# "New Methods of Statistical Economics," Revisited: Short versus Long Tails and Gaussian versus Power-Law Distributions

*The standard "Brownian" model of competitive markets asserts that the increments of price (or of its logarithm) are statistically independent and Gaussian, implying that price itself is a continuous function of time. This model arose in 1900, at an immediately high level of perfection, in the work of L. Bachelier. In many fields of science it became a classic. But for financial prices it is sharply—even overwhelmingly—contradicted by conspicuous and strong symptoms of non-independence, nonGaussianity, and discontinuity. Since 1963, the author has been tackling those symptoms one by one: first, by incorporating strongly non-Gaussian marginal distributions (1963), then by incorporating strong long-term dependence (1965), and finally by combining those two features by introducing the new notion of multifractality (1968 and since). The goal of this article is modest: largely borrowing from the author's previous articles and books, much of it collects and adapts a few brief "teasers" or "appetizers" meant to promote and assist the future development of a framework for a realistic description of actual financial fluctuations. © 2008 Wiley Periodicals, Inc. Complexity 14: 55–65, 2009*

**Key Words:** risk; the starkly non-Brownian character of competitive prices; new models for mild versus wild statistical variability

**BENOIT B. MANDELBROT**

*Benoit B. Mandelbrot is Sterling Professor Emeritus of Mathematical Sciences at Yale University, New Haven, Connecticut 06520-8283 (e-mail: benoit.mandelbrot@yale.edu)*

## 1. INTRODUCTION

To elaborate on the title, the bulk of this article sketches and comments on two 45-years-old articles of mine. They were widely read, provoked a storm of criticism to be exemplified in this article, but have been slow to catch on. However, under the inescapable pressure of events, they are becoming very influential. *New Methods of Statistical Economics*—to be referred to as *New Methods*—is the title of [Ref. 1]. *The Variation of Certain Speculative Prices* [2] will be referred to as *Certain Prices*.

*This is an invited article for the Special Issue—Econophysics, Guest Editors: Martin Shubik and Eric Smith.*


DOI 10.1002/cplx.20264
Published online 25 November 2008 in Wiley InterScience
(www.interscience.wiley.com)

These and other articles that much expanded on the same general topic (including [3–8]), have been republished in a book [9] that centers on slightly copyedited reprints (whose nonedited originals remain easily available).

The key themes common to these articles related to the flippant saying that "The Winner Takes (Almost) All." In many forms, this is a fact familiar in intuitive finance but flatly contradicted in the standard quantitative elaborations. The "winner" depends on the specific case. In the most vital context, much of the total change in a competitive price over a long enough period occurs during a few particularly active trading days.

Also mentioned in the title are related and more precise key themes common to those 1963 articles and the present one: "short versus long distribution tails" and "Gaussian versus power-law distributions."

My early texts were more often criticized than elaborated on. But that period has ended and they have already moved—or seem ready to move—into the limelight. They permeate a nonmathematical book that I coauthored, *The (Mis)behavior of Markets* [10]. Also, the theme of fat tails dominates several recent books written for an even broader audience, such as [Ref. 11]

On a detailed and technical level, those and related themes have already been revisited repeatedly in a number of publications ([12–29]). But they did not "catch" until recently. This is why the main theme deserves to be visited again.

It remains as a challenge to any effort to quantify the economics of inequality—an issue that has long been put aside but has by now become ubiquitous.

The great pioneer, Vilfredo Pareto, discovered in the 1880s that the percentage of earners with a personal income above $u$ has a hyperbolic distribution. It takes the form $Fu^{-\alpha}$, which has two parameters: the exponent $\alpha$ and the prefactor $F$. This negative power of $u$ is referred to as a "power law" distribution. Plotting the logarithm of that percentage against the logarithm of income, Pareto obtained a straight line of slope $-\alpha$. In contrast with the tail of the Gaussian randomness, which is vanishingly short, power law randomness has a very long tail. Its "fatness" expresses that, while extreme events are rare, they follow a law and are of great and often overwhelming importance. My early article introduced issues and tools that were new and sophisticated for their time and transformed the power-law distribution from a curio to a key fact in economics.

The methods, though not the technical results, have as key ingredient an explicit reliance on statistical thermodynamics, either in its classical form or as broadened in various ways, as I began to do in 1951 to account for a fact far from economics, namely Zipf's law for word frequencies [See figure in Ref. 30]. Some of the oldest areas of physical science, especially turbulence, raise problems that recall those of economics and that the usual methods of physics still struggle to tackle. I foresee a broader development that, instead of contrasting physical and other sciences, will contrast two very different forms of randomness, which I called "mild" and "wild." This contrast will be explored.

In the courses on statistics that underlie economics, a dominant role continues to be played by the Gaussian. This fosters would-be models that also imply that extreme values are negligible—an assumption that is thoroughly distressing.

In addition, another article of mine appeared only in French, with a title that ends with *Vérification des Hypothèses* [29]. Altogether, many aspects of *New Methods* and those other old articles have not yet reached the audience they can assist; they should therefore benefit from being restated, as is done in this article. To simplify its organization, sections are short and fairly independent.

## 2. PARETO'S LAW OF INCOME DISTRIBUTION IS THE ONLY ONE THAT COULD POSSIBLY BE SATISFIED BY THE MANY ALTERNATIVE DEFINITIONS OF PERSONAL INCOME

Vilfredo Pareto's discovery was, arguably, the earliest empirical quantitative law to actually originate in economics. It also remains the most solidly confirmed law in its field. Not a reason to minimize credit due to Pareto, he believed the exponent $\alpha$ to be universal, but later authors showed that it is not. In fact, its value provides an important quantitative comparative measure of inequality among high earners.

When trying to gain a fuller appreciation of Pareto's law, *New Methods* was the first to raise a disturbing preliminary issue. How could income follow any universal law whatsoever? The hyperbolic law seemed to me too good to hold, because different countries and periods define and report income in arbitrary and highly variable ways and tolerate different kinds and degrees of creative accounting.

The "universality" of the hyperbolic law helps explain why Pareto could discover it so early, but the fact that such "dirty" data follow a universal law seemed too good to hold. It made me wonder. Forgetting for a moment its cause or causes, what can possibly account for its universality?

That is, could any other quantitative empirical law that is specific to economics boast of the same degree of "robustness," in terms of accuracy and universality? What *New Methods* argued by careful mathematical analysis is that the answer is *No*. For very "dirty" data, Pareto's power law is *the only form* that a probability distribution could consistently take. This property does not concern the prefactor $F$ but the exponent $\alpha$. It holds for all $\alpha$ from 0 to infinity—though it is mostly dangerous when $\alpha$ is small and the tail is long. Consider first different components of income separately, then their arbitrarily weighted additive aggregates or mixtures. I showed that only the power law is effectively "invariant" or "robust"



DOI 10.1002/cplx

under these transformations and analogous ones that reflect diverse existing conditions of observation. For its proof I must refer to *New Methods*. When first reported, it created enough surprise to be featured as an exercise in a celebrated textbook by Feller [31].

