Mandelbrot Makes Sense: A Book Review Essay

A discussion of Benoit Mandelbrot's T*he (Mis)Behavior of Markets* by Nassim Nicholas Taleb

I closed this book feeling that it was the first book in economics that spoke directly to me. Not only that, but the astonishing simplicity, realism, and relevance of the subject makes it the *only* general work in finance I've ever read that seemed to *make sense*.

Benoit Mandelbrot makes sense. Just as he used us common readers outside the ivory tower to force his fractal ideas into science (where they became "part of the scientific consciousness"[1]); he may just be the one to help turn economics into something real.

This first essay is non-technical and general[2] (i.e. can be read by someone without a mathematical background) and focuses around the topics covered in this book. The second one is more technical and it goes deeper into the epistemological problems of "fat tails", concentration, and extreme events.

What do fern leaves, commodity prices, computer book sales, income distribution, the coast of Britain, cauliflowers, and the intricacies of the vascular system have to do with one another? Mandelbrot's work revolves around the simple practical application of a concept called "fractal" in replacement for more complicated mathematical tools that are universally used without empirical justification.

Triangles, squares, circles, and other geometric concepts that caused many of us to yawn in the classroom, may be beautiful and pure notions; but they seem more present in the mind of mathematicians and schoolteachers than in nature itself. Mountains are not triangles or pyramids; trees are not circles; straight lines are almost never seen anywhere. To figure out how the world operates, we need a different geometry than the classical one developed by Euclid of Alexandria some 2400 years ago. Drawing on a list of then obscure (but subsequently made famous) mathematicians, BM coined the word *fractal geometry* to describe these objects that are jagged yet *self-similar* in the sense that small parts resemble, to some degree, the whole (a more mathematically appropriate designation would be the broader "self-affine" but, somehow, designations are sticky and, in this discussion, self-similarity should be held to be "self-affine"). Leaves look like branches; branches look like trees; rocks look like small mountains. If you look at the coast of Britain from an airplane, it resembles what you get using a magnifying glass. This character of self-affinity implies that one deceivingly short and simple rule of iteration can be used, either by a computer, or more randomly, by Mother Nature, to build shapes of seemingly large complexity. He designed, or rather, according to Sir Roger Penrose[3], *discovered* an object, known as the "Mandelbrot set", which became popular with followers of chaos theory as it generated pictures of ever increasing complexity using a deceptively minuscule recursive rule, one that can be reapplied to itself repeatedly. You can look at the set at smaller and smaller resolutions without "ever" reaching the limit; you will continue to see the recognizable shapes.

The introduction of fractals was not initially welcomed by the mathematical establishment. This method of pictorial presentation did not seem to correspond to what seemed "to be mathematics" in the self-defining discipline. It is thanks to its popularity with physicists and other applied scientists, themselves following the lead of the general public (mostly computer "geeks"), that fractal geometry vindicated its way into the now-broadened field of mathematics. For *The Fractal Geometry of Nature* made a splash when it came out a quarter century ago. It spread across the artistic circles and led to studies in aesthetics, architectural designs, even large industrial applications. BM was even offered a position at a medical school! His talks were invaded by all manner of artists[4], earning him the nickname "the rock star of mathematics". The computer age thus helped him become one of the most influential mathematicians in history, in terms of the applications of his work, way before his acceptance by the ivory tower. We will see that, in addition to its universality, his work possesses an unusual attribute: it is remarkably easy to understand.

A Polish-Lithuanian Jew who found refuge in France as a child, BM is also a refugee from the French mathematical establishment protective of the "purity" of mathematics. To borrow from the late probabilist and probability *thinker* E. T. Jaynes (a man who went deeply into the subject), it was said that "the French did quite useful mathematics before Bourbaki" – as the secretive guild-like organization installed a truly top-down view of the subject matter, insuring no corruption by earthly material. Indeed many physicists have been horrified at the extent and side effects of such purism, with Murray Gell-Mann calling it the "Bourbaki Plague", and attributing the divergence between pure mathematics and science to the obscure language of the Bourbakists[5].

In a way, the separation between geometry and algebra can be seen as the separation of images and words in human expression and thought – just imagine a world in which images were barred. The Bourbaki-inspired purblindness does not just limit the tools of analysis. Just like blindness, one of its effects is to reduce contact with reality. Platonic top-down approaches are interesting but they tend to choke under the occasional irrelevance of their pursuits. It is telling that BM's hero is Antaeus, son of Gaia the mother Earth, who needed periodic contact with earth to replenish his strength.

Owing to the vicissitudes of a clandestine life during the Nazi occupation of France, the young Benoit was spared some of the conventional Gallic education with the uninspiring algebraic drills, becoming largely self-taught with some assistance from his uncle Szolem, a prominent member of the French mathematical hierarchy and professor at the College de France. Instead, he developed an encyclopedic knowledge of the history of mathematical thought. He also gave free course to his geometric bent. Untrained in the usual equation solving techniques, he passed the entrance exam to the

elite École Normale using purely geometric intuitions (this should be a hint for educators: consider how much more intuition you can develop with images instead of words). But he left after two days. Already stubborn, unruly and unmanageable, he moved to the more engineering-oriented École Polytechnique. He then settled in the United States, working most of his life as an industrial scientist for IBM, with a few transitory and varied academic appointments. Indeed, thanks to the computer, he could let the potent machine express his geometric hunches and lead through the subject matter's natural course. Indeed, the computer played two roles in the new science he helped conceive. First, these fractal objects, as we will see, could be generated with a simple rule applied to itself which is ideal for the automation of a computer (or mother nature). Second, in the generation of visual intuitions lies a dialectic between the mathematician and the objects generated. A mathematical scientist *par excellence*, in a subject matter that did not (then) exist institutionally, he was held to be a mathematician for scientists and a scientist (particularly a physicist) by the mathematical establishment. And while mathematicians burn out in their twenties, he received his first academic tenure at Yale when he was 75 years old. Indeed, after a stint at Harvard where computer and mathematics are subjected to a conceptual separation, it is at Yale that BM[6] got his dream job as a Professor of *Mathematical Sciences*. And it took him half a century to fully realize what his work was united by an attribute: *roughness*, not just as a quality of objects, but as a standalone field of study. It is impressive to see him as the embodiment of a scientific

