

decisions around that.

This idea is often erroneously called Pascal's wager, after the philosopher and (thinking) mathematician Blaise Pascal. He presented it something like this: I do not know whether God exists, but I know that I have nothing to gain from being an atheist if he does not exist, whereas I have plenty to lose if he does. Hence, this justifies my belief in God.

Pascal's argument is severely flawed theologically: one has to be naïve enough to believe that God would not penalize us for false belief. Unless, of course, one is taking the quite restrictive view of a naïve God. (Bertrand Russell was reported to have claimed that God would need to have created fools for Pascal's argument to work.)

But the idea behind Pascal's wager has fundamental applications outside of theology. It stands the entire notion of knowledge on its head. It eliminates the need for us to understand the probabilities of a rare event (there are fundamental limits to our knowledge of these); rather, we can focus on the payoff and benefits of an event if it takes place. The probabilities of very rare events are not computable; the effect of an event on us is considerably easier to ascertain (the rarer the event, the fuzzier the odds). We can have a clear idea of the consequences of an event, even if we do not know how likely it is to occur. I don't know the odds of an earthquake, but I can imagine how San Francisco might be affected by one. This idea that in order to make a decision you need to focus on the consequences (which you can know) rather than the probability (which you can't know) is the *central idea of uncertainty*. Much of my life is based on it.

You can build an overall theory of decision making on this idea. All you have to do is mitigate the consequences. As I said, if my portfolio is exposed to a market crash, the odds of which I can't compute, all I have to do is buy insurance, or get out and invest the amounts I am not willing to ever lose in less risky securities.

Effectively, if free markets have been successful, it is precisely because they allow the trial-and-error process I call "stochastic tinkering" on the part of competing individual operators who fall for the narrative fallacy—but are effectively collectively partaking of a grand project. We are increasingly learning to practice stochastic tinkering without knowing it—thanks to overconfident entrepreneurs, naïve investors, greedy investment bankers, and aggressive venture capitalists brought together by the free-market system. The

next chapter shows why I am optimistic that the academy is losing its power and ability to put knowledge in straitjackets and that more out-of-the-box knowledge will be generated Wiki-style.

In the end we are being driven by history, all the while thinking that we are doing the driving.

I'll sum up this long section on prediction by stating that we can easily narrow down the reasons we can't figure out what's going on. There are: a) epistemic arrogance and our corresponding future blindness; b) the Platonic notion of categories, or how people are fooled by reductions, particularly if they have an academic degree in an expert-free discipline; and, finally c) flawed tools of inference, particularly the Black Swan-free tools from Mediocristan.

In the next section we will go deeper, much deeper, into these tools from Mediocristan, into the "plumbing," so to speak. Some readers may see it as an appendix; others may consider it the heart of the book.

-
- * This chapter provides a general conclusion for those who by now say, "Taleb, I get the point, but what should I do?" My answer is that if you got the point, you are pretty much there. But here is a nudge.
 - * Dan Gilbert showed in a famous paper, "How Mental Systems Believe," that we are not natural skeptics and that not believing required an expenditure of mental effort.
 - * Make sure that you have plenty of these small bets; avoid being blinded by the vividness of one single Black Swan. Have as many of these small bets as you can conceivably have. Even venture capital firms fall for the narrative fallacy with a few stories that "make sense" to them; they do not have as many bets as they should. If venture capital firms are profitable, it is not because of the stories they have in their heads, but because they are exposed to unplanned rare events.
 - * There is a finer epistemological point. Remember that in a virtuous Black Swan business, what the past did not reveal is almost certainly going to be good for you. When you look at past biotech revenues, you do not see the superblockbuster in them, and owing to the potential for a cure for cancer (or headaches, or baldness, or bad sense of humor, etc.), there is a small probability that the sales in that industry may turn out to be monstrous, far larger than might be expected. On the other hand, consider negative Black Swan businesses. The track record you see is likely to overestimate the properties. Recall the 1982 blowup of banks: they appeared to the naïve observer to be more profitable than they seemed. Insurance companies are of two kinds: the regular diversifiable kind that belongs to Mediocristan (say, life insurance) and the more critical and explosive Black Swan-prone risks that are usually sold to reinsurers. According to the data, reinsurers have lost money on underwriting over the past couple of decades, but, unlike bankers, they are introspective enough to know that it actually could have been far worse, because the past twenty years did not have a big catastrophe, and all you need is one of those per century to kiss the business good-bye. Many finance academics doing "valuation" on insurance seem to have missed the point.



Part 3

THOSE GRAY SWANS OF EXTREMISTAN

It's time to deal in some depth with four final items that bear on our Black Swan.

Primo, I have said earlier that the world is moving deeper into Extremistan, that it is less and less governed by Mediocristan—in fact, this idea is more subtle than that. I will show how and present the various ideas we have about the formation of inequality. *Secondo*, I have been describing the Gaussian bell curve as a contagious and severe delusion, and it is time to get into that point in some depth. *Terso*, I will present what I call Mandelbrotian, or fractal, randomness. Remember that for an event to be a Black Swan, it does not just have to be rare, or just wild; it has to be unexpected, has to lie outside our tunnel of possibilities. You must be a sucker for it. As it happens, many rare events can yield their structure to us: it is not easy to compute their probability, but it is easy to get a *general* idea about the possibility of their occurrence. We can turn these Black Swans into Gray Swans, so to speak, reducing their surprise effect. A person aware of the possibility of such events can come to belong to the non-sucker variety.