It follows that in seeking a distribution for income, Pareto did not really deal with a long list of possible alternatives. In particular, the widely promoted log-normal distribution fails to qualify. It is robust under multiplication (an operation that has no concrete meaning) but it is not at all invariant under aggregation.

Therefore, *either* all variants of income follow a power law *or* there is a total mess that varies in time and space. The same holds for all too many phenomena in economics.

Invariances are the most fundamental tools in both mathematics and physics, and much of my life work has consisted in extending their scope. The fact that the first great empirical finding in economics involves an invariant law completely ceased to be surprising—though this property in no way suffices to explain it.

A closely related aspect of economic uncertainty does not concern a quantity's distribution but the functional relation between different quantities. This issue was tackled in *Vérification des Hypothèses* [29].

In the hundred years since Pareto, no other empirical law of comparable robustness has been discovered. Could a kind of "anthropic principle" be at play—meaning that no distribution other than a power law will ever be firmly established in economics? Is this not ominous? The issue must not be prejudged but faced—not as an afterthought but early in basic economic arguments.

As background, think of any early result in physics, for example Ohm's law that links the intensity of electric current in a wire to the difference of potential between its endpoints. When it was discovered around 1820, the errors of measurement were dreadful, yet, the Ohm law was uncontested. When, much later, professional statistics arose, Ohm's errors must have been found to be more or less Gaussian and his law was surely found to hold at a high level of significance, not to be challenged.

Things are harder in economics. All too often, tests that assume Gaussianity are applied without checking whether the error terms are really Gaussian. If so, relations inferred from data may seem to hold for a while but after a few years prove to be completely wrong. What if the errors' distribution follows a power law? Formally, the expressions that concern statistical estimation have been extended by several authors. I am sure they will be improved upon. But even assuming that a relation holds in terms of populations, one should perhaps expect that, in terms of samples, it will fail "every so often." This issue deserves fresh thought.

The above concept of robustness under change of definition is one that I adapted from two sources far removed from economics. One is the theory of turbulence, with the works of the mathematician Kolmogorov and the physicist Onsager. The other is in pure mathematics with the works of Augustin Cauchy in 1853, Lévy [32], and others. Later these methods became widely used in physics, for example in the theory of critical phenomena; they came somehow to be dubbed "renormalization."

## 3. ZIPF'S "RANK-SIZE" RESTATEMENT OF PARETO'S LAW

Pareto's law has been shown to apply to many quantities other than personal incomes, a particularly persistent collector and seeker having been Zipf [see Figures in Ref. 30]. For some reason, perhaps, the fact that he was particularly concerned with the distribution of word frequencies, he chose to proceed in terms of a nonstandard restatement called "rank-size law." It also takes the form of a power law but has created much confusion and therefore deserves a brief digression.

Let $U$ denote a quantity whose value is random, for example, a person's income, a man's height, or an oil reservoir's capacity. Probabilists use the corresponding lower case letter, $u$, to denote the sample value, as measured in dollars, inches, or barrels. Consider a sample of $N$ values, and denote by the number $\#\{U \geq u\}$ the number of cases where $U \geq u$. When this number follows the hyperbolic formula

$$\#\{U \geq u\} = NFu^{-\alpha}, \qquad (1)$$

the behavior of $U$ is call Pareto-like.

Assuming $U$ to be random and assimilating the relative number of cases – $\frac{\#\{U \geq u\}}{N}$ – to a probability, (1) yields

$$P(u) = \Pr\{U \geq u\} = \text{the probability that } U \geq u = Fu^{-\alpha}. \qquad (2)$$

Take "items," they may be the above examples, or words, or city or firm sizes, and order them by decreasing frequency, population, or income. This ordering attributes to each item a rank $r$. For example, the biggest city has rank $r = 1$, the second biggest has rank $r = 2$, and so on.

So far so good, but one is free to plot (1) on transparent paper and then turn the sheet back up. In the resulting inverted coordinates, (1) becomes

$$u = r^{-1/\alpha}(NF)^{1/\alpha}. \qquad (3)$$

Equation (3) expresses the size of the city of rank $r$ as a function of the rank in this ordering. Zipf did not view (3) as simply the result of inverted probabilistic coordinates, but as a powerful free-standing way of quantifying complicated reality. My account of (3) will be described in Section 6.

*Nota.* Perhaps influenced by Pareto, Zipf chose to assume *both* a universal exponent, namely, $\alpha = 1$, *and* a universal prefactor, namely, $(NF)^{1/\alpha} = 1/10$. These estimates were soon discarded, in particular, because of their strong implication that the total number of distinct words is finite and universal. "Common sense" and the data, and also my model suggest otherwise.


DOI 10.1002/cplx

## 4. A KEY PROBLEM—INEQUALITY IN PARTITION—AND A KEY DISTINCTION—BETWEEN DIFFERENT "STATES" OF RANDOMNESS: "MILD" AS EXEMPLIFIED BY THE GAUSSIAN AND "WILD" AS EXEMPLIFIED BY THE POWER LAW

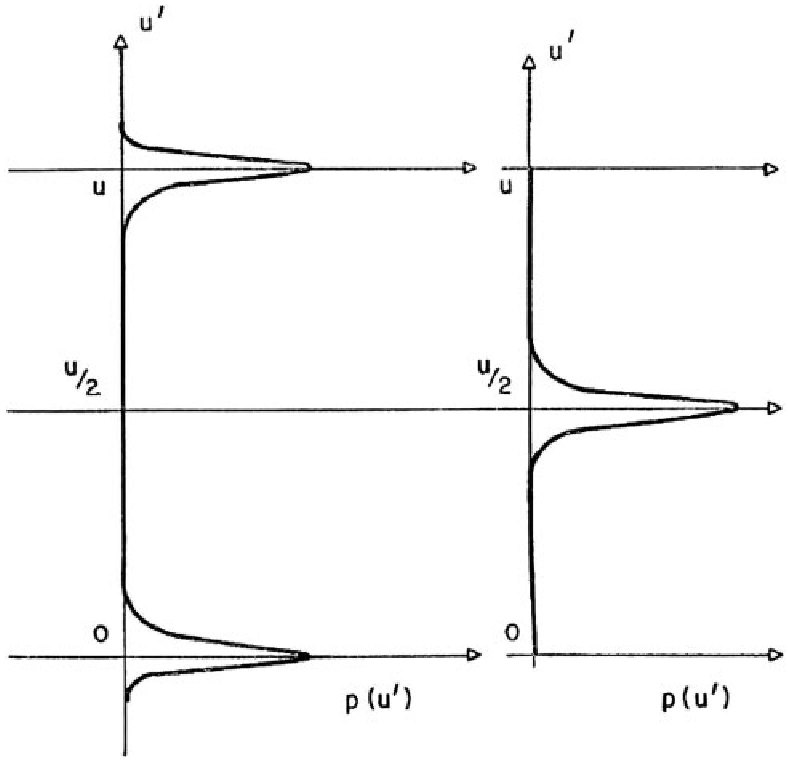
Fundamental to my view of statistics—in economics and many other fields—is a sharp distinction between different "states" of randomness.