thinker who had the luxury to take his time to grow his ideas. (Charmingly, BM, in his scientific writings, when discussing a contribution made by a mature mathematician, mentions his age, such as "Cauchy, at the age of 64..."). It is thanks to such maturation that he joins that category of the classical, pre-academic specialization of the wisdom-generating *natural philosophers.*

What does it all have to do with finance? Can we extend the concept of fractals and self-similarity to statistical frequencies? It would make the concept of astonishing universality. This would make BM the true Kepler of the social sciences. The analogy to Kepler is at two levels, first in the building of insights rather than mere circuitry, second because you can step on his shoulders – the title of Kepler or "Newton of the social sciences" is one so many thinkers with grand ideas have tried to grab (Marx for one aimed at being the Newton of the sciences of man). I am not in the business of defining genius, but it seems to me that the mark of a genius is the ability to pick up pieces that are fragmented in people's mind and binding them together in one, a meta-connection of the dots.

Do probabilities (more exactly, cumulative frequencies) scale like cauliflowers? If so, the implication is not trivial as we may be on to something general, working across sciences and fields. And if so, then the statistical attributes of financial markets can be made far more understandable than by the complicated and middlebrow so-called "Gaussian" framework. Indeed there is something about BM's work that makes him and his ideas far more understandable to the common man

than the theories of financial economists, and, which is worrisome, more *understandable* by the common man more than by the classically trained economist – just as the computer graphic designer or a computerized teenager could get the point far more easily than a classically trained mathematician.

It is not a well-known fact that before his involvement with the roughness in the geometry in nature, BM started his career focusing on problems in social science and finance; it is certainly there that most of his ideas were refined. He initially wrote papers in the 1960s presenting his ideas on "infinite

This would make BM the true Kepler of the social sciences ... first in the building of insights rather than mere circuitry, second because you can step on his shoulders

variance", getting some early acceptance, but rapidly causing anxiety in financial economics circles. He then moved to the less harmful fields of geometry and physics, returning to finance in 1995 when he started a very active production of scientific papers on financial risk. At eighty, he shows no sign of relenting, producing, as I said, the deepest and most realistic finance book ever printed. By writing *The (Mis)Behavior of Markets* in collaboration with Richard Hudson, a long time journalist at the *Wall Street Journal*, he seems to be employing the same

strategy of going straight to practitioners and the general public and bypassing the academic establishment, a task that might appear easy with economics given that the public and professional standing of economists in general and finance academics in particular is one of the lowest of any specialty. So the mission of toppling these fake and empirically invalid beliefs seems trivial.

Or is it? Finance academia, unlike the physics establishment, seems to work more like a religion than a science, with beliefs that have so far resisted any amount of empirical evidence (actually this statement is quite mild; it works just *like* a religion totally impervious to news from reality). The closest field to finance in the history of science would be pre-Baconian medicine as practiced in the Middle Ages, either disdainful of observations or spinning them with theological arguments. financial theory being a fad, not a science, it will take a fad, and not necessarily a science, to unseat its current set of beliefs.

BM wrote his doctoral thesis on what seemed to be two subjects at once: mathematical linguistics and statistical thermodynamics (de

Broglie was the head of the thesis committee). Before the advent of Information Theory as a discipline, such mixing seemed quite strange. A quip goes to the effect that, of his two topics, the first did not exist yet and the second no longer existed. But the unity between the two was the so-called "fat tails" and "power laws" that are now becoming increasingly popular in physics and social science, though not in economics. The spark came from the so-called Zipf's "law" in linguistics, after the works of one George Zipf on the relative ranking in the frequency of words in a vocabulary. BM debunked Zipf's belief in the separation, thanks to these laws, between social and the natural sciences: these "fat tailed" phenomena also existed in physics. We are just blind to them.

BM later built on the works of the (then) unknown mathematician Paul Lévy and, to a lesser degree, the trader-economist Wilfredo Pareto to whom the original power law is attributed. The designation "L-Stable" distributions, (for "Lévy-Stable"), a.k.a. Pareto-Lévy distributions comes from Mandelbrot. I prefer to use the designation "PLM" (Pareto-Lévy-Mandelbrot) for the more general case of a random series with both independent and non-independent increments.

Let us see how power laws, with their scalability, i.e., the asymptotic settling of a series to a constant limit in the relationship between likelihood of events, can be seen as an application of fractal geometry. Consider wealth in America. Assuming we reached the "tail" , the number of people with more than two million will be around a quarter of those with more than one million. Likewise the number of persons
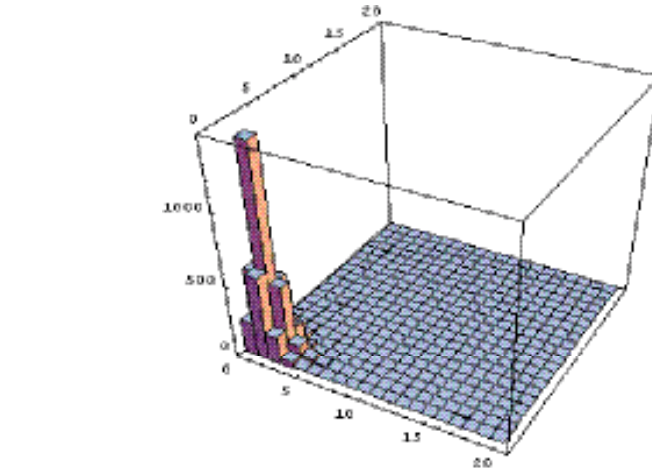