Finally, I will present the ideas of those philosophers who focus on phony uncertainty. I organized this book in such a way that the more technical (though nonessential) sections are here; these can be skipped without any loss to the thoughtful reader, particularly [Chapters 15, 17](#), and the second half of [Chapter 16](#). I will alert the reader with footnotes. The reader less interested in the mechanics of deviations can then directly proceed to [Part 4](#).

Chapter Fourteen

FROM MEDIOCRISTAN TO EXTREMISTAN, AND BACK

I prefer Horowitz—How to fall from favor—The long tail—Get ready for some surprises—It's not just money

Let us see how an increasingly man-made planet can evolve away from mild into wild randomness. First, I describe how we get to Extremistan. Then, I will take a look at its evolution.

The World Is Unfair

Is the world that unfair? I have spent my entire life studying randomness, practicing randomness, hating randomness. The more that time passes, the worse things seem to me, the more scared I get, the more disgusted I am with Mother Nature. The more I think about my subject, the more I see evidence that the world we have in our minds is different from the one playing outside. Every morning the world appears to me more random than it did the day before, and humans seem to be even more fooled by it than they were the previous day. It is becoming unbearable. I find writing these lines painful; I find the world revolting.

Two “soft” scientists propose intuitive models for the development of this inequity: one is a mainstream economist, the other a sociologist. Both simplify a little too much. I will present their ideas because they are easy to understand, not because of the scientific quality of their insights or any consequences in their discoveries; then I will show the story as seen from the vantage point of the natural scientists.

Let me start with the economist Sherwin Rosen. In the early eighties, he wrote papers about “the economics of superstars.” In one of the papers he conveyed his sense of outrage that a basketball player could earn \$1.2 million a year, or a

television celebrity could make \$2 million. To get an idea of how this concentration is increasing—i.e., of how we are moving away from Mediocristan—consider that television celebrities and sports stars (even in Europe) get contracts today, only two decades later, worth in the hundreds of millions of dollars! The extreme is about (so far) twenty times higher than it was two decades ago!

According to Rosen, this inequality comes from a tournament effect: someone who is marginally “better” can easily win the entire pot, leaving the others with nothing. Using an argument from [Chapter 3](#), people prefer to pay \$10.99 for a recording featuring Horowitz to \$9.99 for a struggling pianist. Would you rather read Kundera for \$13.99 or some unknown author for \$1? So it looks like a tournament, where the winner grabs the whole thing—and he does not have to win by much.

But the role of luck is missing in Rosen’s beautiful argument. The problem here is the notion of “better,” this focus on skills as leading to success. Random outcomes, or an arbitrary situation, can also explain success, and provide the initial push that leads to a winner-take-all result. A person can get slightly ahead for entirely random reasons; because we like to imitate one another, we will flock to him. The world of contagion is so underestimated!

As I am writing these lines I am using a Macintosh, by Apple, after years of using Microsoft-based products. The Apple technology is vastly better, yet the inferior software won the day. How? Luck.

The Matthew Effect

More than a decade before Rosen, the sociologist of science Robert K. Merton presented his idea of the Matthew effect, by which people take from the poor to give to the rich.* He looked at the performance of scientists and showed how an initial advantage follows someone through life. Consider the following process.

Let’s say someone writes an academic paper quoting fifty people who have worked on the subject and provided background materials for his study; assume, for the sake of simplicity, that all fifty are of equal merit. Another researcher working on the exact same subject will randomly cite three of those fifty in his bibliography. Merton showed that many academics cite references without having read the original work; rather, they’ll read a paper and draw their own citations from among its sources. So a third researcher reading the second article

selects three of the previously referenced authors for *his* citations. These three authors will receive cumulatively more and more attention as their names become associated more tightly with the subject at hand. The difference between the winning three and the other members of the original cohort is mostly luck: they were initially chosen not for their greater skill, but simply for the way their names appeared in the prior bibliography. Thanks to their reputations, these successful academics will go on writing papers and their work will be easily accepted for publication. Academic success is partly (but significantly) a lottery.*

It is easy to test the effect of reputation. One way would be to find papers that were written by famous scientists, had their authors' identities changed by mistake, and got rejected. You could verify how many of these rejections were subsequently overturned after the true identities of the authors were established. Note that scholars are judged mostly on how many times their work is referenced in other people's work, and thus cliques of people who quote one another are formed (it's an "I quote you, you quote me" type of business).

Eventually, authors who are not often cited will drop out of the game by, say, going to work for the government (if they are of a gentle nature), or for the Mafia, or for a Wall Street firm (if they have a high level of hormones). Those who got a good push in the beginning of their scholarly careers will keep getting persistent cumulative advantages throughout life. It is easier for the rich to get richer, for the famous to become more famous.

In sociology, Matthew effects bear the less literary name "cumulative advantage." This theory can easily apply to companies, businessmen, actors, writers, and anyone else who benefits from past success. If you get published in *The New Yorker* because the color of your letterhead attracted the attention of the editor, who was daydreaming of daisies, the resultant reward can follow you for life. More significantly, it will follow *others* for life. Failure is also cumulative; losers are likely to also lose in the future, even if we don't take into account the mechanism of demoralization that might exacerbate it and cause additional failure.