Mild randomness is exemplified by the Gaussian distribution—one which is commonly called "normal," but in fact deserves less and less to be considered as such. A Gaussian's key feature is a graph of $\log[p(u)]$ that is "convex," like an extended sign ∩.

Wild randomness is exemplified by the power law. A typical key feature of the power law distribution is a graph of $\log[p(u)]$ that is "concave," like a portion of the sign ∪.

In early terms I had used in 1964, the Gaussian exemplifies a "first stage of indeterminism," one that promises a high precision but is in many cases unrealistic and has been deeply misleading. The power law exemplifies a "second stage of indeterminism," in which a lesser precision is involved, but can be delivered. Let me now show that the Gaussian characterizes near-equality between the samples and absence of any substantial concentration, whereas the power law characterizes clear-cut inequality and concentration. Those two solutions of what I called the problem of partition are illustrated by Figure 1 excerpted from my first article [3] on the distribution of income.

**FIGURE 1**




Conditional probability density of the addend $U'$ when the value $u$ of $U' + U''$ is large, for two distributions of $U'$ and $U''$: a bell with an asymptotically scaling tail (left) and the Gaussian (right). For the exponential, $p(u')$ is a constant.

To explain Figure 1, take a homogenous population of adult males and focus on some characteristic, such as height or last year's income, that is variable and can be visualized as random. The common notation of probability theory denotes a random variable by an uppercase letter, say, $U$, and denotes two independent values of $U$, call them $U'$ and $U''$ – by the corresponding lowercase letters $u'$ and $u''$. Finally suppose that you know the sum $u = u' + u''$, for example, the sum of the two males' heights or their last year's incomes.

Now face a question that may at first seem contrived but in fact is deep and far-reaching. Knowing the sum $u$ , what is the form of the so-called conditional distribution of the parts, $U'$ or $U''$? By construction, it is always symmetric relative to $u/2$. But otherwise, Figure 1 shows that changing the distribution of $U$ can affect the conditional distribution of the parts drastically.

In the case of heights, the distribution of $U$ is approximately Gaussian and in this approximation the conditional distribution of "$U'$ knowing $u$" is Gaussian as well, with a thinner bell around $u/2$.

*Proof.* If the *a priori* distribution of $U'$ is Gaussian with mean $M$ and variance $\sigma^2$, then the *a priori* distribution of $U$ is Gaussian with mean $2M$ and variance $2\sigma^2$. The conditional distribution of $u'$ given $u$ is then given by

$$p(u'|u) = \frac{\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(u'-M)^2}{2\sigma^2}\right]\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(u-u'-M)^2}{2\sigma^2}\right]}{2^{-1/2}\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(u-2M)^2}{4\sigma^2}\right]}$$

$$= 2^{1/2}\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(u'-u/2)^2}{\sigma^2}\right]$$

Thus, the conditional probability is itself Gaussian with expectation $u/2$, which depends on $M$, and variance $\sigma^2/2$, which *does not depend on $M$.* ■

But the Paretian last year's income yields an entirely different result. If the sum of two incomes is two million dollars, the conditional value of $U'$ is highly



DOI 10.1002/cplx

bimodal. In the figure's left half, corresponding to the smaller of the two incomes, the conditional distribution lies on a curve very close to the distribution of each income taken separately. To the contrary, the larger of the two incomes is very large and close to two million.

Indeed, approximate the distribution of income by $P(u) = \text{Prob}(U' > u) = u^{-\alpha}$ restricted to $u \geq 1$, hence approximate the density by $2\alpha u^{-\alpha-1}$ . It is easy to check that – for large $u$ – the density of $U = U' + U''$ is $\sim 2\alpha u^{-\alpha-1}$. It follows that

$$p(u'|u) = \alpha u'^{-\alpha-1}\alpha(u-u')^{-\alpha-1}/2\alpha u^{-\alpha-1}.$$

At the endpoints $u' = 1$ and $u' = u - 1$, this conditional density is about $\alpha/2$. At the midpoint $u' = u/2$, the density is proportional to $u^{-\alpha-1}$, which is enormously smaller than $\alpha/2$ and tends to 0 as $u$ increases. The same holds, to take an example, for the probability in the central interval that can be defined, for an arbitrary fixed ratio $q$, as ranging from $qu$ to $(1-q)u$.

The very unequal split is a key characteristic of income and finds counterparts in most other phenomena of economics. The opposite behavior, namely, near equal split, is very common in the established physical sciences, but highly atypical in economics.

## 5. SOME NOTORIOUS FEATURES OF THE POWER LAW: LACK OF A MEANINGFUL "TYPICAL VALUE," AND ITS EFFECT ON "LYING WITH STATISTICS"

Among the distributions of probability, the Gaussian is a model child: it is symmetric so that its expectation, median, and most probable values all coincide, and its "standard" deviation around the mean is meaningful and closely related to, say, the interquartile. Furthermore, knowing the sample mean and variance suffices for all purposes, where the term "sufficient" can either take the everyday meaning or a technical one provided by statistics.

To the contrary, a key fact about the income distribution is that every one of the above-listed attributes is contradicted. The most probable value lies in the low-income skew bell. The sample mean and median differ considerably and both lie between that bell and the tail, somewhere in a range that is typical of nothing and does not draw attention to itself. The mean–median discrepancy is one reason behind the notorious fact that to lie with statistics takes little effort.

Other typical specifications of nice distributions include moments. This is the point where Pareto's law and more generally the power law distribution stand apart and encounter especially ferocious *a priori* opposition. The exponent $\alpha$ is "critical," in the sense that all population moments of order above $\alpha$ are infinite; they are said "not to exist."