Figure 1 The Cauliflower theory of frequencies. This is the result of the application of power laws dynamics to a wealth process. If you divide the area in smaller ordered sub-samples, you will see the same inequality prevailing.

with wealth in excess of 20 million will be approximately the same in relation to those with more than 10 million: about a quarter. This relation (here the square of the ratio) is called a scaling law, as it is retained at all levels, no matter how large the number becomes (say two billion in relation to one billion). What is critical here is that it does not vanish – frequencies get lower for higher wealth levels, but the ratios between two arbitrarily high numbers do not decrease!

Cauliflower? If you separate the frequencies you will find that the sub-samples resemble each other in the degree of inequality in the different ordered sub-sections, as can be shown in Figure 1.

Note that the "tail" is the point where the outcomes become scalable in cumulative probability; it does not have to be a transition point (it can be an asymptotic property as we tend towards it). This scalability seems to apply to a variety of phenomena like book sales, nodes on Google, the relative size of cities, the number of times an academic paper is cited, the number of casualties in

wars, and, of course, market movements. The implication of these power laws is that, for most, there is generally no "standard" deviation from the norm. In the previous example of wealth, if there are more than 1/4 the number of people with a 2 times a given level of wealth than with a given level (more technically, when the tail exponent is higher than 2 since doubling the wealth threshold here leads to an incidence of more than the square of the ratio), then we are dealing with undefined variance. Now, worse, when the frequency in the previous example drops by less than half, then we are in a situation of extreme fat tails: there is no known average. Any arbitrary large number can take place that can disrupt the mean. The concept of average is meaningless, totally meaningless as a characterization of the attributes of a very fat-tailed process, such as computer firms. The notion of a "typical" computer company has nothing to do with anything. Likewise characterizing a "typical" writer provides no information. Just consider how unstable these variables can be: imagine what

would the arrival of Bill Gates to a town do to the average wealth there.

It is worrisome because every student of statistics learns about mean and variance as the foundations of their methods.

The Gaussian, in contrast, is not scalable. Most observations hover around the mediocre, and deviations either way become increasingly rare, to the point of there being events of an impossible occurrence. Take the number of adults heavier than 300lbs and those heavier than 150lbs. The relation between the two numbers is not the same as the one prevailing between 600 and 300lbs. The latter will be considerable smaller. It gets smaller as the number get larger – meaning that there is no self-affinity. Deviations from the norm decrease very rapidly, *at an increasing rate*, to the point where some high number becomes literally impossible. The increase of the rate of the decrease is what prevents scalability. BM calls this type of randomness "mild", as compared to the "wild" one generated by power laws. There is a beautiful sentence in the book differentiating between the two: "Markets often leap, don't glide".

To further see the link between finance and fractal geometry, pick a financial chart. Just like the coast of Britain, self-similar at all resolutions, monthly prices look "like" (i.e. present an affinity with) hourly charts. One has to shrink the timescale more than the price scale in order to get the same effect. Furthermore, if the stretching is done in a random manner, itself fractal, one ends up with what Mandelbrot calls multifractal.

In 1963 BM wrote a paper on the properties of financial prices and

found them to be scaling power laws of the anxiety-causing types – the "infinite variance" variety. The paper was initially endorsed by the orthodox finance establishment, accepting the implications that there is not "standard" risk, no known risk. But suddenly, these academics started looking the other way as "modern portfolio theory", linking risk and return, was born. There had to be a *measure* of risk, even if it presents the fatal contradiction of not working when you need it. The bell curve describes the equivalent of the odds of an uncomfortable airplane ride, nothing about the risk of crash – but operators thought thanks to "science" they were now in control.

If you asked for the bridge between the arts and science, the notion of fractal would come up. If you ask about what bridges hard and social sciences, the same scalable laws would come up. Doesn't this make BM the universal scientist?

Most of the effects of NonGaussianism flow from the consequence that a small number of observations might contribute disproportionately to the total mean and variance. Pending on the gravity, you either need a very large, possibly infinite, sample to track the properties. Indeed, if ten days in a decade represent 40 per cent of the returns, which we tend to see routinely with financial securities, much of conventional sampling theory goes out of the window. Consider that under a Gaussian regime, since these outliers represent a small share of total variations, you should be able to obtain the properties of , say, the stock market by being in it a small sub-segment of the time. Diversification, too suffers from the consequences of scalability. Since fat tails create a winner-take-all environ-

ment, you may not be diversified as much as conventional theory indicates. And conventional statistical theory might make you jump to consequences too quickly: your sample size is smaller than you think.

I was trying to explain the difference between two modes of thinking, the broad and the narrow, to an investor. Remarkably, it corresponds to the difference between power laws and Gaussians. As a method of risk management, he follows the conventional methodology of collecting past returns, building a database, and simulating by drawing from the past, thanks to bootstrapping-style methods. Using such an approach would make him select the largest possible deviation in the simulation as the worst scenario. A method of say, fitting an "empirical probability distribution", would do almost the same. This is an interpolative method – of course the worst possible move in the future is going to be similar to the one in the past, though these moves did not take place in the past's past. After the stock market crash of 1987, they simulate using 22 per cent as the worst daily deviation. Don't they realize that before the crash they would have used the preceding worst case and missed on such a big event?

Both the Gaussian and our conventional wisdom are interpolative. Power laws are *extrapolative*. You look at the ratio of millionaires to bi-millionaires and can translate it into the ratio of 10-millionaires to 20-millionaires. Likewise the ratio between 5% and 10% moves allows you to infer the incidence of moves in excess of 20%.