Note that art, because of its dependence on word of mouth, is extremely prone to these cumulative-advantage effects. I mentioned clustering in [Chapter 1](#), and how journalism helps perpetuate these clusters. Our opinions about artistic merit are the result of arbitrary contagion even more than our political ideas are. One person writes a book review; another person reads it and writes a commentary

that uses the same arguments. Soon you have several hundred reviews that actually sum up in their contents to no more than two or three because there is so much overlap. For an anecdotal example read *Fire the Bastards!*, whose author, Jack Green, goes systematically through the reviews of William Gaddis's novel *The Recognitions*. Green shows clearly how book reviewers anchor on other reviews and reveals powerful mutual influence, even in their wording. This phenomenon is reminiscent of the herding of financial analysts I discussed in [Chapter 10](#).

The advent of the modern media has accelerated these cumulative advantages. The sociologist Pierre Bourdieu noted a link between the increased concentration of success and the globalization of culture and economic life. But I am not trying to play sociologist here, only show that unpredictable elements can play a role in social outcomes.

Merton's cumulative-advantage idea has a more general precursor, "preferential attachment," which, reversing the chronology (though not the logic), I will present next. Merton was interested in the social aspect of knowledge, not in the dynamics of social randomness, so his studies were derived separately from research on the dynamics of randomness in more mathematical sciences.

Lingua Franca

The theory of preferential attachment is ubiquitous in its applications: it can explain why city size is from Extremistan, why vocabulary is concentrated among a small number of words, or why bacteria populations can vary hugely in size.

The scientists J. C. Willis and G. U. Yule published a landmark paper in *Nature* in 1922 called "Some Statistics of Evolution and Geographical Distribution in Plants and Animals, and Their Significance." Willis and Yule noted the presence in biology of the so-called power laws, tractable versions of the scalable randomness that I discussed in [Chapter 3](#). These power laws (on which more technical information in the following chapters) had been noticed earlier by Vilfredo Pareto, who found that they applied to the distribution of income. Later, Yule presented a simple model showing how power laws can be generated. His point was as follows: Let's say species split in two at some constant rate, so that new species arise. The richer in species a genus is, the

richer it will tend to get, with the same logic as the Matthew effect. Note the following caveat: in Yule's model the species never die out.

During the 1940s, a Harvard linguist, George Zipf, examined the properties of language and came up with an empirical regularity now known as Zipf's law, which, of course, is not a law (and if it were, it would not be Zipf's). It is just another way to think about the process of inequality. The mechanisms he described were as follows: the more you use a word, the less effortful you will find it to use that word again, so you borrow words from your private dictionary in proportion to their past use. This explains why out of the sixty thousand main words in English, only a few hundred constitute the bulk of what is used in writings, and even fewer appear regularly in conversation. Likewise, the more people aggregate in a particular city, the more likely a stranger will be to pick that city as his destination. The big get bigger and the small stay small, or get relatively smaller.

A great illustration of preferential attachment can be seen in the mushrooming use of English as a lingua franca—though not for its intrinsic qualities, but because people need to use one single language, or stick to one as much as possible, when they are having a conversation. So whatever language appears to have the upper hand will suddenly draw people in droves; its usage will spread like an epidemic, and other languages will be rapidly dislodged. I am often amazed to listen to conversations between people from two neighboring countries, say, between a Turk and an Iranian, or a Lebanese and a Cypriot, communicating in bad English, moving their hands for emphasis, searching for these words that come out of their throats at the cost of great physical effort. Even members of the Swiss Army use English (not French) as a lingua franca (it would be fun to listen). Consider that a very small minority of Americans of northern European descent is from England; traditionally the preponderant ethnic groups are of German, Irish, Dutch, French, and other northern European extraction. Yet because all these groups now use English as their main tongue, they have to study the roots of their adoptive tongue and develop a cultural association with parts of a particular wet island, along with its history, its traditions, and its customs!

Ideas and Contagions

The same model can be used for the contagions and concentration of ideas. But

there are some restrictions on the nature of epidemics I must discuss here. Ideas do not spread without some form of structure. Recall the discussion in [Chapter 4](#) about how we come prepared to make inferences. Just as we tend to generalize some matters but not others, so there seem to be “basins of attraction” directing us to certain beliefs. Some ideas will prove contagious, but not others; some forms of superstitions will spread, but not others; some types of religious beliefs will dominate, but not others. The anthropologist, cognitive scientist, and philosopher Dan Sperber has proposed the following idea on the epidemiology of representations. What people call “memes,” ideas that spread and that compete with one another using people as carriers, are not truly like genes. Ideas spread because, alas, they have for carriers self-serving agents who are interested in them, and interested in distorting them in the replication process. You do not make a cake for the sake of merely replicating a recipe—you try to make *your* own cake, using ideas from others to improve it. We humans are not photocopiers. So contagious mental categories must be those in which we are prepared to believe, perhaps even programmed to believe. To be contagious, a mental category must agree with our nature.

NOBODY IS SAFE IN EXTREMISTAN

There is something extremely naïve about all these models of the dynamics of concentration I've presented so far, particularly the socioeconomic ones. For instance, although Merton's idea includes luck, it misses an additional layer of randomness. In all these models the winner stays a winner. Now, a loser might always remain a loser, but a winner could be unseated by someone new popping up out of nowhere. Nobody is safe.

Preferential-attachment theories are intuitively appealing, but they do not account for the possibility of being supplanted by newcomers—what every schoolchild knows as the decline of civilizations. Consider the logic of cities: How did Rome, with a population of 1.2 million in the first century A.D., end up with a population of twelve thousand in the third? How did Baltimore, once a principal American city, become a relic? And how did Philadelphia come to be overshadowed by New York?