These negative properties are widely viewed as "anomalous" and "abnormal." In physics, many moments *must* be finite and subjected to conservation laws. The tail's shortness is a key reason why physics could begin with "macro" theories stripped to typical values and only later add "micro" refinements linked to sample variability. But economics proper has no conservation law, so that divergent moments are not excluded. This makes it necessary to rethink the relations between micro and macro theories.

The nonavailability of typical values in economics makes this split raise serious difficulties. The most tempting policy is to paper the difficulties over by truncation, as will be exemplified in Sections 6 and 7; yet, the search for typical values is a continuing intellectual challenge. To the contrary, I have always thought that hiding those challenges also hides some features fundamental to economics, while facing them raises many new and hard problems but opens many windows and brings much knowledge.

Gaussian-based models are part of a belief that economics can aim at an ultimate level of precision comparable to that of standard physics. But this belief has been extremely unhelpful. Facing those challenges promises predictions that are considerably less precise but also more realistic. Incidentally, analogous thoughts also apply in many corners of physics that have also developed slowly. As already sketched, the contrast is not merely a matter of degree but of my long held and often expressed opinion [33] that "indeterminism in science must move on to a second stage," or that randomness can be either in a mild or a wild state.

## 6. THE TEMPTING BUT UNCALLED-FOR RESORTS TO CUTOFF OR TRUNCATION

Among Pareto's income data sets, the Duchy of Oldenburg stood out as an exception for lack of very large values of income. This income distribution motivated Pareto to multiply the power law $u^{-\alpha}$ by a factor $\exp(-\delta u)$ that amounts to a gradual "truncation" of the tail. Statisticians liked this truncation, but not for Pareto's empirical reason, rather for the mathematical reason that all moments are automatically made finite. Some recent econophysicists also began by reviving the exponential factor, but later withdrew it. Other writers promoted at one time some alternative gradual truncations or even an abrupt cutoff declaring that values of $u > u_{\text{cutoff}}$ never occur. Indeed, truncation makes all moments finite but it also greatly affects their values in a way that is usually unnatural.

Incidentally, physics had long been very familiar with power laws, beginning with Newton's universal power law of attraction. But many physicists viewed the resulting divergences as unacceptable and exponential truncations were often proposed. In addition, the notion of quantum of energy was originally introduced by Planck to justify a contrived truncation meant to avoid the "ultraviolet catastrophe." The contrivance worked in that instance, but need not work in all cases.

*An exception: a cutoff is justified in the case of word frequencies.* The power law distribution of word frequencies has already been mentioned and will be mentioned again. For it, a cutoff is unavoidable and can take a form I


DOI 10.1002/cplx

devised very early thanks to a convergence of experiment and theory.

Indeed, a glance at the log–log graphs shows that the generalized Zipf's formula (with $\alpha$ not necessarily 1) is grossly inaccurate for small $r$. Indeed, if one uses the parameters $F$ and $\alpha$ estimated from middle or large values of $r$, interpolating for all $r$, suggests a total "probability" $\sum Fr^{-1/\alpha}$ that is well above 1. However, this difficulty vanishes and the fit is enormously improved if Zipf's formula is replaced by the Zipf-Mandelbrot expression $F(r+V)^{-1/\alpha}$. That is precisely the formula my theory reaches by mathematical necessity: a suitable $V$—not an additional parameter but a function of $\alpha$ and $F$—insures that the sum $\Sigma F(r+V)^{-1/\alpha}$ is equal to 1. That is, in data plotted on doubly logarithmic coordinates, the quantity $V$ takes care—in addition to the cutoff at $r = 1$—of a bend or crossover present for small $r$.

*Cases beyond the exceptional example of word frequencies.* Other contexts bring different temptations for postulating a cutoff. Often, it turns out to be a statistical artifact that careful thought eliminates.

My first encounter with this possibility occurred in 1962 within physics, in the study of the intervals between successive errors in data transmission [34]. Early plots of the evidence showed a power law range extending over several orders of magnitude—an exceptionally wide span—followed by a rapid decay. Many reasons tempted my coauthor to suggest truncation, but I convinced him to first look closer. What we found is that some samples had been discarded by our provider of data as being obviously anomalous: "outliers" of no statistical significance. I insisted on getting hold of those data and adding them to our samples. This attention to detail was rewarded: it eliminated the need for a cutoff. Recent users of power laws in the earth sciences encountered the same issue and reached the same conclusion.

To elaborate, consider a large number of samples drawn from a power-law distribution. Using ranked data, well-known theorems establish that sample variability is small for large values of the rank $r$, but for small values of $r$, especially for the extreme value $r = 1$, it is very large.

As a result, the value observed for $r = 1$ is expected to be way out in an occasional sample, giving the impression it is an "outlier" that can be thrown out.

But for the great majority of samples, the value for $r = 1$ will be close to the values for larger $r$. Those cases create an an apparent crossover that seems to demand a truncation or cutoff of power law behavior. In increasingly large samples, however, this would-be outlier proves to be of a size just right to extend the scaling range. To make sure that a bend is intrinsic and defines two separate scaling ranges, one must observe that—when two very different sample sizes are compared—the bend does not change.

## 7. A CASE AGAINST THE ULTIMATE CROSSOVER: LOGNORMAL DISTRIBUTION

The power law is only concerned with the high incomes tail, but claims to represent it in precise, enlightening, and useful fashion. However, for over a hundred years since its discovery, the Pareto law for income received less acclaim than disdain or dismissal. There was a broad consensus that the data's bell and tails can be fitted just as well, or even better, by the lognormal distribution.

Let $V = \log \Lambda$ be Gaussian, of probability density

$$\frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(\nu-\mu)^2}{2\sigma^2}\right].$$

Then the probability density of the lognormal $\Lambda = \exp V$ is

$$p(\lambda) = \frac{1}{\sigma\sqrt{2\pi}}\frac{1}{\lambda} \exp\left[-\frac{(\log\lambda-\mu)^2}{2\sigma^2}\right].$$

In some respects the lognormal distribution is of great simplicity. This is one reason why, next to the Gaussian, it is widely viewed as the practical statistician's best friend; it continually competes with the power law distributions in the many fields of science where skew long-tailed histograms are a fact of life and concentration ratios are not small. Competition persists because in the midrange the difference between the alternatives is far less than vastly different analytic forms would suggest.

In a way, the lognormal is the ultimate truncation of the power law. It even keeps being suggested as substitute for Zipf's law of word frequencies. The professional statisticians' conventional view is exemplified on pp 101–210 of [35]. We read that "A number of distributions are given by Zipf, who uses a mathematical description of his own manufacture [sic; was Pareto forgotten?]. . .; in fact, however, it is likely that many of these distributions can be regarded as lognormal, or truncated lognormal, with more prosaic foundations in normal probability theory." Those prosaic foundations reside indeed in "normal" probability theory combined with a "principle of proportional effort."