I will rapidly go through the details of the book .

The first part of the *Mis-Behavior of Markets*, out of three, presents the

very sad history of modern finance. It ends with the presentation of the evidence against these models. It does not take a lot of empiricism to figure out that such risk measure is useless: the stock market crash of 1987 had, according to their models, such a low probability, one in several billion billion billion years , that it should not have happened (probabilities that low are no longer measurable; it is meaningless to argue whether to assign a 10-23 or a 10-12 probability to these). You do not need a lot of empirical work to realize that a model is wrong: one single instance suffices to invalidate it.

Another piece of evidence among many is the hedge fund Long Term Capital Management that went bust in 1998. It employed 25 PhDs, and two "Nobel" medalists in economics for their work in finance. Aside from the fact that their "Nobel" was mistakenly presented for inventing a "formula" – the formula has been there for a while; what they did is make it fit into the prevailing economic arguments. They used complicated mathematical models – they should have had on their staff more street-smart cabdrivers who do are privileged to not know economics. LTCM is a milestone as a catastrophe that was caused by the pseudo-science of economics, much like the side effects of those medieval medical remedies.

The second part discusses the fractals theory and its relation to the power laws. Those familiar with BM's ideas from James Gleick's *Chaos* will see the usual themes presented. It ends with the multifractal model where BM presents a memory of prices similar to those of the floods by the Nile river; what happened a decade ago stays lurking in

memory – in other words we are no longer dealing with serially independent draws. The mathematics is more intuitive and more realistic than what we are used to; indeed there is no mathematics but graphs and geometric intuitions. He presents the usual attacks on his model that consist in saying, "daily prices might be nonGaussian but in the long term things become Gaussian". Long term? After the bankruptcy? Long Term Capital Management was a "long term" idea as well. Under leverage there is no such thing as long term.

The third part wraps up with more railing against finance theory and some suggestions for further research. It includes a scene with journalistic overtones of a visit to the laboratory of randomness specialist Richard Olsen in Zurich.

<center>***</center>

This book has a crisp message about risk. The reviews were quite favorable, but distressing for us empiricists as few commentators got the point. People have difficulty dealing with the idea that one can write a general book on a financial topic without telling people about a new foolproof (and secret) technique about how to double their money in 21 days. My book *Fooled by Randomness* generated hundreds of letters with the following class of complaints: "you tell us that it is mostly luck, which seems reasonable, but you don't tell us how to make money out of this luck". People are so conditioned by advice-offering charlatans in business books that anything remotely away from it seems, as I was told, quite "odd". BM's, of course, does not give you a recipe. It was therefore amusing to see the book reviews complaining about the "now what?" –

how can we take these ideas home? The answer is clear: get out of the markets as we understand them less, far less than we are led to believe. That would be a significant first step. What this book is about is the variability of markets and their risks, period.

The central idea about risk management that preoccupies me currently is as follows. If you save people in the process of drowning you are considered a hero. If you prevent people from drowning by averting a flood you are considered to have done nothing for them. Such asymmetry is apparent: you do not get bonus points for telling agents to *avoid* investing. They want "something tangible".

Likewise you do not go very far by telling people "we do not gain anything by talking about the variance". They want a risk number, a correlation number and BM takes it away from them (notice that undefined variance also means undefined correlation).

A simple implication of the confusion about risk measurement applies to the research-papers-and-tenure-generating equity premium puzzle. It seems to have fooled economists on both sides of the fence (both neoclassical and behavioral finance researchers). They wonder why stocks, adjusted for risks, yield so much more than bonds and come up with theories to "explain" such anomalies. Yet the risk-adjustment can be faulty: take away the Gaussian assumption and the puzzle disappears. Ironically, this simple idea makes a greater contribution to your financial welfare than volumes of self-canceling trading advice.

The possibility of "infinite variance" (or more appropriately "unde-

fined variance"), implies that when you take a sample from a long series, every sub-sample yields a different measure of volatility.

Nor does it look like the fudging of the finance models can produce real results. I will omit discussing the repackaging of the Capital Asset Pricing Model under the newer "Arbitrage Pricing Theory", except to bemoan that, seven years after LTCM, the most recent issue of the *Journal of Economic Perspectives*[7] celebrates with some pomp the 40th anniversary of Modern Portfolio Theory. It is saddening to see that so few realize its epistemological dangers. One insightful and honest article, by Fama and French, talks about the poor "empirical" results, accept-

the Swedes for his GARCH process, made the following comment: whether or not you include the stock market crash of 1987 or not makes a huge difference to the choice of model and its parameters. GARCH, extremely fragile in its calibration, is very sensitive to the inclusion of such large observations from the deep past. Does it smell like undefined variance to you?

I am dedicating my next book, *The Black Swan*, to BM for his 80th birthday. I can now safely say, in spite of my having had discussions with hundreds of hotshots, that he is the first person who ever taught me anything meaningful about my subject matter of uncertainty. More specifically, it was the first time in

"mean-divergence" , and the other more gullible "short volatility" who believe in models, "mean-reversion", "arbitrage", the self-canceling activity called statistical arbitrage, and similar things. In other words, there is the naive and the skeptic. Scientists and academics tend to squarely fall in the second category, even when they trade, while veteran traders and real practitioners have the first mindset. It was a surprise to encounter BM, a scientist of the "long-vol" category.

It was also refreshing to find someone who shared the same allergies. It was not just the notion of variance; small details can be revealing. For instance, we both got independently offended by the same

People readily mistake irreverence towards some class of accepted heroes for arrogance. A fair approach would be to examine the targets of such irreverence

ing the notion that empirical implies in practice, *out of sample* and realizing that, in the end, its appeal lies far more in teaching MBA students than anything else.