A Brooklyn Frenchman

When I started trading foreign exchange, I befriended a fellow named Vincent who exactly resembled a Brooklyn trader, down to the mannerisms of Fat Tony, except that he spoke the French version of Brooklynese. Vincent taught me a few tricks. Among his sayings were “Trading may have princes, but nobody stays a king” and “The people you meet on the way up, you will meet again on the way down.”

There were theories when I was a child about class warfare and struggles by innocent individuals against powerful monster-corporations capable of swallowing the world. Anyone with intellectual hunger was fed these theories, which were inherited from the Marxist belief that the tools of exploitation were self-feeding, that the powerful would grow more and more powerful, furthering the unfairness of the system. But one had only to look around to see that these large corporate monsters dropped like flies. Take a cross section of the dominant corporations at any particular time; many of them will be out of business a few decades later, while firms nobody ever heard of will have popped onto the scene from some garage in California or from some college dorm.

Consider the following sobering statistic. Of the five hundred largest U.S. companies in 1957, only seventy-four were still part of that select group, the

Standard and Poor's 500, forty years later. Only a few had disappeared in mergers; the rest either shrank or went bust.

Interestingly, almost all these large corporations were located in the most capitalist country on earth, the United States. The more socialist a country's orientation, the easier it was for the large corporate monsters to stick around. Why did capitalism (and not socialism) destroy these ogres?

In other words, if you leave companies alone, they tend to get eaten up. Those in favor of economic freedom claim that beastly and greedy corporations pose no threat because competition keeps them in check. What I saw at the Wharton School convinced me that the real reason includes a large share of something else: chance.

But when people discuss chance (which they rarely do), they usually only look at their own luck. The luck *of others* counts greatly. Another corporation may luck out thanks to a blockbuster product and displace the current winners. Capitalism is, among other things, the revitalization of the world thanks to the opportunity to be lucky. Luck is the grand equalizer, because almost everyone can benefit from it. The socialist governments protected their monsters and, by doing so, killed potential newcomers in the womb.

Everything is transitory. Luck both made and unmade Carthage; it both made and unmade Rome.

I said earlier that randomness is bad, but it is not always so. Luck is far more egalitarian than even intelligence. If people were rewarded strictly according to their abilities, things would still be unfair—people don't choose their abilities. Randomness has the beneficial effect of reshuffling society's cards, knocking down the big guy.

In the arts, fads do the same job. A newcomer may benefit from a fad, as followers multiply thanks to a preferential attachment-style epidemic. Then, guess what? He too becomes history. It is quite interesting to look at the acclaimed authors of a particular era and see how many have dropped out of consciousness. It even happens in countries such as France where the government supports established reputations, just as it supports ailing large companies.

When I visit Beirut, I often spot in relatives' homes the remnants of a series of distinctively white-leather-bound "Nobel books." Some hyperactive salesman once managed to populate private libraries with these beautifully made volumes;

many people buy books for decorative purposes and want a simple selection criterion. The criterion this series offered was one book by a Nobel winner in literature every year—a simple way to build the ultimate library. The series was supposed to be updated every year, but I presume the company went out of business in the eighties. I feel a pang every time I look at these volumes: Do you hear much today about Sully Prudhomme (the first recipient), Pearl Buck (an American woman), Romain Rolland, Anatole France (the last two were the most famous French authors of their generations), St. John Perse, Roger Martin du Gard, or Frédéric Mistral?

The Long Tail

I have said that nobody is safe in Extremistan. This has a converse: nobody is threatened with complete extinction either. Our current environment allows the little guy to bide his time in the antechamber of success—as long as there is life, there is hope.

This idea was recently revived by Chris Anderson, one of a very few who get the point that the dynamics of fractal concentration has another layer of randomness. He packaged it with his idea of the “long tail,” about which in a moment. Anderson is lucky not to be a professional statistician (people who have had the misfortune of going through conventional statistical training think we live in Mediocristan). He was able to take a fresh look at the dynamics of the world.

True, the Web produces acute concentration. A large number of users visit just a few sites, such as Google, which, at the time of this writing, has total market dominance. At no time in history has a company grown so dominant so quickly—Google can service people from Nicaragua to southwestern Mongolia to the American West Coast, without having to worry about phone operators, shipping, delivery, and manufacturing. This is the ultimate winner-take-all case study.

People forget, though, that before Google, Alta Vista dominated the search-engine market. I am prepared to revise the Google metaphor by replacing it with a new name for future editions of this book.

What Anderson saw is that the Web causes something *in addition* to concentration. The Web enables the formation of a reservoir of proto-Googlees waiting in the background. It also promotes the *inverse Google*, that is, it allows people with a technical specialty to find a small, stable audience.

Recall the role of the Web in Yevgenia Krasnova's success. Thanks to the Internet, she was able to bypass conventional publishers. Her publisher with the pink glasses would not even have been in business had it not been for the Web. Let's assume that Amazon.com does not exist, and that you have written a sophisticated book. Odds are that a very small bookstore that carries only 5,000 volumes will not be interested in letting your "beautifully crafted prose" occupy premium shelf space. And the megabookstore, such as the average American Barnes & Noble, might stock 130,000 volumes, which is still not sufficient to accommodate marginal titles. So your work is stillborn.

Not so with Web vendors. A Web bookstore can carry a near-infinite number of books since it need not have them physically in inventory. Actually, nobody needs to have them physically in inventory since they can remain in digital form until they are needed in print, an emerging business called print-on-demand.