Aitchinson and Brown were apparently not aware that (see Section 8) Pareto and Zipf had also been reduced to probability theory. Nevertheless, Zipf became an attractive magnet for non-professional dabblers of all kinds, but also a repulsive magnet for the overwhelming majority of professional students of randomness.

As long as the quarrel between power law and lognormal remained a matter of mechanical statistical testing, it became tiresome and seemed to persist without end. This issue deserves a comment without waiting. Statistical tests being by no means intrinsic, different goals call for different tests. The present complication is not only an issue for philosophical discussions by statisticians but something very practical. On close look, lognormality is not what is taken for granted. Each of those drawbacks suffices to make the lognormal misleading and even dangerous.

In particular, sums of lognormals are not lognormal and their analytic expressions are unmanageable. Lognormality is invariant under multiplication, but this is not a useful property because dollars and firm sizes add and subtract but



DOI 10.1002/cplx

do not multiply. This is a severe handicap since each scientific or engineering problem involves many fluctuating quantities, linked by a host of necessary relations. For each quantity, given a long all-purpose list of possibilities, pure curve-fitting will look for the best fit. But compare the best fit with 1-day and 2-day price changes.

The distribution of the sum of 1-day lognormal fits is not lognormal; it is neither analytically manageable nor identical to the distribution of a 2-day fit. More generally, a good theory should satisfy the requirement that the separately best fitting expressions should be linked by the relations the data must satisfy. This is very constraining. The jointly Gaussian case is the classical exception. Another exception is found in my "PLM" model, to be touched on in Section 9.

*Population moments of the lognormal* $\Lambda$. They pose delicate problems – being finite, *but far from robust with respect to deviations of* $\log \Lambda$ *from absolute lognormality*. Being approximately Gaussian is usually good enough, but from this viewpoint the population moments of $\exp X$ are unusual. The known values of $E\Lambda^q$ are simple to evaluate for a lognormal, but extremely sensitive to seemingly insignificant deviations of $\Lambda$ from lognormality. As lognormality can never be verified with absolute precision, the moments of $E\Lambda^q$ are effectively arbitrary. Let me elaborate.

*Sequential sample moments of the lognormal behave extremely erratically; they do approach the population moments, but only in an irrelevant long-run.* John Maynard Keynes observed that in the long-run we shall be all dead. So asymptotic results are fine, but not sufficient. In the middle-run, the scatter of the sequential sample moments of the lognormal is large, ruled by impossibly complicated formulas, and effectively useless. To appreciate that scatter, a good approach is to graph the data on doubly logarithmic coordinates, then focus on the range where they can be approximated by a power law. To extrapolate from that range makes a power law fit useful even for lognormal data!

*It is widely assumed that the lognormal has long been thoroughly "explained" by a straightforward random "proportional effect" argument, in fact, in such approximations the Gaussian is only an approximation valid in a central "bell"—not in the large u tails. Therefore the argument's conclusions concerning* $\Lambda$ *cannot be trusted.* In most scientific problems, the fact that the central limit theorem only concerns a central bell makes little difference. In real problems, however, the number of conceivable multiplicative terms of proportionate effect is not only finite (it always is in science); it is small. Therefore, the Gaussian involved in the limit theorem is at best a distant asymptotic approximation to a preasymptotic reality which (as already mentioned) may be very different.

*Under the lognormal assumption, the basic phenomenon of industrial concentration is downgraded as a transient that can occur in a small sample, but vanishes asymptotically.* In an industry including $N$ firms of lognormally distributed size, the relative size of the largest depends on $N$. It is only asymptotically that this $N$-dependence vanishes. For small $N$, it may be sizeable. Suggesting that concentration is a transient phenomenon that would introduce endless complication and is supported by no compelling evidence.

In essence: the view of the lognormal as the statistician's best friend is by no means obvious. That distribution is to be handled with extreme caution and in general is to be avoided.

*Slow randomness:* Between the mild and wild "states of randomness," I found it useful to introduce an intermediate state of "slow randomness." The lognormal provides an excellent illustration of this state and its pitfalls. It is beloved because it "passes" as mild: moments are easy to calculate and it is easy to take for granted that they play the same role as for the Gaussian. *But they do not.* They hide the difficulties due to skewness and long-tailedness behind limits that are, at best, overly sensitive and overly slowly attained.

## 8. A BROAD CLASS OF TRADITIONAL WOULD-BE "EXPLANATIONS" OF PARETO-ZIPF THAT INVOKE LOG(INCOME) OR THE LIKE: ONE SUCCESS AND MANY DIFFICULTIES

Pareto's discovery was puzzling and efforts to explain the power law began early and continue ceaselessly to this day. On casual inspection, those old or recent efforts differ from one another. But, unless there is something very new that I missed, all involve the same key ingredient: compensation between two exponentials.

In rank-size terms, each involves an auxiliary quantity $k$ and two constants $A > 0$ and $B > 0$ such that rank $\approx A^k$ and probability $\approx B^{-k}$, hence, writing $\alpha = -\log A / \log B$ yields $u \approx r^{-1/\alpha}$.

In income terms, one must provide the expression $V = \log(\text{income})$ with a concrete interpretation. If so, Pareto's power law $u^{-1/\alpha}$ transforms into $\exp(-\alpha \log u) = \exp(-\alpha v)$. The exponential distribution plays a central and well-understood role in physics, especially statistical thermodynamics. In fact, it is encountered all over the sciences and is endowed with a Horn of Plenty of alternative explanations that seem formally very easy to translate into new terms. Those explanations may differ in form but are logically absolutely equivalent.

Examples include the dynamical explanation of the so-called "barometric distribution" that rules the density of air under the combined efforts of gravity and heat motion. This distribution is the steady state in a process that combines diffusion with a reflecting lower barrier.

It is not a surprise that a number of models of scaling are more or less obvious and/or conscious economic translations of various physical models. But this seemingly easy task includes a difficult step: to justify the transformation $V = \log U$.


DOI 10.1002/cplx

*The easy case of word frequencies.* By good fortune, I first met power law distributions in 1951 in the context of word frequencies. This is a rare case where $V$ has a clear intrinsic meaning. Indeed, information theory interprets $\log U = \log(\text{a word's frequency})$ as roughly the number of letters in an optimally coded word. The remaining argument is so straightforward that a sketch of a simple variant can fit here.

