x_0 , l'échantillon ordonné des revenus s'écrit $X_{(1)} \geq X_{(2)} \geq \dots \geq X_{(n)}$. le rapport de hiérarchisation de rang k est défini par la quantité $X_{(k)}$

¹⁰L'économiste italien Corrado Gini a développé cette analyse au cours de ses travaux sur les disparités de revenus. Il a abouti à la construction d'une courbe dite « courbe de concentration » et la détermination d'un ratio, l'indice de Gini.

$/ X_{(k+1)}$. Lorsque les X_k présentent une queue de distribution en fonction puissance du type Pareto d'exposant alpha, on montre que :

$$X_{(k)} / X_{(k+1)} = (k+1/k)^{1/\alpha}$$

Si l'on examine par exemple le rapport de hiérarchisation entre le patrimoine le plus grand et celui qui vient juste après, il vient d'après la relation précédente $X_{(1)} / X_{(2)} = 2^{1/\alpha}$. On voit que le rapport entre les deux patrimoines les plus élevés est de l'ordre de grandeur de $2^{1/\alpha}$. Si $\alpha = 1,5$, on obtient $2^{2/3} \approx 1,6$: le plus grand patrimoine observé sera 1,6 fois plus grand que le second plus grand. Si $\alpha = 0,5$, on obtient $2^2 = 4$: le plus grand patrimoine serait 4 fois plus grand que le second plus grand. Plus alpha diminue et plus ce rapport, donc la hiérarchisation des valeurs, augmente. Les grandes valeurs sont d'autant plus grandes que alpha diminue ou, pour le dire autrement, que l'inégalité de la distribution augmente quand le paramètre alpha décroît. Il est intéressant ici de se rappeler que Pareto lui-même avait commis une erreur dans l'interprétation de la variation du paramètre alpha notamment dans sa querelle avec Sorel¹¹. Pour Pareto, l'inégalité diminue avec alpha (Barbut, 1999 : 93 et 2007 : 197) soit donc le contraire de ce qui advient avec une fonction puissance.

Le paramètre alpha détermine directement la forme de la « pointe de la courbe des revenus » (ici des patrimoines) et donc la concentration de la richesse. On peut déduire de valeurs faibles de alpha une très forte hiérarchisation des valeurs des revenus ou des patrimoines ou, pour le dire autrement, que la « pente » se raidit quand on arrive sur les 1%, 0,1%, 0,01% des patrimoines les plus élevés. Les écarts entre chaque patrimoine sont de plus en plus grands. Cela signifie que la distribution des patrimoines au-delà du seuil x_0 est fortement inégalitaire ou, plus précisément, fortement hiérarchisée.

Demandons-nous à présent quelle serait la « valeur typique » du patrimoine maximum au sens du patrimoine le plus probable. Le rapport entre le plus grand patrimoine et le dernier (qui est le seuil x_0), est $X_{(1)} / x_0 = n^{1/\alpha}$. Il vient que l'ordre de grandeur de $X_{(1)}$ est :

$$X_{(1)} = x_0 n^{1/\alpha}$$

Supposons par exemple que la taille de l'échantillon soit $n = 10\,000$. En choisissant comme précédemment $\alpha = 1,5$, avec une valeur normalisée arbitrairement à $x_0 = 1$, la plus grande valeur observée sera de l'ordre de $10\,000^{2/3} = 465$. Si à présent alpha est inférieur à 1, par exemple $\alpha = 0,5$, alors la plus grande valeur est de l'ordre de $10\,000^2$ soit 100 000 000.

La moyenne probable au-delà d'un seuil et les lois du maximum

Intéressons-nous, comme le proposait Sorel, aux individus qui ont un patrimoine X supérieur à un « seuil de richesse » x_0 donné (par exemple $x_0 = 500\,000$ euros), soit la variable $X > x_0$. Demandons-nous ensuite, parmi ceux dont le patrimoine dépasse ce seuil, quelle est la valeur *moyenne* de leur patrimoine. Il s'agit d'une « moyenne conditionnelle » de X au-delà du seuil x_0 qui est l'espérance mathématique que l'on notera $M(x_0)$. Maurice Fréchet a montré (Barbut, 1989, 1998) que, dans le cas de lois de Pareto, la moyenne conditionnelle au-delà d'un seuil x_0 est un multiple de ce seuil à une constante près, c'est-à-dire une fonction linéaire de ce seuil :

$$M(x_0) = a x_0 + b$$

Si $b = 0$, la distribution est de Pareto de type I, si $b \neq 0$, la distribution est de Pareto de type II. Par exemple si $a=2$ et $b=0$, le patrimoine moyen de ceux dont le patrimoine dépasse 500 000 euros est 1 000 000 d'euros. La moyenne conditionnelle au-delà du seuil x_0 s'écrit :

¹¹ Je remercie l'un des rapporteurs de m'avoir fait cette remarque et cette suggestion de commentaire.

$$M(x_0) = \frac{\alpha}{\alpha-1} \times x_0$$

L'existence d'une moyenne conditionnelle linéaire est une « façon intuitive de percevoir l'inégalité d'une distribution » (Barbut, 1998 : 43). Par exemple, si $\alpha = 1,5$, alors $a = 3$ et $M(x_0) = 3 x_0$. Cela veut dire que la valeur la plus probable, une fois qu'on dépasse le seuil, est le triple du seuil. Le cadre de pensée parétien est un cadre dans lequel toute valeur probable conditionnée par un franchissement de seuil est un multiple de ce seuil. C'est une caractéristique fondamentale des fonctions puissance : la valeur probable d'une quantité – une fois franchi un seuil – est proportionnelle au seuil franchi. L'on appelle cette caractéristique le paradoxe de la valeur probable en univers parétien. En utilisant une analogie météorologique, on pourrait dire que, contrairement à l'adage populaire qui énonce « après la pluie, le beau temps », on se trouve ici davantage dans une structure météorologique où les aléas climatiques seraient tels que, après la pluie, le plus vraisemblable sera l'orage.

Réciproquement, si une relation linéaire de ce type est vérifiée par une variable aléatoire X pour $X > x_0$ (dans l'exemple précédent, X est le montant du revenu ou patrimoine des « riches »), alors la distribution de probabilité de cette variable appartient admet une fonction de queue qui peut prendre plusieurs formes, qualifiées par Barbut (1988) de « parétienne » (fonction puissance), « contra-parétiennes » ou exponentielle. Quand $a > 1$, on se trouve dans le cas de distributions « parétiennes » et $a = \alpha / (\alpha - 1)$. Dans le deuxième cas, $\alpha = 1$ et $b > 0$. Dans ce cas, la distribution de Pareto n'en est plus une car elle se transforme en distribution à queue exponentielle. La moyenne conditionnelle au-delà du seuil x_0 s'écrit :

$$M(x_0) = x_0 + b$$

La valeur moyenne au-delà du seuil est égale au seuil lui-même augmenté de l'écart type de la distribution. Quand x_0 devient « grand », $M(x_0)$ est peu différent de x_0 . Ce qui justifie l'usage des « valeurs typiques » à base de moyennes dans le cas de distributions exponentielles.

Dans le troisième cas, $0 < a < 1$ et la distribution est appelée « contra-parétienne » ou « anti-parétienne » parce que (Barbut, 1998 : 48)

dans la famille étudiée, tout est à l'opposé et en dualité de ce qui se passe pour celles de Pareto :

- La variable est bornée supérieurement au lieu de l'être inférieurement
- Ses moments existent à tous les ordres
- L'exposant est positif au lieu d'être négatif
- Les grandes valeurs sont rigoureusement impossibles

On peut alors faire le lien avec la théorie des valeurs extrêmes et les lois du maximum¹². De la même manière que la convergence des moyennes empiriques vers l'espérance mathématique (la moyenne « théorique ») est donnée par la loi des grands nombres dans sa forme classique gaussienne, il existe une convergence des extrêmes empiriques vers des limites particulières. La répartition des moyennes empiriques autour de leur espérance tend à être normale quand la variance est finie, la répartition des extrêmes empiriques tend aussi vers une loi limite appelée « loi des extrêmes » dont la forme est donnée par le théorème de Fisher-Tippett (1928). A la fluctuation des moyennes empiriques correspond la fluctuation des maxima empiriques.

Cette loi limite prend nécessairement l'une des trois formes suivantes : loi de Fréchet, loi de Weibull, loi de Gumbel. L'exponentielle de la loi de Pareto correspond à une loi de Fréchet. L'exponentielle d'une distribution anti-parétienne correspond à la loi de Weibull.