So as the author of this little book, you can sit there, bide your time, be available in search engines, and perhaps benefit from an occasional epidemic. In fact, the quality of readership has improved markedly over the past few years thanks to the availability of these more sophisticated books. This is a fertile environment for diversity.*

Plenty of people have called me to discuss the idea of the long tail, which seems to be the exact opposite of the concentration implied by scalability. The long tail implies that the small guys, collectively, should control a large segment of culture and commerce, thanks to the niches and subspecialties that can now survive thanks to the Internet. But, strangely, it can also imply a large measure of inequality: a large base of small guys and a very small number of supergiants, together representing a share of the world's culture—with some of the small guys, on occasion, rising to knock out the winners. (This is the "double tail": a large tail of the small guys, a small tail of the big guys.)

The role of the long tail is fundamental in changing the dynamics of success, destabilizing the well-seated winner, and bringing about another winner. In a snapshot this will always be Extremistan, always ruled by the concentration of type-2 randomness; but it will be an ever-changing Extremistan.

The long tail's contribution is not yet numerical; it is still confined to the Web and its small-scale online commerce. But consider how the long tail could affect the future of culture, information, and political life. It could free us from the dominant political parties, from the academic system, from the clusters of the press—anything that is currently in the hands of ossified, conceited, and self-

serving authority. The long tail will help foster cognitive diversity. One highlight of the year 2006 was to find in my mailbox a draft manuscript of a book called *Cognitive Diversity: How Our Individual Differences Produce Collective Benefits*, by Scott Page. Page examines the effects of cognitive diversity on problem solving and shows how variability in views and methods acts like an engine for tinkering. It works like evolution. By subverting the big structures we also get rid of the Platonified *one way* of doing things—in the end, the bottom-up theory-free empiricist should prevail.

In sum, the long tail is a by-product of Extremistan that makes it somewhat less unfair: the world is made no less unfair for the little guy, but it now becomes extremely unfair for the big man. Nobody is truly established. The little guy is very subversive.

Naïve Globalization

We are gliding into disorder, but not necessarily bad disorder. This implies that we will see more periods of calm and stability, with most problems concentrated into a small number of Black Swans.

Consider the nature of past wars. The twentieth century was not the deadliest (in percentage of the total population), but it brought something new: the beginning of the Extremistan warfare—a small probability of a conflict degenerating into total decimation of the human race, a conflict from which nobody is safe anywhere.

A similar effect is taking place in economic life. I spoke about globalization in [Chapter 3](#); it is here, but it is not all for the good: it creates interlocking fragility, while reducing volatility and giving the appearance of stability. In other words it creates devastating Black Swans. We have never lived before under the threat of a global collapse. Financial institutions have been merging into a smaller number of very large banks. Almost all banks are now interrelated. So the financial ecology is swelling into gigantic, incestuous, bureaucratic banks (often Gaussianized in their risk measurement)—when one falls, they all fall.* The increased concentration among banks seems to have the effect of making financial crisis less likely, but when they happen they are more global in scale and hit us very hard. We have moved from a diversified ecology of small banks, with varied lending policies, to a more homogeneous framework of firms that all resemble one another. True, we now have fewer failures, but when they occur ...

I shiver at the thought. I rephrase here: we will have fewer but more severe crises. The rarer the event, the less we know about its odds. It means that we know less and less about the possibility of a crisis.

And we have some idea how such a crisis would happen. A network is an assemblage of elements called nodes that are somehow connected to one another by a link; the world's airports constitute a network, as does the World Wide Web, as do social connections and electricity grids. There is a branch of research called "network theory" that studies the organization of such networks and the links between their nodes, with such researchers as Duncan Watts, Steven Strogatz, Albert-Laszlo Barabasi, and many more. They all understand Extremistan mathematics and the inadequacy of the Gaussian bell curve. They have uncovered the following property of networks: there is a concentration among a few nodes that serve as central connections. Networks have a natural tendency to organize themselves around an extremely concentrated architecture: a few nodes are extremely connected; others barely so. The distribution of these connections has a scalable structure of the kind we will discuss in [Chapters 15](#) and [16](#). Concentration of this kind is not limited to the Internet; it appears in social life (a small number of people are connected to others), in electricity grids, in communications networks. This seems to make networks more robust: random insults to most parts of the network will not be consequential since they are likely to hit a poorly connected spot. But it also makes networks more vulnerable to Black Swans. Just consider what would happen if there is a problem with a major node. The electricity blackout experienced in the northeastern United States during August 2003, with its consequential mayhem, is a perfect example of what could take place if one of the big banks went under today.

But banks are in a far worse situation than the Internet. The financial industry has no significant long tail! We would be far better off if there were a different ecology, in which financial institutions went bust on occasion and were rapidly replaced by new ones, thus mirroring the diversity of Internet businesses and the resilience of the Internet economy. Or if there were a long tail of government officials and civil servants coming to reinvigorate bureaucracies.

REVERSALS AWAY FROM EXTREMISTAN

There is, inevitably, a mounting tension between our society, full of concentration, and our classical idea of *aurea mediocritas*, the golden mean, so it is conceivable that efforts may be made to reverse such concentration. We live in a society of one person, one vote, where progressive taxes have been enacted precisely to weaken the winners. Indeed, the rules of society can be easily rewritten by those at the bottom of the pyramid to prevent concentration from hurting them. But it does not require voting to do so—religion could soften the problem. Consider that before Christianity, in many societies the powerful had many wives, thus preventing those at the bottom from accessing wombs, a condition that is not too different from the reproductive exclusivity of alpha males in many species. But Christianity reversed this, thanks to the one man–one woman rule. Later, Islam came to limit the number of wives to four. Judaism, which had been polygenic, became monogamous in the Middle Ages. One can say that such a strategy has been successful—the institution of tightly monogamous marriage (with no official concubine, as in the Greco-Roman days), even when practiced the “French way,” provides social stability since there is no pool of angry, sexually deprived men at the bottom fomenting a revolution just so they can have the chance to mate.