Take an alphabet of $M + 1$ letters, $L_m$. Have "typing monkeys" use this alphabet to produce a random text in which (a) the improper letter $L_0$ is a "space" used with the probability $p_0$, and (b) each of the $M$ proper letters is used with the probability $(1 - p_0)/M$. A word made of $k$ proper letters followed by a space will have the probability $p_0[(1 - p_0)/M]^k \sim B^{-k}$, by definition of $B$. Such a word's rank $r$ is $\sim M^k$. Therefore $k = \log r/\log M$, and the word's probability takes the form $\Pr \approx p_0 \exp(-\log r \log B/\log M)$. That is, $\Pr \approx r^{-1/\alpha}$ with

$$\frac{1}{\alpha} = \frac{\log B}{\log M} = \frac{-\log(1 - p_0) + \log M}{\log M} = 1 + |\log_M(1 - p_0)| > 1. \quad (4)$$

*There is nothing more to Zipf's law for words.* For example, I showed that this argument extends to the case when the letters are not independent and their sequence is Markovian. In other generalizations the same result is obtained asymptotically.

The compensation between the two exponentials can be phrased in several additional ways. There is a "thermodynamical" or "information-theoretical" restatement that reduces the exponentiality of the distribution of $\log u$ to entropy maximization—entropy being often reinterpreted as a Shannon information. This interpretation brings nothing new, but appears learned.

In 1953, shortly after I explained Zipf's law for word frequencies, some linguist friends believed that this law might impact syntax or perhaps semantics. It does not but this investigation somehow started me on a path that led to fractals.

Every generalization involves a complication and there are also cases where Zipf's law remains altogether unexplained.

*A prelude to arguments claiming to explain the exponential distribution of $\log U$ as being a barometric distribution. A surrogate for diffusion based on a biased random walk in integer time $t$.* This surrogate restricts $\log_e U$ to be the form $kc$, where $k$ is an integer and $c > 0$. Between times $t$ and $t + 1$, it allows the following possibilities:

1. $\log_e U$ can increase by $c$, the probability being $p$.
2. $\log_e U$ can decrease by $c$, the probability being $1 - p$.
3. $p < 1/2$, When used to explain the barometric distribution in the atmosphere, this inequality expresses the "force of gravity."
4. There is a reflecting barrier that can indifferently be interpreted in either of two principal ways. The firms that go below $\log_e U = \log_e \tilde{u} - (1/2)c$ are given a new chance to start in life at the level $\tilde{u}$, or become lost but replaced by a steady influx of new firms starting at the level $\tilde{u}$.

Condition 4 is essential. Otherwise, conditions 1–3 make firm sizes decrease on the average. Hence, to insure that the number of firms remains time-invariant above a lowest value $\tilde{u}$, one needs a counterpart of the closed bottom in the barometric distribution of physics.

*A necessary condition for equilibrium.* For a distribution of $\log U$ to be invariant under the above transformations, it is *necessary* that the expected number of firms growing from size $e^{kc}$ to size $e^{(k+1)c}$ be equal to the expected number of firms declining from size $e^{(k+1)c}$ to size $e^{kc}$. This can be written

$$\frac{\Pr\{\log_e U > (k+1)c\} - \Pr\{\log_e U > kc\}}{\Pr\{\log_e U > kc\} - \Pr\{\log_e U > (k-1)c\}} = \frac{p}{1-p}.$$

A steady-state solution in which $\Pr(\log U > v) = \exp(-\alpha v)$ requires $\exp(-\alpha c) = p/(1 - p)$. The condition $\alpha > 0$ requires the following system of equations:

$$\begin{aligned}
&P(k,t) \\
&\quad = \Pr\{\log_e U = kc \text{ at time } t\} \\
&P(k,t+1) \\
&\quad = pP(k-1,t) + (1-p)P(k+1,t) \\
&\quad \text{if } k > k_0 \log(\tilde{u}/c) \\
&P(\log_e(\tilde{u}/c), t+1) \\
&\quad = (1-p)P(\log_e(\tilde{u}/c), t) + (1-p) \\
&\quad P(1 + \log_e(\tilde{u}/c), t).
\end{aligned}$$

These equations are of the second order, that is, "diffusive." Therefore, they have a steady-state limit function $P(k, t)$, independently of the initial conditions at a preassigned starting time $\tilde{t}$. That limit is exponential. (*Proof*: Under the steady-state condition $P(k, t + 1) = P(k, t)$, the second equation yields $pP(k, t + 1) = (1 - p)P(1 + k_0, t)$. Then the first equation gives the same identity by induction on $k_0 + 1, k_0 + 2$, and so on.) Therefore, conditions 1.,2.,3. and 4. provide a possible generating model of the exponential distribution for $\log U$.

*Reminder of reasons why diffusion is a reasonable idea in physics.* One reason is that, while the limit takes a large number of exchanges of energy to be attained, each interaction is extremely fast and the overall duration in (say) seconds may be small enough to make an asymptotic model worth examining closely.

A second reason is that the energy of even the most energetic gas molecule is negligible with respect to a total energy. Therefore, the details of its distribution hardly matter. One can rely on something called a "mean field argument" that trusts that all kinds of things can happen to one molecule, while elsewhere hardly anything changes. Even the most energetic molecule is only affected by the "mean field" of all the others.

This simplicity helped the theory of gases to develop and to be applied. It makes no difference whether the total "energy" $\sum V$ in a gas is fixed (as it is in a "microcanonical" system) or allowed to



DOI 10.1002/cplx

fluctuate (as it is in a "canonical" system.) This is a reason why, in physics, the diffusion model's conclusions do not contradict its premises.

*Reasons why diffusion of* log $U$ *seems* not to be *a reasonable idea in economics, at least a priori.* A first reason is that the number of steps needed for a limit theorem to apply may well be too large for the limit to be relevant. A second reason is that the transformation $V = \log U$ might seem innocuous, but is not. It introduces major changes, especially when $\alpha < 2$ and $U$ is what I called "wildly random."

For example, a country's largest city may "typically" include 15% of the total population, but the sample variability is very high. Examine the diffusion model's key assumption, that the largest firm or city can grow or wane without influencing a whole industry or country. Ex-post, if the diffusion argument could somehow be inverted, it would be a prediction about observable facts—interesting but arguably incorrect. Ex-ante, this presumed property is surely no more obvious than scaling itself, hence cannot be safely inserted in an explanatory model of scaling.

The diffusion of log $U$ yields baffling economic predictions. Additionally, the argument ceases to be grounded in thermodynamics, because the latter does not handle situations where canonical and microcanonical models do not coincide.

It is true that neither the success nor the failure of a model in physics can guarantee its success or failure in economics.