Now there have been fixes to these equations to accommodate fattails, to no avail. Every option trader knows that volatility is variable – but models such as GARCH, with close to 10,000 academic publications, do not seem to bring us closer to anything. Making volatility variable is more complicated than we think: there is the problem of the specification of such variability. At the last ICBI Madrid Derivatives Convention, Robert Engle, freshly medalled by

my life that I had a conversation with someone who can naturally hold that the notion of "variance" is meaningless in characterizing uncertainty – and we could move on to a more meaningful discussion of the subject. I finally found someone I could talk to without feeling deep strain and tension.

There is more. He could communicate with the trader in me. I was taken aback by how easily his ideas spoke to me, down to the very practical. We traders divide persons into two categories: those with a "long volatility" frame of thought, who, in general, never rule out blowups, change, trends, conspiracies, and

statement that "nature does not make jumps".

So time lost was made up and it was refreshing to discover the personal charm of the universal philosopher and be privileged to his conversation partner. BM only lives five kilometers away from my house, which means that we spent more time talking on the telephone than meeting in person (this is how these things work). Conversations with him are punctuated by opened-and-closed parentheses, with tours of classical literature, history, science, music, back to science, with digressions rarely left hanging. Not surprisingly, he is an independent thinker in just

about everything; he is a pack of intuition; he is encyclopedic and is a universal conversationalist. If you manage to age well, you actually get better because you know so many more things. And he has an astonishing memory ("*une memoire d'éléphant*").

Having read descriptions of his personality, I was taken aback by the difference between the real man and the reputation of "arrogance" – which to me (as I am familiar with such accusations) comes merely from his targeted irreverence and lack of willingness to put up with established truths and established gods. People readily mistake irreverence towards some class of accepted heroes for arrogance. A fair approach would be to examine the targets of such irreverence.

In a way BM is the exact opposite of what I call the academic clerk: someone who is there to work on research like an obedient tax accountant. BM is a maverick, tenacious, and idiosyncratic in his approach; he seems to scorn formalities. It is all-natural that he would have had to counter resistance from the clerks. I was in for a surprise: I had the feeling of talking to a trader, capable of revising his views at a blip. And the man was simple, friendly, charming, the reverse of arrogant – except for his colorful irreverence. Consider that one of his colleagues, Michael Frame[8] who was also told that BM was "arrogant", accounts for his surprise upon having to contradict BM on a critical point. BM's reply was "Marvelous. The problem is more interesting than I had expected".

One final remark about recognition. When Daniel Kahneman received the Nobel medal many people congratulated him on such an honor. My reaction was to congratulate the Nobel committee: finally, these Swedes seem to be serious about their prize. Not only have they helped to make economics more of a science, but they also gave it the credentials to help enrich other disciplines. The abundance of data makes the field of economics an ideal laboratory to develop insights and quantitative tools helpful to other sciences – we can develop insights about human nature from economic choices (Kahneman and Tversky); we can also learn new mathematical methods (Mandelbrot). I hereby ask the Swedes to take some perspective and think of those whom, a century from now, will be identified as having changed the way we view the world.
■ *The (Mis)Behavior of Markets: A Fractal View of Risk, Ruin, and Reward* by Benoit Mandelbrot & Richard Hudson, Basic Books.

FOOTNOTES

1 Kenneth Falconer, *Nature*, 430/ 1 July 2004.
2 A shorter version of this book review was withdrawn from the *Los Angeles Times*, partly because I got too close to Mandelbrot after writing the review and did not want to bear the risk of personal conflict.
3 Roger Penrose, 2005, *The Road to Reality*, New York: Knopf.
4 John Brockman, 2005, Discussion with Benoit Mandelbrot, www.edge.org
5 See the posthumous *Probability Theory: The Logic of Science* by E.T. Jaynes, 2003, Cambridge University Press.
6 See Mandelbrot's essay on www.edge.com
7 *Journal of Economic Perspectives* Vol. 18, No. 3, Summer 2004
8 See the personal testimony in Michel L. Lapidus (editor): Fractal geometry and applications: A Jubilee of Benoit Mandelbrot, *Proceedings of Symposia in Pure Mathematics*, 72, 1, American Mathematical Society.

Fat Tails, Asymmetric Knowledge, and Decision Making

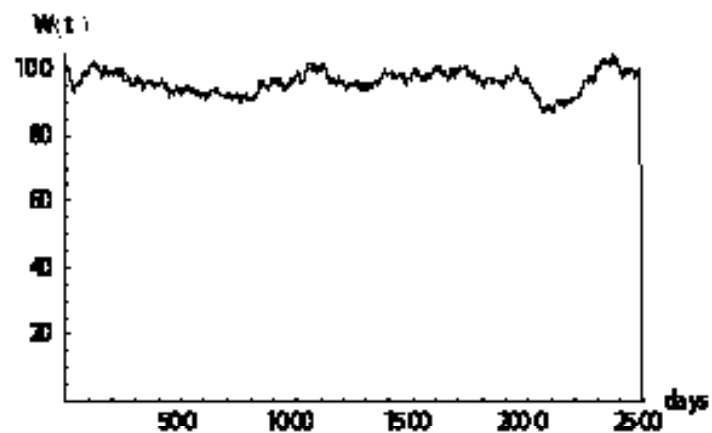Nassim Nicholas Taleb's Essay in honor of Benoit Mandelbrot's 80th birthday

Figure 1: A dataset of 2,500 prices. Infer the attributes.

Introduction

Consider the following thought experiment. You show an agent a set of data of 2,500 days worth of returns (the resulting asset price $W(t)$ being represented in Figure 1) and ask him to infer the attributes of what he saw.