But I find the emphasis on economic inequality, at the expense of other types of inequality, extremely bothersome. Fairness is not exclusively an economic matter; it becomes less and less so when we are satisfying our basic material needs. It is pecking order that matters! The superstars will always be there. The Soviets may have flattened the economic structure, but they encouraged their own brand of *übermensch*. What is poorly understood, or denied (owing to its unsettling implications), is the absence of a role for the *average* in intellectual production. The disproportionate share of the very few in intellectual influence is even more unsettling than the unequal distribution of wealth—unsettling because, unlike the income gap, no social policy can eliminate it. Communism could conceal or compress income discrepancies, but it could not eliminate the superstar system in intellectual life.

It has even been shown, by Michael Marmot of the Whitehall Studies, that those at the top of the pecking order live longer, even when adjusting for disease. Marmot’s impressive project shows how social rank alone can affect longevity. It was calculated that actors who win an Oscar tend to live on average about five

years longer than their peers who don't. People live longer in societies that have flatter social gradients. Winners kill their peers as those in a steep social gradient live shorter lives, regardless of their economic condition.

I do not know how to remedy this (except through religious beliefs). Is insurance against your peers' demoralizing success possible? Should the Nobel Prize be banned? Granted the Nobel medal in economics has not been good for society or knowledge, but even those rewarded for *real* contributions in medicine and physics too rapidly displace others from our consciousness, and steal longevity away from them. Extremistan is here to stay, so we have to live with it, and find the tricks that make it more palatable.

-
- * These scalable laws were already discussed in the scriptures: "For onto everyone that hath shall be given, and he shall have abundance; but from him that hath not shall be taken away even that which he hath." Matthew (Matthew 25:29, King James Version).
 - * Much of the perception of the importance of precocity in the career of researchers can be owed to the misunderstanding of the perverse role of this effect, especially when reinforced by bias. Enough counterexamples, even in fields like mathematics meant to be purely a "young man's game," illustrate the age fallacy: simply, it is necessary to be successful early, and even very early at that.
 - * The Web's bottom-up feature is also making book reviewers more accountable. While writers were helpless and vulnerable to the arbitrariness of book reviews, which can distort their messages and, thanks to the confirmation bias, expose small irrelevant weak points in their text, they now have a much stronger hand. In place of the moaning letter to the editor, they can simply post their review of a review on the Web. If attacked ad hominem, they can reply ad hominem and go directly after the credibility of the reviewer, making sure that their statement shows rapidly in an Internet search or on Wikipedia, the bottom-up encyclopedia.
 - * As if we did not have enough problems, banks are now more vulnerable to the Black Swan and the ludic fallacy than ever before with "scientists" among their staff taking care of exposures. The giant firm J. P. Morgan put the entire world at risk by introducing in the nineties RiskMetrics, a phony method aiming at managing people's risks, causing the generalized use of the ludic fallacy, and bringing Dr. Johns into power in place of the skeptical Fat Tonys. (A related method called "Value-at-Risk," which relies on the quantitative measurement of risk, has been spreading.) Likewise, the government-sponsored institution Fanny Mae, when I look at their risks, seems to be sitting on a barrel of dynamite, vulnerable to the slightest hiccup. But not to worry: their large staff of scientists deemed these events "unlikely."

Chapter Fifteen

THE BELL CURVE, THAT GREAT INTELLECTUAL FRAUD*

Not worth a pastis—Quételet's error—The average man is a monster—Let's deify it—Yes or no—Not so literary an experiment

Forget everything you heard in college statistics or probability theory. If you never took such a class, even better. Let us start from the very beginning.

THE GAUSSIAN AND THE MANDELBROTIAN

I was transiting through the Frankfurt airport in December 2001, on my way from Oslo to Zurich.

I had time to kill at the airport and it was a great opportunity for me to buy dark European chocolate, especially since I have managed to successfully convince myself that airport calories don't count. The cashier handed me, among other things, a ten deutschmark bill, an (illegal) scan of which can be seen on the next page. The deutschmark banknotes were going to be put out of circulation in a matter of days, since Europe was switching to the euro. I kept it as a valedictory. Before the arrival of the euro, Europe had plenty of national currencies, which was good for printers, money changers, and of course currency traders like this (more or less) humble author. As I was eating my dark European chocolate and wistfully looking at the bill, I almost choked. I suddenly noticed, for the first time, that there was something curious about it. The bill bore the portrait of Carl Friedrich Gauss and a picture of his Gaussian bell curve.



The last ten deutschmark bill, representing Gauss and, to his right, the bell curve of Mediocristan.

The striking irony here is that the last possible object that can be linked to the German currency is precisely such a curve: the reichsmark (as the currency was previously called) went from four per dollar to *four trillion* per dollar in the space of a few years during the 1920s, an outcome that tells you that the bell curve is meaningless as a description of the randomness in currency fluctuations. All you need to reject the bell curve is for such a movement to occur once, and only once—just consider the consequences. Yet there was the bell curve, and next to it Herr Professor Doktor Gauss, unprepossessing, a little stern, certainly

not someone I'd want to spend time with lounging on a terrace, drinking pastis, and holding a conversation without a subject.