All told, the models of $U$ based on the diffusion of log $U$ leave an embarrassment of questions. The *utter pessimist* will dismiss flawed models as not worth further discussion. The *moderate pessimist* will not take the diffusion of log $U$ seriously but will play with it. The *moderate optimist* will not want to look too closely at a gift horse; even in physics, it is common that stochastic models yield results that contradict the reality behind their assumptions. In any event, I view the moderately optimistic position as very hard to adopt. The widespread very optimistic and unquestioning view of diffusion as sufficing to explain power laws is altogether indefensible.

### 8.1. Conclusion

The concentration characteristic of wild randomness has very concrete consequences. The difficulties associated with the diffusion of log $U$ have helped me to adopt the policy of viewing scaling as a postulate that brings economies of thought, good fit and a useful basis for practical work. It must be thoroughly understood, and its consequences explored, without waiting for a definite explanation.

## 9. MORE PRECISE MODELING OF PRICE VARIATION AND THE ISSUE OF WHETHER OR NOT VARIANCE IS FINITE, I.E., $\alpha > 2$: THE 1963 PLM PROCESS, THE 1965 HHM PROCESS, AND THE 1972 MULTIFRACTALITY

*Certain Prices* appeared simultaneously with *New Methods*, and in many ways it is a far more important contribution to economics. It is also far better known—though this fact is often hidden by imaginative alternate terms.

Recall that around 1960 the old Ph.D. thesis of Louis Bachelier, which dated to 1900, suddenly became very popular again. It postulated that price follows what later became known as a Brownian motion: price changes are Gaussian and their values over nonoverlapping time spans are independent.

This thesis was admirable and of the greatest possible historical interest, but in economic data all its assumptions are contradicted and its principal consequences are falsified. A counter theory was needed and the most important discoveries reported in *Certain Prices* were that the distribution of price changes is long-tailed and that price variation is discontinuous. The relations between some old articles raise issues that remain important and deserve comment.

As already mentioned, *New Methods* is founded on a principle of invariance—the scaling principle—which it applies mostly to the asymptotic range of large values of $u$. To the contrary, *Certain Prices* and some later article of mine postulate scaling for all values of $u$. A key difference is that asymptotic scaling applied to independent variables allows $\alpha$ to range from 0 to $\infty$ but uniform scaling restricts $\alpha$ to stay between 0 and 2.

For example, take the sum of $N$ variables following the power law distribution with $\alpha = 2.9$. That exponent being greater than 2.0 implies finite variance. Hence the central limit theorem kicks in asymptotically, and a Gaussian bell arises *in the limit*. However, it is true for all values of $\alpha$ that power law tails persist and vanish only in a practically irrelevant asymptotic regime. For finite $N$ they only help in the central bell. At that time, such refined arguments were new (and—as already noted—W. Feller quoted this one in an exercise of his celebrated textbook) and used to belong to esoterica. The need to be concerned with esoteric minutia in applied work has been a token of my work in all fields.

The fact that I began by studying cotton was largely accidental but extraordinarily fortunate because in that case $\alpha$ is about 1.7 , that is, clearly less than 2, which is the critical value for the finiteness of the variance. As a result, the fact I began with cotton sheltered me from being forced to face long tails and long dependencies at the same time. But the sign of $\alpha - 2$ soon became disturbing, as the value of $\alpha$ that I observed for wheat prices was uncomfortably close to 2. So it was for the security prices to which E. Fama extended my work.

An old rule of warfare is: if you can, face your foes one by one. In effect—returning to Bachelier's Brownian motion to face its flaws—chance made me obediently follow this rule. It has often proven very effective in the sciences, but requires exquisite care. In the wake of *Certain Prices*, many authors itemized the flaws of the Brownian and tackled each separately. This yielded a pile of incoherent (and often mutually contradictory) fixes. My count of the ad-hoc parameters needed in a published article


DOI 10.1002/cplx

halted when it reached 37. Those parameters' effects could not be separated and none could be estimated.

Instead, my work in the 60s adopted the opposite extreme. In 1963, I tackled the long tailedness of the margined distribution of price changes, which I memorably termed the "Noah effect." Shortly later, in 1965 [36, 37], I tackled a form of dependence, which I called the "Joseph Effect"—by introducing a then-shocking notion that dependence can be infinitely long. A dream I had held throughout and often stated was finally fulfilled when I saw that the two effects can be solved simultaneously by injecting multifractals.

Moreover, the same tool was applied throughout: some form of scaling that I borrowed more or less unconsciously from the study of turbulence—well before it flourished in the study of critical phenomena.

## 10. HALF-WAY THROUGH ITS HUNDRED YEARS LONG HISTORY THE POWER LAW—SCALING—CHANGED FROM AN IDLE CURIO TO AN IMPORTANT SPUR TO ECONOMICS

Arguably, my 1951 explanation of Zipf's law and its sequel have naturally split the history of Pareto's law into two periods of roughly equal length.

During its first fifty-odd years, as noted, it was reasonably widely known but not really believed and not important. A mere curio examined with unsophisticated tools, it had little impact on either the theory or the practice of economics. The early literature was impressive in quantity (especially in French and Italian), but limited to arguing against the lognormal alternative, finding new examples (this being the key contribution of Zipf) and trying to "explain" the behavior of log(income) by some random multiplicative argument.

I took a different tack, injected altogether more sophisticated tools and methods, and all-too often encountered controversy. Directly or not, a high proportion of my research was not devoted to the causes of the power law, but to its surprisingly extensive consequences. They have by now grown into something that promises to contribute to a much needed new foundation for those essential parts of economics that involve inequality.

So bold an assertion demands immediate elaboration and—most unfortunately—first demands a one paragraph long detour through the swamp of methodology. As soon as some important and pressing questions are identified, there is a strong temptation to attack them without lengthy and seemingly irrelevant preparation. My reading of the history of science suggests that this strategy has often proved ill-conceived. It amounts—to use an alpinist's words—to attempt to learn mountain climbing by attacking the most visible and challenging peaks. Clearly, this would be an unpromising and dangerous idea. It is wiser to first gain skills and confidence on far less glamorous but easier slopes.

True, this view has an uncomfortable resemblance with the joke about "looking for a lost key under the lamp post rather than in the dark corner." However, in the case of economics, the direct search for a well understood and satisfactory framework has surely not been an unmitigated success and a new framework is needed.

Commenting on *Certain Prices*, Paul Cootner blasted me *as a revolutionary* [who] *like Prime Minister Winston Churchill before him, promises us not utopia but blood, sweat, toil and tears. If* [he] *is right, almost all our statistical tools are obsolete... Almost without exception, past econometric work is meaningless* [38]. This verbal blast is sufficient to explain why the early work revisited in this paper was not immediately accepted. In the context of wild randomness, the standard statistical tools for mild randomness have continued to receive lip service. But in fact, the Gaussian (or lognormal) was overloaded with fixes, each meant to account for a feature of my model, such as discontinuity and variable variance.