Odds are that he would tell you that the log-returns are Gaussian. 2,500 days data set represents an ample sample size by any measure, enough for the distribution to reveal itself to us. Clearly all the attributes of a mild distributions are there: no excess Kurtosis over that of a Normal, no outliers, no jumps, no gaps; a histogram of the returns would reveal the Platonic Bell Shape.

Now we continue with the rest of the story. We add one day, number 2,501; one single day can show a quite different picture.

Picture 2 shows the information-al increase by that one day. The generating process for these draws is a mere switching process, built around a Gaussian, to which was added the occasional drawing, once in 2,500 days, from an infinite variance kick. This implies that the total is of infinite variance. Those who have not seen any such situation should take a look at emerging market currencies (those in a managed regime). It can also apply to a hedge fund returns: The properties of the late hedge fund LTCM are not too different from what we just saw. The bigger the divergence between the two regimes (the "normal" and the "unusual"), the worse the epistemological picture as more people will tend to be fooled by what they saw.

The central problem of uncertainty

What I call the central epistemological problem of uncertainty[1] is sum-

marized as follows: we do not observe probability distributions, only random draws from an unspecified generator. So we need data to figure out the probability distribution. How do we gauge the sufficiency of the size of the sample? Well, from the probability distribution. If at the same time one needs data to figure out the probability distribution, and the probability distribution to figure out if we have enough data, then we have a severe circular epistemological problem.

Note here that fat tails are contagious. If you combine two random variables each following a power law distribution but with different exponents, the result is a power law distribution with, for tail exponent, the lower of the two. Here we have two processes, one of finite, the other of infinite variance; accordingly the infinite variance will prevail.

A traditional philosophical way to deal with the regress argument, if one follows the epistemological traditions, would be to either 1) put your hands up and bemoan the Problem of Induction, and find theological arguments to have some unquestioned belief or 2) proceed to a systematic layering: One can pose a meta-distribution, one that would

take into account the probability of the candidate distribution being the wrong one. You can use priors and probabilize with series of meta-probabilities. Neither handy, nor convincing, and it implies as Elie Ayache[2] put it in this magazine "trying to find a random generator behind the random generator". And it does not escape the attacks by classical Pyrrhonian skeptics: we seem to be either 1) justifying belief with reference of other belief, itself justified by other belief, all the way up until some unargued dogma, which could be fragile (in this case some "known" distribution or generator for the time series) 2) justifying belief somewhere in the loop with another previously derived belief and falling back into severe circularity; finally [3]) the regress may never end and we stay at the beginning.

Note that the quantitative-statistical literature is not thoughtful enough or self-critical "to be even wrong" on the subject. How? Conventional tests of normality study the square errors from a Gaussian and use a Gaussian-inspired distribution (a special case of the Gamma distribution, the Chi-Square, which is the distribution of

the squared Gaussian variate) . This is exceedingly circular and reflects a severe lack of awareness of such circularity.

An easier solution

As an operator first and last, I believe that there are, however, far more elementary (and practical) ways to deal with this problem, or at least to protect ourselves from its ill effects. How? I propose two approaches.

First, consider Pascal's wager. We can change our payoff structure to accommodate what absence of knowledge we suffer from, and with respect to which moments of the distribution. For instance, if the data has "infinite" (or undefined) variance, one can avoid exposure to such infinite tail by clipping the sensitivity to the offending part of the distribution. Purchasing a simple derivative(say, an extremely out-of-the-money call), if it such product is available, may provide a solution. Our doubt can be targeted and remedied by transactions. *Tout simplement*.

Second, what we call the masquerade problem. The data cannot tell us what is the probability distribution generating it; but it can easily tell us what such probability distribution *is not* (or is not likely to be), and which moments of the distributions we may not be able to compute.

Portfolios, infinite variance, and epistemic opacity

What many academic philosophers do not realize is that the limits of some knowledge may be of small moment. I would rather use my energy in changing my payoff structure rather than getting into intractable issues and playing philosophaster. My colleague, another option trader and empirical philosopher Rabbi Anthony ("Tony") Glickman (also a

Talmudic scholar), explains quite eloquently that being an option trader gives someone a philosophical approach along "long gamma" lines, or, more formally in the decision theory literature: along a mind-set focused on the convexity of payoffs. One comment I make here about Tony is that his definition of philosopher is similar to mine (and Mandelbrot's): a philosopher is someone who specializes in ideas, not in other people's ideas – like stamp collecting. Professional philosophers can be like parasites. To Tony, like for me, *being long an option in the tail (or more generally "long convexity") eliminates the need to try to figure out what we don't know*[3]. Only an option trader could understand that – that's what I am trying to generalize to all decision making under uncertainty and convey to nontraders in my forthcoming *The Black Swan*.

It is key that we operators and decision makers are capable of insulating ourselves from nasty parts of the distribution. It is a fact that a portfolio constituted of securities that have infinite variance does not need to have infinite variance. How? If you are short a call spread with the position strike K, described as short a call struck at K, long another call at $K + y$, you are "short volatility", but you are not exposed to infinite variance. Your payoff is capped. Furthermore: *the properties of your strategy are not fragile to parametric assumptions or choice of mode*l.