Shockingly, the bell curve is used as a risk-measurement tool by those regulators and central bankers who wear dark suits and talk in a boring way about currencies.

The Increase in the Decrease

The main point of the Gaussian, as I've said, is that most observations hover around the mediocre, the average; the odds of a deviation decline faster and faster (exponentially) as you move away from the average. If you must have only one single piece of information, this is the one: the dramatic increase in the speed of decline in the odds as you move away from the center, or the average. Look at the list below for an illustration of this. I am taking an example of a Gaussian quantity, such as height, and simplifying it a bit to make it more illustrative. Assume that the average height (men and women) is 1.67 meters, or 5 feet 7 inches. Consider what I call a *unit of deviation* here as 10 centimeters. Let us look at increments above 1.67 meters and consider the odds of someone being that tall.*

10 centimeters taller than the average (i.e., taller than 1.77 m, or 5 feet 10): 1 in 6.3

20 centimeters taller than the average (i.e., taller than 1.87 m, or 6 feet 2): 1 in 44

30 centimeters taller than the average (i.e., taller than 1.97 m, or 6 feet 6): 1 in 740

40 centimeters taller than the average (i.e., taller than 2.07 m, or 6 feet 9): 1 in 32,000

50 centimeters taller than the average (i.e., taller than 2.17 m, or 7 feet 1): 1 in 3,500,000

60 centimeters taller than the average (i.e., taller than 2.27 m, or 7 feet 5): 1 in 1,000,000,000

70 centimeters taller than the average (i.e., taller than 2.37 m, or 7 feet 9): 1 in 780,000,000,000

80 centimeters taller than the average (i.e., taller than 2.47 m, or 8 feet 1): 1 in 630,000,000,000,000

Higher than €8 million: 1 in 4,000
Higher than €16 million: 1 in 16,000
Higher than €32 million: 1 in 64,000
Higher than €320 million: 1 in 6,400,000

The speed of the decrease here remains constant (or does not decline)! When you double the amount of money you cut the incidence by a factor of four, no matter the level, whether you are at €8 million or €16 million. This, in a nutshell, illustrates the difference between Mediocristan and Extremistan.

Recall the comparison between the scalable and the nonscalable in [Chapter 3](#). Scalability means that there is no headwind to slow you down.

Of course, Mandelbrotian Extremistan can take many shapes. Consider wealth in an extremely concentrated version of Extremistan; there, if you double the wealth, you halve the incidence. The result is quantitatively different from the above example, but it obeys the same logic.

Fractal Wealth Distribution with Large Inequalities

People with a net worth higher than €1 million: 1 in 63
Higher than €2 million: 1 in 125
Higher than €4 million: 1 in 250
Higher than €8 million: 1 in 500
Higher than €16 million: 1 in 1,000
Higher than €32 million: 1 in 2,000
Higher than €320 million: 1 in 20,000
Higher than €640 million: 1 in 40,000

If wealth were Gaussian, we would observe the following divergence away from €1 million.

Wealth Distribution Assuming a Gaussian Law

People with a net worth higher than €1 million: 1 in 63

Higher than €2 million: 1 in 127,000

Higher than €3 million: 1 in 14,000,000,000

Higher than €4 million: 1 in 886,000,000,000,000,000

Higher than €8 million: 1 in
16,000,000,000,000,000,000,000,000,000,000

Higher than €16 million: 1 in ... *none of my computers is capable of handling the computation.*

What I want to show with these lists is the qualitative difference in the paradigms. As I have said, the second paradigm is scalable; it has no headwind. Note that another term for the scalable is power laws.

Just knowing that we are in a power-law environment does not tell us much. Why? Because we have to measure the coefficients in real life, which is much harder than with a Gaussian framework. Only the Gaussian yields its properties rather rapidly. The method I propose is a general way of viewing the world rather than a precise solution.

What to Remember

Remember this: the Gaussian–bell curve variations face a headwind that makes probabilities drop at a faster and faster rate as you move away from the mean, while “scalables,” or Mandelbrotian variations, do not have such a restriction. That’s pretty much most of what you need to know.*

Inequality

Let us look more closely at the nature of inequality. In the Gaussian framework, inequality decreases as the deviations get larger—caused by the increase in the rate of decrease. Not so with the scalable: inequality stays the same throughout. The inequality among the superrich is the same as the inequality among the simply rich—it does not slow down.†

Consider this effect. Take a random sample of any two people from the U.S. population who jointly earn \$1 million per annum. What is the most likely breakdown of their respective incomes? In Mediocristan, the most likely combination is half a million each. In Extremistan, it would be \$50,000 and \$950,000.

The situation is even more lopsided with book sales. If I told you that two authors sold a total of a million copies of their books, the most likely combination is 993,000 copies sold for one and 7,000 for the other. This is far more likely than that the books each sold 500,000 copies. *For any large total, the breakdown will be more and more asymmetric.*

Why is this so? The height problem provides a comparison. If I told you that the total height of two people is fourteen feet, you would identify the most likely breakdown as seven feet each, not two feet and twelve feet; not even eight feet and six feet! Persons taller than eight feet are so rare that such a combination would be impossible.

Extremistan and the 80/20 Rule

Have you ever heard of the 80/20 rule? It is the common signature of a power law—actually it is how it all started, when Vilfredo Pareto made the observation that 80 percent of the land in Italy was owned by 20 percent of the people. Some use the rule to imply that 80 percent of the work is done by 20 percent of the people. Or that 80 percent worth of effort contributes to only 20 percent of results, and vice versa.