¹² La théorie des valeurs extrêmes fut imaginée aux Pays-Bas dans les années 1930 pour modéliser les risques de crue et d'inondations dans les polders. Pour une introduction à la théorie des valeurs extrêmes, on pourra consulter Embrechts, Klüppelberg & Mikosch (1997).

L'exponentielle d'une exponentielle correspond à la loi de Gumbel. Ces trois formes ont été unifiées par une loi généralisée du maximum dans la représentation de Jenkinson-von Mises (1954). Un paramètre de cette loi (noté ξ) résume les trois possibilités et représente le poids des queues de distribution. Dans le premier cas, celui de la loi de Fréchet, ξ est positif et correspond à $1/\alpha$. Plus alpha diminue, plus ξ augmente, plus la queue de distribution s'épaissit et plus l'inégalité s'accroît (on peut penser que Pareto avait à l'esprit ξ et non alpha lorsqu'il commit son erreur d'interprétation). Dans le deuxième cas, celui de la loi de Weibull, ξ est négatif et correspond à $-1/\alpha$. La queue de distribution est fine, comme dans la loi bêta. Le troisième cas dans lequel ξ est nul représente un cas limite des deux précédents quand ξ tend vers zéro. C'est la loi de Gumbel, dont la queue est à décroissance exponentielle comme dans la loi normale.

Nous disions au début que la démarche de Pareto admettait deux regards mathématiques possibles, soit considérer les valeurs les plus grandes de X et, dans ce cas, cela revenait à introduire la variable aléatoire $\{X | X > x_0\}$, soit considérer les excédents $X - x_0$ et, dans ce cas, cela revenait à examiner la variable aléatoire $\{X - x_0 | X > x_0\}$. Ces deux approches correspondent aux lois des maxima et aux lois des dépassements de seuil. Par exemple, si les maxima sont modélisables par une loi de Fréchet, les dépassements de seuil suivent une loi de Pareto de type I ou de type II. Si les grandes valeurs des revenus ou des patrimoines au-delà de seuils suffisamment élevés peuvent être modélisées par une loi des extrêmes, alors les théorèmes de Gnedenko (1943), Balkema et De Haan (1974) et Pickands (1975) montrent que la loi des dépassements de seuils peut être modélisée par une loi de Pareto généralisée.

Le hasard parétien

En 1959, Mandelbrot se saisit de la question de Pareto sur les revenus et publie deux *Comptes rendus à l'Académie des sciences* (séances du 3 août et du 23 novembre 1959) sur la modélisation de la distribution des revenus (Mandelbrot, 1959a, 1959b). Ces deux communications vont renouveler en profondeur la manière de comprendre en économie les distributions à queue étirée et introduire la notion de hasard parétien.

La loi de Pareto et la représentation de Lévy-Khintchine

Le point de départ de la réflexion de Mandelbrot est de chercher à lever les obstacles résultant des deux limitations de la loi de Pareto, que l'on rappelle. D'une part, choisir cette loi revient à devoir opérer une scission de la répartition des revenus en deux régimes distincts, les revenus élevés et les autres. D'autre part, contrairement à la loi normale dont l'un des intérêts principaux provient de son statut probabiliste de loi limite (existence du théorème de la limite centrée sous sa forme classique), la loi de Pareto n'apparaissait pas, aux yeux des économistes de l'époque, avoir de justification théorique semblable dans la théorie des probabilités. Alors que la gaussienne s'obtient par application du théorème de la limite centrée, la loi de Pareto semblait ne pas posséder ce statut probabiliste de loi limite.

Le 3 août 1959, Mandelbrot propose de représenter la répartition *entière* des revenus par une loi de probabilité alpha-stable de Lévy, totalement asymétrique et de valeur moyenne finie, qui représente une limite en probabilité pour une loi des grands nombres généralisée. Une modélisation qui lève d'un coup les deux obstacles de la loi de Pareto. Cette proposition est à ce point révolutionnaire qu'il est nécessaire d'être précis sur les raisons qui lèvent ces obstacles. Contrairement à la loi normale et aux principales distributions de probabilité connues par leur fonction de densité, une distribution alpha-stable ne possède pas de fonction de densité connue. Pour la décrire, on passe dans l'espace de Fourier par une transformation qui aboutit à ce qu'on

appelle la fonction caractéristique, qui est la transformée de Fourier de la densité¹³. Quelle que soit une loi de probabilité donnée, la fonction caractéristique existe toujours tandis que la densité n'existe pas toujours. Mandelbrot décrit la distribution des revenus par la transformée de Fourier de leur densité en précisant la forme de sa fonction caractéristique.

La fonction caractéristique utilisée par Mandelbrot est celle de l'exposant caractéristique¹⁴ d'une loi alpha-stable. La transformée de Fourier de la densité d'une loi alpha-stable est généralement représentée par son exposant caractéristique dont on connaît l'expression, c'est la représentation dite « de Lévy Khintchine ». On utilise en général¹⁵ une *forme intégrée* de l'exposant caractéristique. Cette forme intégrée ne donne pas accès à l'intuition directe de la présence d'une loi de Pareto. Pour avoir accès à la présence de Pareto dans ces lois, mieux vaut utiliser une *forme non intégrée* (Le Courtois et Walter, 2012 : 90) utilisant une fonction appelée la « mesure de Lévy ». La mesure de Lévy est une fonction qui « règle » le poids des queues de distribution, aussi bien les fluctuations des sauts que l'asymétrie et la courbure. Avec une loi alpha-stable, la mesure de Lévy notée $\nu(x)$ s'écrit comme une loi de Pareto de type I :

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

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

On notera la différence avec la formule originale de Pareto due à l'exposant $1+\alpha$ alors que la formule de Pareto indiquait alpha seul. Cette différence provient de ce que Pareto utilisait la fonction de répartition tandis que la mesure de Lévy utilise la fonction de distribution, soit la dérivée de la fonction de répartition. Mais il s'agit bien de la même loi. La mesure de Lévy est dissymétrique. Les deux valeurs différentes C_- et C_+ expriment la « taille » des queues de distribution à gauche et à droite. Le paramètre alpha donne la « forme » de la distribution. Quand alpha = 2, il s'agit d'une loi normale. Quand alpha = 1, il s'agit d'une loi de Cauchy. Le paramètre alpha est l'exposant de Pareto. Comme la mesure de Lévy est parétienne, et que la mesure de Lévy définit le comportement « à l'infini » du processus aléatoire, les distributions alpha-stables sont asymptotiquement parétiennes avec l'indice de queue alpha.

L'ajustement par une loi alpha-stable dissymétrique de la distribution des revenus étant vérifié sur toute la distribution, la queue de distribution étant parétienne d'indice alpha, Mandelbrot décide de désigner cette loi comme une « loi de Pareto-Lévy » (Mandelbrot, 1959a : 613) et il conclut que « la distribution de Pareto-Lévy paraît être une bonne première approximation dans tout le domaine de variation » du revenu (*id.*). Sur la partie gauche de la distribution des revenus, Mandelbrot précise que la distribution alpha-stable dissymétrique « décroît plus vite que toute exponentielle » (Mandelbrot, 1959b : 2153). L'aspect nouveau de la proposition de Mandelbrot vient donc de la réunification des deux parties de la courbe, les revenus très faibles comme les revenus très élevés. Il n'est plus nécessaire de chercher un seuil de séparation entre deux régimes de revenus, ni de combiner deux lois de probabilités pour ces deux régimes.

L'autre propriété des lois alpha-stables est leur stabilité par addition. Cela veut dire que « la même loi paraît s'appliquer au revenu tout entier ainsi qu'à ses diverses composantes prises séparément » (Mandelbrot, 1959a : 613). Quand on cherche à expliquer un revenu par ses

¹³ Si Y est une distribution de probabilité, la transformée de Fourier de sa densité est $\Phi_Y(u) = E[\exp(iuY)]$ ou « exp » est la fonction exponentielle et « i » l'unité imaginaire dont le carré vaut (-1) : tel que $i^2 = -1$.

¹⁴ Si $\Phi_Y(u)$ est la fonction caractéristique, l'exposant caractéristique est $\Psi_Y(u)$ tel que $\Phi_Y(u) = \exp(\Psi_Y(u))$ où « exp » représente la fonction exponentielle. Il est très important de ne pas confondre la fonction caractéristique et l'exposant caractéristique : $\Psi_Y(u) = \ln \Phi_Y(u)$ où « ln » est le logarithme népérien.