Note here, in the earlier thought experiment, that the moments of the distribution are very precarious; the loss L (taken in Log returns) is so large that the moments are insensitive to the probability of the big loss π. Indeed the pair πL (probability times the payoff) is so large that we
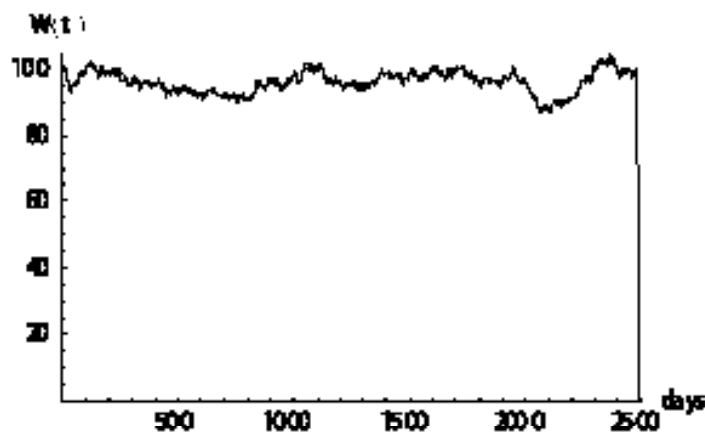

Figure 2: A dataset of 2,501 prices. What is the informational increase?

may never care about the size of the probability. It is so obvious that we should work to control L – or, if we can't, to only enter transactions where such L can be controlled.

Now the question: what if we can't insulate ourselves from such distributions? The answer is "do something else", all the way to finding another profession. Risk managers frequently ask me what to do if the commonly accepted version of Value-at-Risk does not work. They still need to give their boss *some* number. My answer is: clip the tails if you can; get another job if you can't. "Otherwise you are defining yourself as a slave". If your boss is foolish enough to want you to guess a number (patently random), go work for a shop that eliminates the exposure to its tails and does not get into portfolios first then look for measurement after. Indeed if like me you think that Modern Portfolio Theory is charlatanism (as confirmed by my trader's observations and empirical research, and Mandelbrot's work), use portfolios that do not depend on their measurements. It is so easy to avoid traps.

The asymmetric masquerade problem

A power law (as we saw in the thought experiment) can easily masquerade as a Gaussian but not the reverse (at least not easily). We can reject the Gaussian more easily than we can accept it. More generally, a distribution with fat tails can show milder tails than its "true" properties, except, of course, when it is too late. It will even tend to do so. The small sample properties of these processes are such that we are not likely to encounter large moves in them. We can call that problem an "epistemic headwind". To answers

some questions put to me in this magazine about skepticism and asymmetric knowledge[4], I will use the argument that it is always easier to figure out what the distribution is not than what it is. Compare that to the attributes of humans: a criminal can masquerade as an honest citizen; an honest citizen cannot as easily fake being a criminal. Many extensions of this point are accepted in many fields: one single event constitutes a catastrophe; one needs many days without an event to pronounce an environment as catastrophe-free.

This asymmetry is at the core of skeptical empiricism: our body of knowledge is more readily increased by negative observations than by confirming ones. Remarkably, we can do something with this; it leads us to a ranking of the robustness of results. And remarkably, it is because I elect to behave operationally as if the market followed a Mandelbrot-stable process that I can build portfolios that I am comfortable with. A Mandelbrot-stable variable is simply here what is called a Levy stable, but with non-serially independent draws (what BM calls multifractal). We will return to the situation.

The α problem

Take X a random variable, we have a power law $P[X > x_0] \sim O(x_0^{-\alpha})$. Clearly we are told that if the first and second moments of the distribution are defined, i.e., $\alpha > 2$, then, under aggregation the series becomes Gaussian so we can use the conventional tools of analysis. Note here that this only holds if we have independent increments.

BM came up with papers in the 1960s[5] showing cotton prices with tail $\alpha < 2$, in other words implying Levy-stability; the distribution has fat tails and does not become a Gaussian

under aggregation. There have been series of papers[6] disagreeing with Mandelbrot's early work and its conclusions. Researchers tend to be "skeptical" about the Lévy regime hypothesis producing, for more than a quarter century now, "evidence" to the effect that Mandelbrot's early characterization of infinite variance is wrong – people seem to very badly need a Gaussian in order for them to operate with the current academic framework. Their methodology is based on two arguments, first, the "observation" of $\alpha > 2$ and, second, the examination of the behavior of the data when they lengthen the time observation period.

These studies are either inconsequential or wrong in their inferences. First, it does not make much difference whether or not we are in a Lévy regime since we don't really stay in the Gaussian regime in the parts of the distribution that matter. Second, we do not "have evidence" that we are not in a Lévy regime. Third, we need to go beyond the "Lévy regime" and consider the Mandelbrot regime by lifting the too-restrictive assumption of independent increments. I will get into the details of the arguments next.

Point 1: The slowness in the rate of convergence makes a cubic α very seriously NonGaussian.
If we accept that α is approximately 3, "outside the Lévy regime", we are still in trouble with respect to the convergence to the Gaussian. Finite second moment implies convergence under aggregation, but we need to remember that with $\alpha < 4$ have an undefined 4th moment. The implication is rather serious. Consider that the 4th moment is the variance, corresponds to the error of the measurement in the variance (what we option traders call the "Vvol"). It will be infinite! This implies a quite nasty rate of convergence. There will always be a NonGaussian jump in the extreme tail to make the tail scalable.

Another way to view it is that the observations that we are adding are likely to be biased towards the middle of the distribution, making it converge in the body but much more slowly in the tails. We can examine this quantitatively. Take $\alpha = 3$. It is easy to show[7] that, in standard deviation terms, outside (Log (n)), with n the number of observations, we stay in a scalable regime. Even if you add up 1 million days, the Gaussian
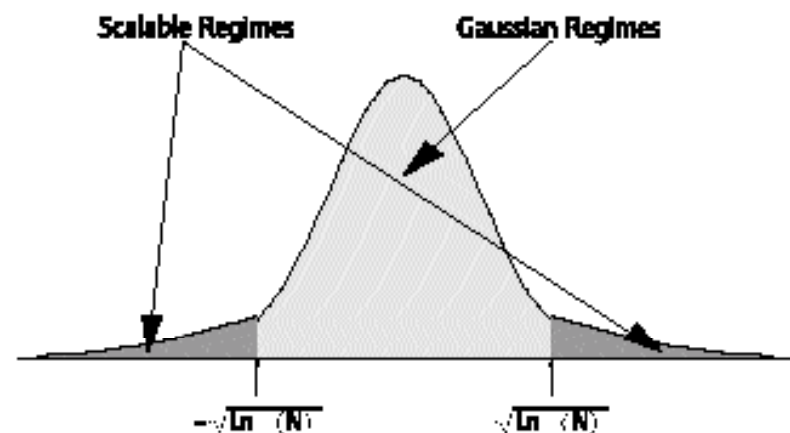

Figure 3: The regime densities.

regime stops at 3.7 sigmas! Typical penny-wisdom since the consequences of outside such moves are disproportionately large. Figure 3 shows the two-regime densities.