Since the 60s, however, the economists' confidence in their work has definitely not increased. The standard statistical tools had promised a good deal but delivered little. The "new methods" I have been advocating promise less precision but offer a launching pad for a much needed fresh attempt at a quantitative understanding of economics. They have earned the right to be put to continuing test.


## REFERENCES

1. Mandelbrot, B. New methods in statistical economics. J Polit Econ 1963, 71(Octob.), 421–440.
2. Mandelbrot, B. The variation of certain speculative prices. J Business Univ Chicago 1963, 36(October), 394–419.
3. Mandelbrot, B. The Pareto-Lévy law and the distribution of income. Int Econ Rev 1960, 1(May), 79–106.
4. Mandelbrot, B. Stable Paretian random functions and the multiplicative variations of income. Econometrica 1961, 29, 517–543.
5. Mandelbrot, B. Forecasts of future prices, unbiased markets and "martingale models". J Business Univ Chicago 1966, 39, 242–255.
6. Mandelbrot, B. Long-run linearity, locally Gaussian processes, H-spectra and infinite variances. Int Econ Rev 1969, 10, 82–111.
7. Mandelbrot, B.; van Ness, J.W. Fractional Brownian motions, fractional noises and applications. SIAM Rev 1968, 10, 422–437.
8. Mandelbrot, B. When can price be arbitraged efficiently? a limit to the validity of the random-walk and martingale models. Rev Econ Stat 1971, 53, 225–236.
9. Mandelbrot, B. Fractals and Scaling in Finance: Discontinuity, Concentration, Risk; Springer: New York, 1997.





DOI 10.1002/cplx

10. Mandelbrot, B.; Hudson, R. The (Mis)Behavior of Markets, A Fractal View of Risk, Ruin, and Return; Basic Books: New York, 2004.
11. Taleb, N.N. The Black Swan; Random House: New York, NY, 2007.
12. Mandelbrot, B. Statistical dependence in prices and interest rates. In: Papers of the Second World Congress of the Econometric Society; Cambridge: England, 1970.
13. Mandelbrot, B. Possible refinement of the lognormal hypothesis concerning the distribution of energy dissipation in intermittent turbulence, Volume 12 of Lecture Notes in Physics, Rosenblatt; M.; Van Atta, C., Eds.; In: Statistical Models and Turbulence (La Jolla, California); Springer: New York, 1972; pp 335–351.
14. Mandelbrot, B. Statistical methodology for non-periodic cycles: from the covariance to R/S analysis. Ann Econ Soc Meas 1972, 1, 257–288.
15. Mandelbrot, B. A subordinated stochastic process model with finite variance for speculative prices. Econometrica 1973, 41, 157–160.
16. Mandelbrot, B.; Calvet, L.; Fisher, A. The multifractal model of asset returns, Discussion Paper 1164 of the Cowles Foundation for Research in Economics; Yale University: New Haven CT, 1997.
17. Calvet, L.; Fisher, A.; Mandelbrot, B. Large deviations and the distribution of price changes, Discussion Paper 1165 of the Cowles Foundation for Research in Economics; Yale University: New Haven CT, 1997.
18. Fisher, A.; Calvet, L.; Mandelbrot, B. Multifractality of deutsch-mark/US dollar exchange rates, Discussion Paper 1166 of the Cowles Foundation for Research in Economics; Yale University: New Haven CT, 1997.
19. Mandelbrot, B. Renormalization and fixed points in finance, since 1962, Volume A263 of Physica, Iagolnitzer, D., Ed., In: Statistical Physics 20, International IUAP Conference: Paris, 1999; pp 477–487.
20. Mandelbrot, B. Scaling in financial prices, I: Tails and dependence. Quant Finan 2001, 1, 113–123.
21. Mandelbrot, B. Scaling in financial prices. II. Multifractals and the star equation. Quant Finan 2001, 1, 124–130.
22. Mandelbrot, B. Scaling in financial prices. III. Cartoon Brownian motions in multifractal time. Quant Finan 2001, 1, 427–440.
23. Mandelbrot, B. Scaling in financial prices. IV. Multifractal concentration. Quant Finan 2001, 1, 641–649.
24. Mandelbrot, B. Stochastic volatility, power-laws and long memory. Quant Finan 2001, 1, 558–559.
25. Mandelbrot, B. Parallel cartoons of fractal models in finance. Ann Finan 2005, 1, 179–192.
26. Mandelbrot, B.; Calvet, L.; Fisher, A. Multifractality of deutschmark/us dollar exchange rates, Cowles Foundation Discussion Papers, 2008, p 1166.
27. Mandelbrot, B. Parallel cartoons of fractal models in finance. Ann Finan 2005, 1, 99–100.
28. Mandelbrot, B. The inescapable need for fractal tools in finance. Ann Finan 2005, 1, 193–195.
29. Mandelbrot, B. Épistémologie du hasard dans les sciences sociales: Invariance des lois et vérification des prédictions, Encyclopédie de la Pléiade (Gallimard): Logique et Connaissance Scientifique, Piaget, J., Ed.; Gallimard: Paris, 1967; pp 1097–1113.
30. Zipf, G.K. Human Behavior and the Principle of Least Effort; Addison-Wesley: Cambridge, MA, 1949.
31. Feller, W. An Introduction to Probability Theory and its Applications; Wiley: New York, NY, 1971.
32. Lévy, P. Théorie de l'addition des Variables Aléatoires, 2nd Edition; Gauthier-Villars: Paris, 1954.
33. Mandelbrot, B. Towards a second stage of indeterminism in science. Interdiscplinary Sci Rev 1987, 12, 117–127.
34. Berger, J.M.; Mandelbrot, B. A new model for the clustering of errors on telephone circuits. IBM J Res Dev 1963, 7, 224–236.
35. Aitchinson, J.A.C.; Brown, J. The Lognormal Distribution; The University Press: Cambridge UK, 1957.
36. Mandelbrot, B. Self-similar error-clusters and the concept of conditional stationarity. IEEE Trans Commun Techol 1965, C.O.M.-13(June), 71–90.
37. Mandelbrot, B. Very long-tailed probability distributions and the empirical 'distribution of city sizes', Massarik, F., Ratoosh, P., Eds.; In: Mathematical Explorations in Behavioral Science, R. D. Irwin, Inc.; Clarendon Press: Homewood, IL, 1965; pp 322–332.
38. Cootner, P.H. The Random Character of Stock Market Prices; MIT Press: Cambridge, MA, 1964.


DOI 10.1002/cplx