¹⁵ Dans la littérature économique ou sociologique, ou encore dans les travaux d'histoire de la pensée économique.

composantes, on « espère pouvoir appliquer quelque théorème limite des probabilités » (Mandelbrot, 1959a : 614). De manière générale, lorsqu'on agrège des quantités économiques, s'il existe une loi par échelle d'observation, alors on ne pourra pas donner du phénomène une explication globale. D'où l'importance de pouvoir représenter un phénomène économique par une loi qui soit *invariante par agrégation*, c'est-à-dire *invariante par changement d'échelle* d'observation. En 1967, Mandelbrot revint sur cette propriété : « étant donné les difficultés notoires de l'observation dans les sciences humaines, comment se fait-il que divers observateurs ne puissent jamais se mettre d'accord sur le même modèle stochastique pour représenter l'activité spontanée des systèmes auxquels ils s'intéressent ? ». Sa réponse fut : « sous certaines conditions assez générales, nous pensons que la réponse à cette question est que seules les lois d'une certaine famille sont susceptibles d'être jamais observées, parce qu'elles sont *les seules à être invariantes par rapport aux diverses transformations de la réalité* que l'on rencontre lorsque l'on passe d'un observateur à un autre. Ces lois sont liées à une célèbre expression statistique dite loi de Pareto'' (Mandelbrot, 1967 : 1098, nos italiques).

Une distribution alpha-stable est aussi une loi limite et, dans ce sens, les lois alpha-stables résolvent le problème des domaines d'attraction des lois comme celle de Pareto. Ainsi, la loi de Pareto acquiert le statut de loi limite et Pareto peut qualifier sa loi de nouvelle « loi des erreurs »¹⁶ (Barbut, 1991a : 42). Mais une loi des erreurs régie par d'autres mécanismes que la loi normale. De la même manière que la loi des erreurs de Gauss définit un « hasard gaussien » (tirages aléatoires gaussiens), il est possible de définir avec la loi de Pareto un « hasard parétien » (tirages aléatoires parétiens) que Mandelbrot introduira en 1962 (cf. ci-dessous).

Enfin, si l'on considère la suite des revenus successifs dans le temps, les différentes lois de distribution, stables par addition, constituent un processus aléatoire alpha-stable de Lévy (Lévy, 1937) ou encore une « marche au hasard »¹⁷. Partant de l'analyse de la distribution des revenus, Mandelbrot a introduit l'outil probabiliste qui va lui permettre d'aborder la question de la variation des prix boursiers. Avec le passage des lois alpha-stables des revenus aux prix boursiers, la loi de Pareto pénètre dans la modélisation financière.

Le passage de la loi de Pareto en finance et le hasard parétien

La première mise en évidence du caractère parétien des fluctuations des prix cotés sur les marchés boursiers apparaît dans la désormais classique étude des variations du prix du coton entre 1880 et 1958. Cette étude, constituant une partie d'un rapport interne de recherche à IBM¹⁸, fut publiée sous forme réduite dans les *Comptes rendus à l'Académie des Sciences*, séance du 4 juin 1962 (Mandelbrot, 1962). C'est la première trace d'une considération explicite de l'existence de phénomènes parétiens en finance, à partir de l'étude des grandes variations des prix. L'utilisation par Mandelbrot des processus alpha-stables de Lévy comme modèle de variations des cours, à une époque où l'ensemble des études statistiques entreprises sur les marchés cherchaient à établir que les variations des cours pouvaient être représentées par des mouvements browniens, permet de qualifier cette représentation nouvelle de « modèle de Mandelbrot ».

Dans sa communication de 1962, Mandelbrot s'intéresse aux variations des différences premières des logarithmes des prix de clôture du coton sur τ jours. En notant $S(t)$ le prix coté

¹⁶ La première loi des erreurs étant la loi normale de Laplace-Gauss.

¹⁷ Une marche au hasard est définie comme un processus aléatoire à accroissements indépendants et stationnaires. Il n'est pas nécessaire d'ajouter des conditions supplémentaires sur l'existence des moments. Le mouvement brownien n'est qu'une marche au hasard parmi d'autres. L'hypothèse de la marche au hasard en finance est souvent confondue avec l'hypothèse brownienne.

¹⁸ Research Note NC-87, 26 mars 1962, Centre de recherche IBM, Yorktown Heights, New York.

du coton (d'un actif financier) à l'instant t , le logarithme népérien noté $X(t) = \ln S(t)$ représente le taux de rentabilité continu de l'actif coté entre 0 et t (avec la convention $X(0)=1$). De ce point de vue, en faisant l'hypothèse selon laquelle les variations de l'actif peuvent être représentées par une variable aléatoire qui dépend du temps, alors $X(t)$ est le processus aléatoire du prix de l'actif et τ le « pas de temps » de l'analyse de la variation des prix. Mandelbrot examine l'expression¹⁹ :

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

$Z(t, \tau)$ représente la variation du taux de rentabilité continu de l'actif coté sur la période de τ jours, c'est-à-dire l'accroissement du processus aléatoire $X(t)$ sur cette période.

Mandelbrot choisit pour la durée τ deux valeurs : 1 mois et 1 jour. Puis il considère les queues de distributions empiriques et la fréquence cumulée des grandes variations empiriques observées, notée Fr en calculant les expressions $\text{Fr}(Z(t, \tau) > u)$ pour les grandes variations positives (donc la fréquence cumulée des grandes variations positives) et $\text{Fr}(Z(t, \tau) < -u)$ pour les grandes variations négatives (donc la fréquence cumulée des grandes variations négatives), exactement comme Pareto le fit sur les revenus. Il représente, comme Pareto, les données sur un graphique bilogarithmique en précisant qu'« il nous paraît étonnant que – tout au moins à notre connaissance – les variations des prix n'aient pas été présentées de cette façon jusqu'ici » (Mandelbrot, 1965 : 3968). Les fréquences cumulées sont ajustables par des droites de pentes voisines de $\alpha = 1,7$ soit, pour les valeurs positives :

$$\ln [\text{Fr}(Z(t, \tau) > u)] \approx -\alpha \ln u + \ln C(\tau)$$

Mandelbrot observe aussi que les droites correspondant aux distributions des pas de temps quotidien et mensuel sont parallèles. Il en déduit alors que les lois de distribution des variations quotidienne et mensuelle ne diffèrent que par un *changement d'échelle*. Autrement dit, le phénomène étudié est stable par changement d'échelle, ou encore « invariant d'échelle » : « la loi de distribution de $Z(t, \tau)$ est stable par changement de τ » (Mandelbrot, 1962 : 3968). Mandelbrot propose alors son interprétation : la stabilité du phénomène entre 1 jour et 1 mois peut se comprendre comme la trace de la stabilité au sens de Lévy, et dans ce cas la fonction caractéristique de $Z(t, \tau)$ devrait nécessairement avoir la forme de la fonction caractéristique d'une distribution alpha-stable de Lévy. Autrement dit, le processus aléatoire des cours $X(t)$ pourrait s'analyser comme un processus aléatoire alpha-stable de Lévy.

Comme les distributions alpha-stables sont asymptotiquement paréliennes avec l'indice de queue α , Mandelbrot en conclut que « le caractère parétien de $Z(t, \tau)$ est “prédit” ou “confirmé” par la stabilité » (Mandelbrot, 1962 : 3969). La dernière vérification concerne les constantes sur lesquelles doit aussi s'appliquer l'invariance d'échelle. Montrant que la relation $C(\tau) = \tau C(1)$ est à peu près observée, Mandelbrot propose alors une interpolation de son modèle en temps continu. La représentation des dynamiques boursières par des processus aléatoires de Lévy à queues paréliennes est née.

¹⁹ Nous réécrivons ici l'analyse de Mandelbrot avec les conventions de notation modernes.