The situation is reminiscent of the value at risk problem. The tail of the distribution is where our errors compound. That is where ironically people like the precision.

Point 2: Absence of Evidence is Not Evidence of Absence: The Small Sample Bias Problem

Measuring an $\alpha > 2$ does not imply with any confidence that the "true" α is not <2. I avoid the discrepancies here in the measurement results from the various estimators, whether Hill and Log-Log linear regression. It just takes time (and data) for these distributions to reveal themselves. Simulate a series of symmetric random draws with $\alpha = 1.9$ and you will recover an α close to 3 with 106 samples. This is an argument well known to many traders[8] and discussed in Weron (2003).[9]

As we saw with the 2,500 day properties in the thought experiment, matters can be even more complex with a mixed process. In short, the fat tailed process tend to show the underestimation of the observed volatility.

Point 3: Where the Aggregation Fattens the Tails. Many of these inferences and indeed much of the mathematics we are used to assumes that we have independent draws

Now consider the following intuition: very bad moves generate very large up or down moves. And also consider that this may only happen in extreme circumstances, when the moves exceed a given threshold. Intuitively, a large loss might gener-

ate series of self-causing liquidations. What would that do to the scalability?

Well, in such mechanism, the aggregation *fattens* the tails. Such is the observation made by Sornette concerning the events leading up to the crash of 1987,[10] prompting him to analyze the properties of "drawdowns" independently.

This brings us to the Mandelbrot multifractal generalization[11] that shows that the process can have $1<\alpha<\infty$. Indeed much of the work on stable distributions is restrictive – obsessively relying on the assumption of independence.

Final note and consequences for Financial Engineering and Quantitative Finance.

I conclude by saying that to many of us the field of finance seem to be intricately linked to modern portfolio theory. I showed that it does not have to be so. And it does not take much to fix the problem.

FOOTNOTES

1 See Taleb and Pilpel, 2004. See also Elie Ayache , 2004a

2 Ayache, 2004b.

3 We can extend this convexity argument to the philosophy of probability in general. Take the subjectivist concept of probability as degree of belief attributed to De Finetti. Probability is held to be the price I am willing to fix in such a way as I would equivalently buy or sell a state of the world, making sure I remain consistent and avoid the "Dutch book" problem. Well, you do not have to put yourself in a situation which you have to trade.

4 see Ayache 2004.

5 Mandelbrot (1963) See also Mandelbrot (1997).

6 See Officer (1972), Stanley et al (2000), Gabaix et al (2003)

7 See Sornette (2004) for the proof. See also Bouchaud and Potters (2003).

8 Mark Spitznagel brought this to my attention.

9 Weron (2001).

10 See Sornette (2004) for the argument.

11 Mandelbrot (2000a, 2000b).

REFERENCES

■ E. Ayache,2004a, The Back of Beyond, *Wilmott,* 26-29

■ E. Ayache,2004b, A Beginning, in the end, *Wilmott,* 6-11

■ M. Blyth, R. Abdelal and Cr. Parsons,2005, Constructivist Political Economy, Preprint, forthcoming, 2006: *Oxford University Press*.

■ J.-P. Boucheaud and M. Potters, 2003, *Theory of Financial Risks and Derivatives Pricing, From Statistical Physics to Risk Management*, 2 nd Ed., Cambridge University Press.

■ X. Gabaix, P. Gopikrishnan, V.Plerou & H.E. Stanley, 2003, A theory of power-law distributions in financial market fluctuations, *Nature*, 423, 267-270.

■ B. Mandelbrot, 1963, The variation of certain speculative prices. *The Journal of Business*, 36(4):394–419.

■ B. Mandelbrot, 1997, *Fractals and Scaling in Finance*, Springer-Verlag.

■ B. Mandelbrot, 2001a, *Quantitative Finance*, 1, 113-123

■ B. Mandelbrot, 2001b, *Quantitative Finance*, 1, 124-130

■ R. R. Officer 1972 J. Am. Stat. Assoc. 67, 807–12

■ D. Sornette, 2003, *Why Stock Markets Crash : Critical Events in Complex Financial Systems*, Princeton University Press

■ D. Sornette ,2004, 2 nd Ed., *Critical Phenomena in Natural Sciences, Chaos, Fractals, Self-organization and Disorder: Concepts and Tools*, (Springer Series in Synergetics, Heidelberg)

■ H.E. Stanley, L.A.N. Amaral, P. Gopikrishnan, and V. Plerou, 2000, Scale invariance and universality of economic fluctuations, *Physica A*, 283,31-41

■ N N Taleb and A Pilpel, 2004, I problemi epistemologici del risk management in: Daniele Pace (a cura di) Economia del rischio. Antologia di scritti su rischio e decisione economica, Giuffrè, Milano.

■ R. Weron, 2001,Levy-stable distributions revisited: tail index > 2 does not exclude the Levy-stable regime.*International Journal of Modern Physics C* (2001) 12(2), 209-223.