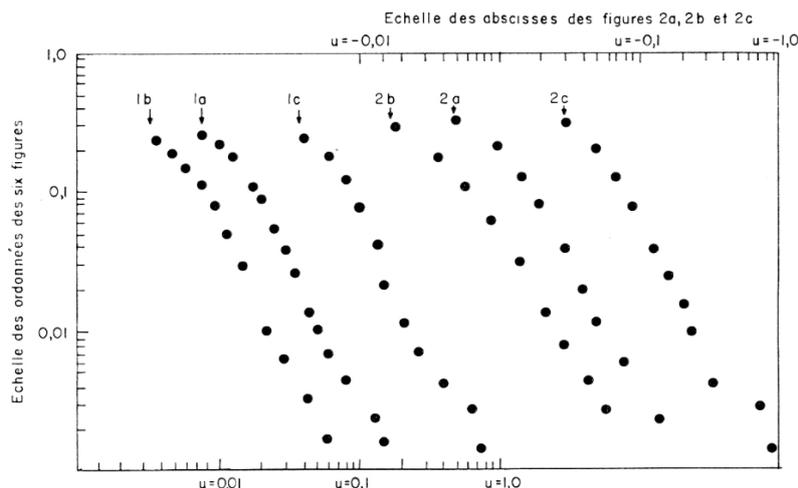


Figure 1. Le parallélisme des fonctions de répartition empiriques $Fr(Z(t, \tau) < -u)$ des variations $Z(t, \tau)$ des prix du coton en représentation bilogarithmique. Les droites sont parallèles de pente $(-\alpha)$. Les constantes $C(\tau)$ sont proportionnelles selon l'échelle τ et la relation $C(\tau) = \tau C(1)$ est à peu près observée, ce qui exprime la stabilité par addition au sens de Lévy soit l'invariance d'échelle de la loi de Pareto. Source : Mandelbrot (1962).

Comme les variations des cours sont dues « au hasard », et qu'il ne s'agit plus d'un hasard gaussien (tirages aléatoires gaussiens) mais d'un hasard régi par une loi de Pareto (tirages aléatoires parétiens), Mandelbrot forge la notion de « hasard parétien », une forme de hasard régie par des mécanismes permettant aux moyennes empiriques de « sauter » de plusieurs ordres de grandeur. Pour cette raison, il appellera le hasard parétien un « hasard sauvage » (*wild chance*) qu'il opposera à la forme classique du hasard, qu'il appellera « hasard sage » (*mild chance*)²⁰. Le hasard sauvage est ainsi un hasard qui correspond à un monde sans moyennes (même si, comme on l'a dit, les calculs de moyennes empiriques sont toujours possibles, bien que la moyennisation produite soit dans beaucoup de cas artificielle).

Le second fait important de cette première mise en évidence empirique de la loi de Pareto en finance concerne la valeur de la pente α . Comme on l'a vu, cette valeur fournit un renseignement sur l'existence des moments de la variable aléatoire car aucun moment d'ordre supérieur à α n'existe. Mandelbrot trouve que α est inférieur à 2 (mais supérieur à 1), ce qui signifie que les distributions $Z(t, \tau)$ ne possèdent pas de moments au-delà du premier. En particulier, le moment centré d'ordre 2, la variance, est infini. Ce qui a pour effet de rendre « inapplicables » (Mandelbrot, 1962 : 3968), la plupart des techniques statistiques utilisées en général pour l'étude des variations boursières. Enfin, les trajectoires des processus alpha-stables de Lévy étant presque partout discontinues, les méthodes usuelles de prise de position sur les marchés et de gestion de ces positions, qui s'appuient sur une hypothèse de continuité des trajectoires boursières, comme les protections de type « *stop loss* », exposent les opérateurs de marché et les gérants de portefeuille à un risque de perte plus important que prévu (ce qui se passera effectivement avec la faillite des gestions en delta neutre en 1987).

²⁰ Pour une définition récente de ces deux notions, voir Mandelbrot & Taleb (2010).

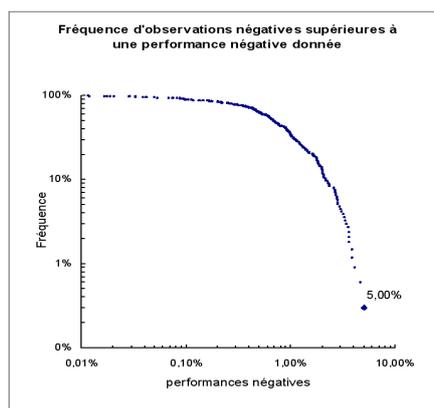


Figure 2. La fonction de répartition empirique $Fr(Z(t, \tau) < -u)$ de la queue de distribution des variations négatives de la performance $Z(t, \tau)$ d'un portefeuille actions en représentation bilogarithmique. Au-delà d'un certain seuil de performance négative (environ -2%), on observe un alignement des points sur une droite de pente $(-\alpha)$ indiquant une loi de puissance sur les grandes valeurs négatives. On retrouve le graphique de Mandelbrot (1962). Source : (Walter, 2013).

À la suite de cette communication, Mandelbrot va déployer son approche des phénomènes économiques et financiers à partir d'un cadre de pensée parétien. Cette étape correspond aux deux articles publiés en 1963 (Mandelbrot, 1963a, 1963b). Dans le premier, il présente son objectif :

Ceci est une présentation informelle et non mathématique de plusieurs approches mathématiques nouvelles pour l'étude des marchés spéculatifs et d'autres phénomènes économiques. Mon hypothèse principale est la suivante : pour pouvoir faire une description exploitable des variations des prix, de la distribution des revenus, de la taille des entreprises etc., il est indispensable d'utiliser des variables aléatoires de variance marginale infinie.

Ce qui a pour effet indirect de réactualiser la loi de Pareto : « qui est omniprésente dans le champ de l'économie, comme dans certains domaines de la physique » (Mandelbrot, 1963b : 421). Aussi, si l'on cherche à réellement représenter la vraie économie, alors « la loi de Pareto s'impose d'elle-même » (Mandelbrot, 1963b : 422). La lecture de ces articles fait apparaître comment l'introduction de la loi de Pareto dans la finance est effectuée par Mandelbrot.

La saga de la loi de Pareto en finance

L'histoire des distributions puis des processus alpha-stables en finance commence. Cette histoire a été racontée ailleurs et on se contente ici de la résumer brièvement, nous permettant de renvoyer le lecteur à des travaux antérieurs (Walter, 2013).

La proposition d'utiliser des lois de Pareto dans la modélisation financière déclencha la « crise leptokurtique » (du grec *leptos*, mince, fin, et *kurtos*, bossu, courbé), une crise de paradigme qui fut suivie par de violentes controverses dans les milieux financiers, tant professionnels qu'universitaires, à la hauteur de la radicalité du changement de cadre conceptuel qui était véhiculé par l'approche parétienne.

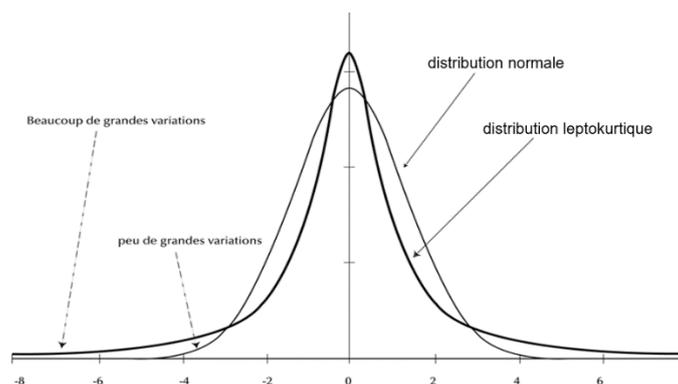


Figure 3. Comparaison entre une distribution normale et une distribution leptokurtique. Par rapport à la distribution normale, la masse de probabilité – égale à 1 par définition – est répartie différemment : elle est « transférée » des flancs (plus creusés) sur la pointe (plus pointue) et les queues de distributions (plus épaisses). Source : Walter.

Dans un premier temps, la résistance des universitaires américains à l'utilisation de la loi de Pareto en finance fut telle que toute observation empirique qui n'entraînait pas dans la modélisation à queues exponentielles était artificiellement modifiée par une troncature appropriée des queues de distribution, cela afin de valider le modèle exponentiel pour éviter les fonctions puissance. Ainsi le futur « prix Nobel » Clive Granger préféra retirer les grandes valeurs des données plutôt que d'accepter l'hypothèse de la loi de Pareto (De Bruin et Walter : 899-901). L'invalidation statistique de la représentation brownienne (donc à loi marginale gaussienne) en finance résista à tout changement de cadre de modélisation, contredisant les règles épistémologiques les mieux établies pour le processus d'évolution des théories scientifiques. Cette résistance au changement représente une « énigme épistémologique » (Walter, 2019) qui ne peut être résolue que par la sociologie.

Il faut noter que l'inexistence d'estimateurs fiables des paramètres des distributions alpha-stables et l'absence d'ordinateurs puissants furent un obstacle sérieux à la pratique statistique, au moins au début. Dès 1965, Eugene Fama lui-même appelait au développement de la recherche d'indicateurs statistiques : « une étape importante qui reste à franchir [pour l'usage des processus alpha-stables en finance] est le développement d'une large panoplie d'outils statistiques pour traiter ces distributions » (Fama, 1965 : 99, notre traduction). Des travaux réalisés sur les marchés financiers dans les années 1980, rendus possibles grâce à de nouveaux résultats obtenus en théorie des probabilités sur les processus aléatoires autosimilaires, en économétrie sur l'étude des séries temporelles, et en statistique sur des méthodes d'estimation robustes des paramètres caractéristiques pour ces séries et processus, levèrent ces obstacles et relancèrent la question de l'utilisation des lois d'échelles en finance, qui apparût de nouveau pertinente, et l'intérêt porté à ce type de lois crût de manière rapide.

Ce renouveau d'intérêt théorique fut par ailleurs soutenu et favorisé par le déploiement dans le champ financier d'une approche institutionnelle des risques de marché et de crédit, dite « prudentielle », qui se traduisit par l'existence de groupes de réflexion internationaux dont l'objectif fut de promouvoir des règles de quantification des risques universellement applicables par les établissements financiers dans tous les domaines de leur activité. Ce souci prudentiel correspondit à l'augmentation rapide des encours d'instruments financiers appelés « dérivés », dont l'instabilité et la difficulté de gestion constituaient des sources grandissantes de préoccupation pour les autorités de tutelle des marchés. Assez rapidement, et d'autant plus que les crises multiples qui agitaient les marchés boursiers venaient poser des problèmes importants et non résolus pour le contrôle des risques, il est apparu à ces instances représentatives que les outils et méthodes classiques, tous fondés sur l'utilisation de techniques gaussiennes, c'est-à-dire instrumentant des probabilisations de risque de perte à partir de la loi normale, ne

permettaient pas de disposer de mesures fiables pour la quantification, et donc la prévention, de ces risques. Un besoin de nouvelles méthodologies et de nouveaux outils apparût, qui contribua à accélérer la recherche théorique sur les lois d'échelle et leur applicabilité opérationnelle. Cette convergence d'intérêt fut à l'origine de la redécouverte de la loi de Pareto en finance.

Dans les années 1990, la loi de Pareto généralisée connût une renaissance fulgurante et apparût rapidement comme un outil pertinent pour le traitement adéquat des queues de distribution. Dans les années 2000, cette loi devint la boîte à outils classique de la modélisation financière pour les queues de distribution (Longin, 2017). Aujourd'hui, dans les pratiques professionnelles des financiers, la loi de Pareto généralisée fait désormais partie des instruments utilisés sans questionnement. En soixante ans, entre l'année de sa réapparition en finance en 1962 et aujourd'hui, la loi de Pareto sera passée d'un « objet non identifié » voire dangereux à un jouet naturel pour ingénieurs financiers et actuaires.

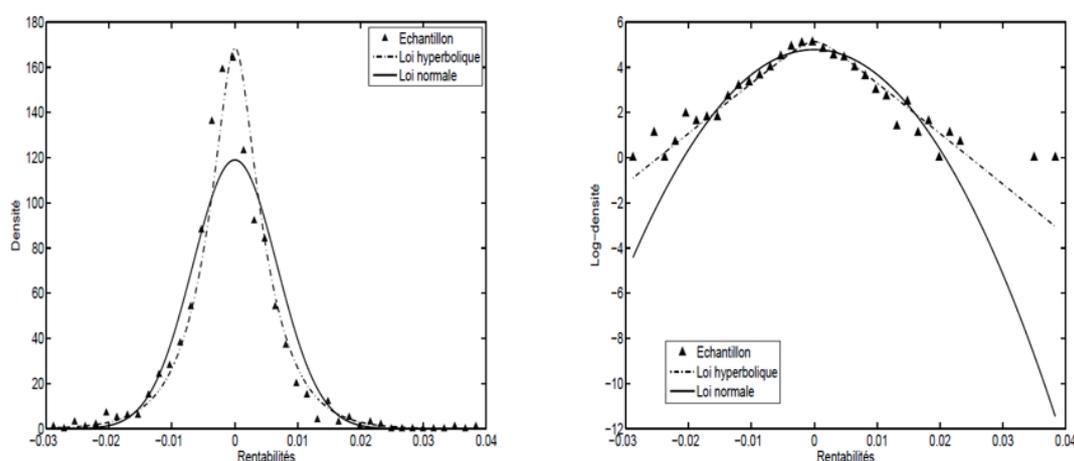


Figure 4 : Ajustement des distributions normale et hyperbolique sur les variations de l'action BNP en graphique classique à gauche et semi-logarithmique à droite. Le graphique semi-logarithmique transforme la gaussienne en parabole et fait mieux apparaître les écarts des points sur les queues de distribution. Les extrêmes gaussiens sont nettement distingués des extrêmes hyperboliques. Source : Le Courtois et Walter (2012).

Les mathématiques « naturelles » du non coté

Il reste un domaine de la finance qui n'a pas, jusqu'à présent, été touché par la modélisation parétienne, celui du capital investissement appelé aussi le « non coté ». Par économie non cotée, on entend l'ensemble des sociétés commerciales ou des infrastructures qui se trouvent en dehors du marché boursier. Investir dans l'économie non cotée signifie prendre une participation à la vie économique par un investissement en « fonds propres » d'une entreprise, au moyen de véhicules de placement appelés « fonds d'investissement »²¹. Nous voudrions ici proposer une idée nouvelle pour montrer l'actualité de la loi de Pareto dans ce secteur. Cette loi pourrait représenter l'outil qui correspondrait aux « mathématiques naturelles » (Chemillier, 2007) des professionnels du non coté.

L'hypothèse faite est la suivante. Avec un cadre parétien, il serait possible de restaurer un calcul probabiliste et une rationalité à l'endroit où tout n'est perçu que comme incertitude radicale ou irrationalité des professionnels et entrepreneurs. Les fonctions puissance des lois de Pareto pourraient permettre de réunifier dans une même construction de la quantification de l'incertitude, dans un même cadre d'analyse, les aspects économiques et spéculatifs de l'activité

²¹ Pour une introduction, voir Demaria, 2020 ; Bédu, Foureault & Palard, 2022 ; et Mauboussin & Callaghan, 2020, qui contient une bibliographie exhaustive.

d'investissement dans le non coté, les attentes des investisseurs et les actions des entrepreneurs. Ces deux dimensions seraient saisies dans un même mouvement intellectuel, dont la clé repose sur le hasard parétien.

Cette proposition est ainsi une tentative de réponse à une exigence rationnelle adressée à l'investissement en capital pour éviter le pessimisme à l'endroit « des mauvais calculs et des déraisons de l'homme économique » (Revue *Esprit*, 2009). Les agents qui présentent des attitudes apparemment irrationnelles en regard de la théorie du choix rationnel (comme les professionnels du non coté) pourraient malgré tout être considérés comme rationnels en admettant qu'ils se représentent l'incertitude économique par des lois de puissance. Notre proposition revient à suggérer que l'« homme extrême » de Pareto (Barbut, 1999) serait rationnel d'une façon extrêmement différente.

Ethnomathématique du capital investissement

Nous adoptons une démarche sociologique et ethnologique, en partant d'entretiens avec les professionnels du non coté, pour saisir ce qu'ils ont à l'esprit, leurs « opérations mentales » quand ils prennent une décision d'investissement. De ce point de vue, la démarche que nous proposons rejoint celle de l'ethnomathématique²², trouver des structures mathématiques dans une activité humaine qui ne se pense pas comme mathématique. Bien qu'appliquée initialement sur les sociétés indigènes, nous étendons cette méthodologie sur la population des professionnels du non coté en cherchant à montrer que, dans la pratique du capital investissement, on trouve la présence de processus de décision et de savoirs pratiques qui semblent présenter un caractère mathématique, mais sans écriture. En un sens, nous cherchons à identifier les « mathématiques naturelles » du capital investissement. Pour cela, nous associons l'analyse mathématique et l'enquête de terrain. Le terrain est celui du monde du capital investissement, l'analyse mathématique est celle du cadre parétien présenté précédemment.

Le modèle mental des professionnels du non coté

Partons d'entretiens que nous avons eus avec des investisseurs en capital en prenant au sérieux la manière que les professionnels du non coté ont de parler de leur activité. Chez les investisseurs en capital, la phrase type la plus souvent exprimée est : « le produit de cession espéré est un multiple du montant investi »²³. Cette phrase soulève de nombreuses questions, à la fois cognitives et mathématiques.

Écrivons mathématiquement cette phrase. Notons V le produit de cession et V_0 la valeur du montant investi au départ et K le « multiple » de l'investissement initial. Alors l'écriture mathématique de la phrase du professionnel interrogé est :

$$E [V | V > V_0] = K \times V_0$$

On retrouve la relation des moyennes conditionnelles linéaires dans le monde parétien. Le modèle qui représente le mieux ce que ce professionnel semble avoir à l'esprit est une représentation parétienne des prix de cession. Le processus cognitif des professionnels du non coté semble formé par une perception intuitive des propriétés parétiennes de l'économie (queues de distribution ultra-étirées). On pourrait dire que le modèle mental des professionnels du non coté « incarne » les propriétés mathématiques de la loi de Pareto. Plus précisément que les *opérations mentales* mises en œuvre par l'investissement en capital reviennent à « faire de la mathématique » parétienne. Ceci veut dire que, lorsque les professionnels du non coté prononcent la phrase « l'espérance du prix de sortie est un multiple du prix d'entrée », tout se

²² Sur l'ethnomathématique, on pourra consulter Vandendriessche (2018).

²³ Entretiens avec l'auteur.

passer comme si l'investissement dans le non coté était pensé en termes parétiens, même sans connaître formellement les mathématiques relatives à la loi de Pareto.

Si l'environnement économique dans lequel évolue la prise de participation est parétien, alors les risques qui affectent les investissements et donc les prix de cession possibles, seront parétiens. C'est-à-dire que le hasard qui gouverne les fluctuations du prix de cession sera parétien. Dans une économie parétienne, très peu de produits de cession sont très élevés et beaucoup de produits de cession sont peu élevés, les produits de cession sont fortement hiérarchisés au-delà d'un certain seuil possible de prix de sortie (on retrouve la « loi des 80/20 »). On peut comprendre cette situation comme un événement singulier A qui, s'il advient, produirait une très forte croissance de la société, et donc de la valeur du produit de cession. En deçà de ce seuil, la hiérarchisation des produits de cession s'atténue : si l'événement singulier A n'advient pas, alors les produits de cession potentiels resteront dans des limites plus classiques. De manière pratique, on peut imaginer une situation dans laquelle le produit de cession sera très élevé dans deux ou trois cas, puis plus faible sinon.

La probabilité du produit de cession qui dépend du franchissement d'un « produit de cession-seuil » peut se calculer dans un cadre parétien. C'est la distribution de probabilité du produit de cession V au-delà du produit de cession V_A , c'est-à-dire si l'événement A advient. Si l'environnement économique est parétien, V suit une loi de Pareto d'exposant alpha. Dans ce cas, $K = \alpha / (\alpha - 1)$ et le prix de cession espéré au-delà du seuil V_A si ce seuil est dépassé est :

$$E [V | V > V_A] = \frac{\alpha}{\alpha - 1} \times V_A$$

Pour le dire autrement, le produit de cession espéré si on atteint un certain niveau d'évaluation est un multiple de ce niveau. Cette manière de penser est la « signature » d'un cadre mental parétien. Par exemple, si $\alpha = 1,5$, alors $K = 3$. Cela veut dire que la valeur la plus probable, une fois atteint le seuil, est le triple du seuil lui-même.

Il est légitime de considérer le langage professionnel des investisseurs en capital comme leur « tradition orale ». Dans cette tradition, on dit que « si à cette date on évalue la société 100, alors on pourra espérer une cession à 300 », ou bien « si l'évaluation dépasse 100, alors l'objectif de cession est de 300 », etc. Une forme de fabrique narrative : on utilise une tradition orale pour raconter une histoire sur le futur prix de la société. Dans un cadre de pensée parétien, tout franchissement de seuil (toute nouvelle évaluation) produit automatiquement un nouvel objectif de prix de cession espéré, toujours plus élevé et d'autant plus que l'exposant alpha de Pareto se rapproche de sa valeur limite 1. Par exemple, si $\alpha = 1,1$ le multiple qui produit l'objectif de produit de cession est $\alpha / (\alpha - 1) = 11$. Les gérants du fonds d'investissement se fixeraient donc un produit de cession espéré onze fois supérieur à leur montant investi. Remarquons que l'on retrouve des dynamiques de prix qui coïncident avec les emballements spéculatifs de marché tels ceux qui se sont produits sur les valeurs Internet lorsque les acheteurs de ces actions raisonnaient en termes de multiples du chiffre d'affaires. Le cadre parétien permet de trouver une même rationalité pour les secteurs cotés et non cotés.

Signification de l'exposant de Pareto pour le non coté

Pour donner une base cognitive réelle à ces propriétés parétiennes abstraites, cherchons à représenter comment le paramètre alpha pourrait s'incarner dans les pratiques du non coté. Considérons un fonds d'investissement j qui va prendre une participation sur une société k à la date t . Nous considérons que le paramètre alpha traduit l'appréhension subjective par l'équipe de gestion du fonds j de l'incertitude sur la société k au temps t et on indice pour cette raison le paramètre alpha par alpha (j, k, t) les indices faisant apparaître l'environnement concret de cette perception subjective. Quatre situations sont alors possibles pour les façons de considérer les

fluctuations affectant la vie économique des sociétés et donc le prix de sortie. Chaque cas correspond à un principe de variabilité pour alpha ou à une catégorie de type d'investissement dans le non coté. Nous employons ci-après la notation « ● » pour symboliser l'homogénéisation de alpha selon l'un ou l'autre des indices retenus²⁴.

Premièrement, on peut imaginer que la perception de l'incertitude sur les sociétés en général (et pas seulement la société examinée) dépend de l'équipe de gestion du fonds qui observe l'économie, que cette perception caractérise l'attitude de l'équipe de gestion, son degré d'appréhension de l'incertitude en soi. On peut complexifier cette idée et imaginer que cette attitude devant le futur évolue au cours du temps (ce qui revient à modifier la distribution de Pareto). Il s'agit d'un point de vue subjectiviste quant à l'équipe de gestion mais homogène quant aux entreprises : toutes les entreprises sont appréciées et soupesées avec la même grille subjective de perception de l'incertitude, qui est propre à une équipe de gestion. Chaque équipe de gestion adopte une politique d'évaluation qui lui est propre. Chacune a sa manière d'appréhender le risque. Des équipes de gestion peuvent avoir une meilleure connaissance d'un secteur économique que d'autres, des déterminants de l'activité considérée, de la capacité à évaluer la direction de l'entreprise etc. Dans ce cas, alpha ne dépend que de l'équipe de gestion et du temps et tout événement peut advenir, qui conduirait les gérants à réajuster leurs prévisions sur le prix de sortie. En particulier chaque équipe de gestion va appréhender différemment l'événement déclencheur de grandes espérances de produit de cession. Cette manière de voir les choses permet d'introduire une composante subjective dans l'évaluation de la participation non encore cédée, c'est-à-dire l'appréciation du seuil de franchissement. Dans ce cas, la distribution conditionnelle des produits de cession n'a *pas d'échelle propre* mais seulement celle de l'équipe de gestion du fonds. Cette absence d'échelle propre est aussi une caractéristique d'une économie parétienne dans laquelle chaque agent projette sa propre échelle, nécessairement subjective, sur le phénomène observé. Dans ce cas, $\alpha = \alpha(j, \bullet, t)$. Le paramètre alpha ne dépend que des investisseurs j (de l'équipe de gestion du fonds) et du temps t , et non de la société.

À l'opposé, on peut imaginer des contextes dans lesquels le paramètre alpha ne dépend que de la société et du temps et pas des investisseurs. L'incertitude sur la société objet de la prise de participation est, soit « objective » soit admise par consensus de place. L'aléa qui quantifie l'incertitude sur la société est propre à cette société : il s'agit d'un point de vue objectiviste. Ce qui revient à dire que tous les investisseurs utiliseraient les mêmes critères d'appréciation de l'incertitude. Dans ce cas, $\alpha = \alpha(\bullet, k, t)$ et ne dépend que de la société et du temps, et non des investisseurs (de l'équipe de gestion du fonds).

Si l'entreprise est incertaine « en soi », indépendamment des points de vue subjectifs des investisseurs et du moment où on la considère, alors on aura $\alpha = \alpha(\bullet, \bullet, \bullet)$ non indicé. Le paramètre alpha est uniforme et indépendant du temps, des investisseurs et des entreprises.

Investisseurs et régulateur : Pareto contre Quetelet

L'insatisfaction des professionnels du non coté pour la manière de mesurer le risque imposée par le régulateur est connue. Elle vient de ce que le régulateur utilise des quantités comme la volatilité qui, aux yeux des professionnels du non coté, ne sont pas adaptées à la réalité de l'activité d'investissement. Pour quantifier le risque d'un investissement dans le non coté, les mesures usuelles à base de moyenne qui reposent sur une représentation normale des risques comme la volatilité, ne sont pas pertinentes car la hiérarchie forte des produits de cession potentiels (la « pointe de la courbe ») est parétienne. La vision à base de moyennes, comme celle imposée par les autorités de tutelle relève d'une *erreur de perspective* sur la nature des

²⁴ Nous suivons ici et développons l'intuition présentée dans Walter & Brian (2008), p. 154.

distributions des phénomènes économiques réels. Les professionnels du non coté n'ignorent pas la question du risque dont ils savent qu'elle est bien réelle et présente dans leur activité, à la fois à un niveau d'investissement sur une seule société mais aussi dans le cas d'un portefeuille d'un fonds d'investissement. Mais les approches existantes les laissent en général insatisfaits. L'ethnomathématique permet d'éclairer l'origine de cette incompréhension.

Il y a discordance entre le modèle mental des professionnels du non coté et le modèle mental des autorités de tutelle. Le raisonnement du régulateur est incompatible avec le raisonnement des professionnels du non coté. En un sens, dans les délibérations entre professionnels et autorités de régulation financière sur ce qui serait « raisonnable » pour la sécurité financière et la société, on assiste à une collision entre deux visions du monde : le monde des moyennes et le monde de Pareto. Il apparaît alors un conflit inéluctable entre les mathématiques non écrites des professionnels du non coté et les mathématiques écrites des autorités de tutelle et de régulation des marchés financiers. Le dialogue institutionnel échoue faute d'une croyance partagée sur le risque en situation d'incertitude radicale.

En utilisant les ressources de la littérature française, nous pourrions considérer que les investisseurs en capital, confrontés aux méthodes et aux outils à base de moyenne et de volatilité forgés pour eux par les autorités de tutelle et les normes réglementaires prudentielles internationales, ressemblent au savetier de la fable tentant d'expliquer au financier que cette manière de compter ne correspond pas à sa pratique effective. Ainsi, Jean de La Fontaine dans *Le Savetier et le Financier* présentait-il les réticences du savetier devant les tentatives d'estimation du financier (Walter & Brian, 2008 : 150) :

- Que gagnez-vous par an ?
- Par an ? Ma foi, Monsieur, [...] ce n'est point ma manière de compter de la sorte.
- Eh bien ! que gagnez-vous, dites-moi, par journée ?
- Tantôt plus, tantôt moins, le mal est que toujours (et sans cela nos gains seraient assez honnêtes), le mal est que dans l'an s'entremêlent des jours [...]

Comme le régulateur aujourd'hui, le financier de la fable cherche à établir un calcul à base de moyennes. Le savetier lui répond qu'il ne calcule pas ainsi. Comme l'investisseur en capital, il ne s'attend pas à une performance régulière. Or les calculs de risque imposés par le régulateur nécessitent des performances régulières. Le système de représentation du savetier, qui est celui d'un monde sans moyenne théorique, contrevient aux hypothèses de régularité du système de représentation du financier, qui est celui d'un monde dans lequel les moyennes théoriques existent (théorème central limite dans sa forme classique). Le risque d'un investissement dans le non coté n'est pas évaluable par les représentations usuelles de la théorie financière classique qui est construite sur la validation de la notion de moyenne « débarrassée de ses accidents » (Feldman, Lagneau & Matalon, 1991 : 20). Le hasard parétien, ou hasard correspond aux lois en fonction puissance, est la plupart du temps un hasard sans moyenne théorique. L'histoire des sciences enseigne que le XIX^e siècle fut le siècle des moyennes sous l'influence de la théorie de l'homme moyen de Quetelet. Contre un régulateur queteletésien, les investisseurs en capital se situeraient donc du côté de Pareto.

On connaît la réponse du financier « riant de la naïveté » du savetier. Trois cents ans après La Fontaine, la mise en évidence de raisonnements parétiens dans les pratiques du non coté pourrait permettre aux professionnels de l'investissement en capital de sourire (ou de se lamenter) des autorités épistémiques de réglementation prudentielle qui imposent un calcul queteletésien à base de moyennes théoriques partout (les *benchmarks* de performance) et qui tentent d'enfermer dans des cadres réglementaires inadaptés – car construit sur des théorisations de la moyenne – les intuitions persistantes des professionnels, en fin de compte moins ignorants des risques réels

encourus que l'application trop rapide de la théorie financière néoclassique chercherait à le faire croire. L'ethnomathématique permet d'établir une correspondance entre certaines propriétés de la loi de Pareto et des modes de penser spécifiques aux investisseurs en capital.

Conclusion

Nous avons présenté en détail le cadre parétien, à la fois dans ses aspects mathématiques mais aussi comme modèle mental pour éclairer d'un regard neuf les opérations mentales mises en œuvre dans la pratique professionnelle du capital investissement, domaine qui échappait jusqu'à présent aux modélisations tentées par la théorie financière néoclassique. L'exposé de la manière dont Pareto a conçu sa loi et de la correspondance entre les fonctions puissance et les moyennes conditionnelles linéaires nous a permis de saisir l'originalité du cadre de pensée parétien et l'arrière-plan cognitif de la notion de hasard parétien.

L'histoire de l'entrée de la loi de Pareto dans la modélisation financière nous a permis de suivre l'évolution de ce cadre de pensée nouveau par rapport au cadre de la théorie financière néoclassique, un cadre nouveau qui fut d'abord violemment rejeté par la communauté universitaire avant d'être finalement accepté avec le développement des approches prudentielles des marchés. Nous avons aussi insisté sur une façon de différencier la finance parétienne de la finance néoclassique, à partir de l'usage des moyennes. Alors que la finance néoclassique repose sur un usage extensif de la notion de moyenne, une finance parétienne s'en éloignerait radicalement en considérant que les moyennes théoriques peuvent ne pas exister à cause des lois de puissance, et donc que toute moyennisation empirique trop rapidement entreprise relèverait plus ou moins d'une artificialisation du phénomène examiné.

Le point par lequel nous voudrions conclure est notre proposition d'aborder le secteur du non coté par une approche ethnomathématique. L'absence de modélisation efficace pour ce secteur est peut-être due au fait que les mathématiques adaptées à cette activité étaient ignorées des modélisateurs qui l'étudiaient, tandis que les professionnels, suivant une tradition orale solidement établie, n'étaient pas en mesure de les décrire et de les étudier. Notre proposition a cherché à relier ces deux mondes et à faire apparaître ainsi une conjecture sur le fonctionnement de la pensée rejoignant de manière indirecte le « principe de Pareto », exprimer une exigence de rationalité s'appliquant la vie sociale en général²⁵. Une conjecture parétienne, en quelque sorte.

Il s'agit d'un champ de recherche encore en chantier à la croisée de la sociologie du non coté et des mathématiques parétiennes. Nous espérons que notre proposition encouragera les travaux dans ce champ encore nouveau, pour faire progresser notre compréhension de la nature mathématique de l'activité d'investissement en capital. En regard des montants investis sur le non coté et des enjeux colossaux de ce secteur d'activité pour le financement de la transition énergétique face au risque climatique, il y a urgence à ce que les mathématiciens et les sociologues s'associent pour analyser cette activité.

²⁵ Je remercie Emmanuel Picavet d'avoir porté à ma connaissance cet aspect de la pensée de Pareto, qu'il développe dans Picavet (2013 : 167).

Références bibliographiques

- Anderson C., 2006, *La Longue Traîne. Quand vendre moins c'est vendre plus*, Paris, Flammarion, coll. « Champs essais ».
- Bak P. 1997, *How nature works. The science of self-organized criticality*, Oxford, Oxford University Press.
- Balkema A. A., de Haan, L. 1974, “Residual lifetime at great age”, *The Annals of Probability*, 2, 5, p. 792-804. DOI :10.1214/aop/1176996548
- Barbut M. 1989, « Distributions de type parétien et représentation des inégalités » *Mathématiques et sciences humaines*, 106, p. 53-69.
- Barbut M. 1991a, « Note sur les moyennes de variables aléatoires », dans J. Feldman, G. Lagneau, B. Matalon (dir.), *Moyenne, Milieu, Centre. Histoire et usages*, Paris, Éditions de l'EHESS, p. 31-43.
- Barbut M. 1991b, « Moyennes et valeurs centrales », dans J. Feldman, G. Lagneau, B. Matalon (dir.), *Moyenne, Milieu, Centre. Histoire et usages*, Paris, Éditions de l'EHESS, p. 135-151.
- Barbut M. 1998, « Une famille de distributions : des parétiennes aux “contra-parétiennes” », *Mathématiques et sciences humaines*, 141, p. 43-72.
- Barbut M. 1999, « Pareto et la statistique. L'homme extrême de Pareto : sa postérité, son universalité », dans A. Bouvier (dir.), *Pareto aujourd'hui*, Paris, Puf, 1999, p. 85-109.
- Barbut M. 2004, « Une famille de distributions : des parétiennes aux contra-parétiennes », *Mathématiques et sciences humaines*, 141. DOI : 10.4000/msh.2777
- Barbut M. 2007, *La mesure des inégalités. Ambiguïtés et paradoxes*, Genève, Droz.
- Bédu N., Foureault F., Palard J.-E. 2022, *Le Private Equity*, Paris, La Découverte, coll. « Repères ».
- Bouchaud J.-P. 2001, “Power-Laws in Economy and Finance: Some Ideas from Physics”, *American Finance Association New Orleans Meetings, 2001*, DOI : 10.2139/ssrn.240608
- Bouchaud J.-P. 2021, *De la physique statistique aux sciences sociales*, Paris, Fayard, coll. « Leçons inaugurales du Collège de France ».
- Bouvier A. (dir.) 1999, *Pareto aujourd'hui*, Paris, Puf.
- Chemillier M. 2007, *Les Mathématiques naturelles*. Paris, Odile Jacob.
- De Bruin B., Walter, C. 2017, “Research habits in financial modelling: the case of non-normality of market returns in the 1970s and the 1980s”, dans E. Ippoliti et C. Ping (eds.), *Methods and Finance. A Unifying View on Finance, Mathematics and Philosophy*, Springer International, p. 79-93.
- Demaria C. 2020, *Introduction au private equity*, Paris, Revue Banque édition.
- Embrechts P., Klüppelberg C., Mikosch T. 1997, *Modelling Extremal Events for Insurance and Finance*, Berlin, Springer.
- Fama E. 1965, “The behavior of stock market prices”, *Journal of Business*, 38, 1, p. 34-105.
- Feldman J., Lagneau G., Matalon B. (dir.) 1991, *Moyenne, Milieu, Centre. Histoire et usages*, Paris, Éditions de l'EHESS.

Fisher R. A., Tippett L. (1928), "Limiting form of the frequency distribution of the largest or smallest member of a sample", *Mathematical Proceedings of the Cambridge Philosophical Society*, 24, 2, p. 180-90. DOI : 10.1017/S0305004100015681

Fréchet M. 1925, "Une nouvelle représentation analytique de la répartition des revenus », XVIe session de l'Institut international de statistique, Rome.

Gabaix X. 2009, "Power Laws in Economics and Finance", *Annual Review of Economics*, 1, p. 255-294.

Gnedenko B. V. 1943, « Sur la distribution limite du terme maximum d'une série aléatoire », *Annals of Mathematics*, 44, p. 423-453.

Ibragimov R. 2005, « On the robustness of economic models to heavy-tailedness assumptions », Working Paper, Yale University.

Koch R. 1997, *Le principe 80/20. Faire plus avec moins*, Paris, De l'homme éditions.

Le Courtois O., Walter C. 2012, *Risques financiers extrêmes et allocation d'actifs*, Paris, Economica.

Lévy P. 1924, « Théorie des erreurs. La loi de Gauss et les lois exceptionnelles », *Bulletin de la société mathématique de France*, 52, p. 49-85.

Lévy P. 1937, *Théorie de l'addition des variables aléatoires*, Paris, Gauthier-Villars.

Longin F. (dir.) 2017, *Extreme Events in Finance: A Handbook of Extreme Value Theory and its Applications*, New Jersey / Canada, John Wiley & Sons, coll. "Handbooks in Financial Engineering and Econometrics".

Mandelbrot B. 1959a, « Variables et processus stochastiques de Pareto-Lévy, et la répartition des revenus », *Comptes rendus à l'Académie des sciences*, 249, p. 613-615.

Mandelbrot B. 1959b, « Variables et processus stochastiques de Pareto-Lévy, et la répartition des revenus », *Comptes rendus à l'Académie des sciences*, 249, p. 2153-55.

Mandelbrot B. 1962, « Sur certains prix spéculatifs : faits empiriques et modèle basé sur les processus stables additifs non gaussiens de Paul Lévy », *Comptes rendus à l'Académie des Sciences*, 254, p. 3968-3970.

Mandelbrot B. 1963a, « The Variation of Certain Speculative Prices », *Journal of Business*, 36, 4, p. 394-419.

Mandelbrot B. 1963b, « New methods in statistical economics », *The Journal of Political Economy*, 71, p. 421-440.

Mandelbrot B. 1967, « Sur l'épistémologie du hasard dans les sciences sociales. Invariance des lois et vérification des prédictions », dans Piaget J. (dir.) *Logique et Connaissance Scientifique*, Paris, Encyclopédie de La Pléiade, p. 1097-1113.

Mandelbrot, B., Taleb N., 2010, « Mild vs. Wild randomness. Focusing on Those Risks That Matters », dans Diebold F., Doherty N. & Herring R., *The Known, the Unknown, and the Unknowable in Financial Risk Management*, Princeton, Princeton University Press, p. 47-58

Mantegna R., Stanley E. 2000, *An Introduction to Econophysics. Correlations and Complexity in Finance*, Cambridge, Cambridge University Press.

- Manzo G. 2009, « Marc Barbut, La mesure des inégalités. Ambiguïtés et paradoxes, Genève-Paris, Droz, 2007 », *L'Année Sociologique*, 59, 2, p. 507-512.
- Mauboussin M., Callaghan D. 2020, "Public to Private Equity in the United States: A Long-Term Look", Morgan Stanley Investment Management.
- Pareto V. 1965 [1895], *Écrits sur la courbe de répartition des richesses*, réunis et présentés par G. Busino, Genève, Droz.
- Picavet E. 2013, « Réflexions sur la normativité collective et l'héritage de l'approche parétienne », *Revue européenne des sciences sociales*, 51-2, p. 149-171.
- Pickands J. 1975, "Statistical inference using extreme-order statistics", *Annals of Statistics*, 3, p. 119-131. DOI: 10.1214/aos/1176343003.
- Revue européenne des sciences sociales*, 51(2), « 1963-2013 : Autour de Pareto ».
- Revue Esprit* 2009, « Les mauvais calculs et les déraison de l'homme économique », juin.
- Samuelson P. 1972, *L'économie*, Paris, Armand Colin.
- Sorel G., 1897, « La loi des revenus », *Le devenir social*, 7, p. 578-607.
- Steiner P. 2013, « La courbe parétienne des revenus et la sociologie économique », *Revue européenne des sciences sociales*, 51, 2, p. 193-211.
- Vandendriessche E. 2018, « Anthropologie des nombres et ethnomathématique », *L'Homme*, 225, p. 185-198, DOI : 10.4000/lhomme.30732
- Walter C. 2013, *Le modèle de marche au hasard en finance*, Paris, Economica.
- Walter C. 2019, "The Brownian motion in finance: an epistemological puzzle", *Topoi*, 40, p. 1-17. DOI : 10.1007/s11245-019-09660-7
- Walter C., Brian E. 2008, *Critique de la valeur fondamentale*, Paris, Springer-Verlag France.
- West G. 2017, *Scale. The Universal Laws of Growth, Innovation, Sustainability, and the Pace of Life in Organisms, Cities, Economies, and Companies*, New York, Penguin Press.
- Zajdenweber D. 2000, *L'économie des extrêmes*, Paris, Flammarion, coll. « Nouvelle bibliothèque scientifique ».