As far as axioms go, this one wasn't phrased to impress you the most: it could easily be called the 50/01 rule, that is, 50 percent of the work comes from 1 percent of the workers. This formulation makes the world look even more unfair, yet the two formulae are exactly the same. How? Well, if there is inequality, then those who constitute the 20 percent in the 80/20 rule also contribute unequally—only a few of them deliver the lion's share of the results. This trickles down to about one in a hundred contributing a little more than half the total.

The 80/20 rule is only metaphorical; it is not a rule, even less a rigid law. In the U.S. book business, the proportions are more like 97/20 (i.e., 97 percent of book sales are made by 20 percent of the authors); it's even worse if you focus on literary nonfiction (twenty books of close to eight thousand represent half the

sales).

Note here that it is not all uncertainty. In some situations you may have a concentration, of the 80/20 type, with very predictable and tractable properties, which enables clear decision making, because you can identify *beforehand* where the meaningful 20 percent are. These situations are very easy to control. For instance, Malcolm Gladwell wrote in an article in *The New Yorker* that most abuse of prisoners is attributable to a very small number of vicious guards. Filter those guards out and your rate of prisoner abuse drops dramatically. (In publishing, on the other hand, you do not know beforehand which book will bring home the bacon. The same with wars, as you do not know beforehand which conflict will kill a portion of the planet's residents.)

Grass and Trees

I'll summarize here and repeat the arguments previously made throughout the book. Measures of uncertainty that are based on the bell curve simply disregard the possibility, and the impact, of sharp jumps or discontinuities and are, therefore, inapplicable in Extremistan. Using them is like focusing on the grass and missing out on the (gigantic) trees. Although unpredictable large deviations are rare, they cannot be dismissed as outliers because, cumulatively, their impact is so dramatic.

The traditional Gaussian way of looking at the world begins by focusing on the ordinary, and then deals with exceptions or so-called outliers as ancillaries. But there is a second way, which takes the exceptional as a starting point and treats the ordinary as subordinate.

I have emphasized that there are two varieties of randomness, qualitatively different, like air and water. One does not care about extremes; the other is severely impacted by them. One does not generate Black Swans; the other does. We cannot use the same techniques to discuss a gas as we would use with a liquid. And if we could, we wouldn't call the approach "an approximation." A gas does not "approximate" a liquid.

We can make good use of the Gaussian approach in variables for which there is a rational reason for the largest not to be too far away from the average. If there is gravity pulling numbers down, or if there are physical limitations preventing very large observations, we end up in Mediocristan. If there are strong forces of equilibrium bringing things back rather rapidly after conditions

diverge from equilibrium, then again you can use the Gaussian approach. Otherwise, fuhgedaboudit. This is why much of economics is based on the notion of equilibrium: among other benefits, it allows you to treat economic phenomena as Gaussian.

Note that I am not telling you that the Mediocristan type of randomness does not allow for *some* extremes. But it tells you that they are so rare that they do not play a significant role in the total. The effect of such extremes is pitifully small and decreases as your population gets larger.

To be a little bit more technical here, if you have an assortment of giants and dwarfs, that is, observations several orders of magnitude apart, you could still be in Mediocristan. How? Assume you have a sample of one thousand people, with a large spectrum running from the dwarf to the giant. You are likely to see many giants in your sample, not a rare occasional one. Your average will not be impacted by the occasional additional giant because some of these giants are expected to be part of your sample, and your average is likely to be high. In other words, the largest observation cannot be too far away from the average. The average will always contain both kinds, giants and dwarves, so that neither should be too rare—unless you get a megagiant or a microdwarf on very rare occasion. This would be Mediocristan with a large unit of deviation.

Note once again the following principle: the rarer the event, the higher the error in our estimation of its probability—even when using the Gaussian.

Let me show you how the Gaussian bell curve sucks randomness out of life—which is why it is popular. We like it because it allows certainties! How? Through averaging, as I will discuss next.

How Coffee Drinking Can Be Safe

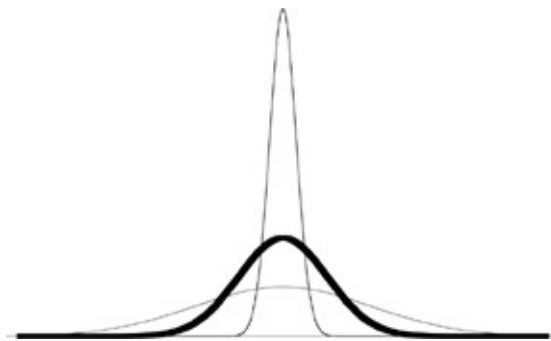
Recall from the Mediocristan discussion in [Chapter 3](#) that no single observation will impact your total. This property will be more and more significant as your population increases in size. The averages will become more and more stable, to the point where all samples will look alike.

I've had plenty of cups of coffee in my life (it's my principal addiction). I have never seen a cup jump two feet from my desk, nor has coffee spilled spontaneously on this manuscript without intervention (even in Russia). Indeed, it will take more than a mild coffee addiction to witness such an event; it would require more lifetimes than is perhaps conceivable—the odds are so small, one

in so many zeroes, that it would be impossible for me to write them down in my free time.

Yet physical reality makes it possible for my coffee cup to jump—very unlikely, but possible. Particles jump around all the time. How come the coffee cup, itself composed of jumping particles, does not? The reason is, simply, that for the cup to jump would require that all of the particles jump in the *same* direction, and do so in lockstep several times in a row (with a compensating move of the table in the opposite direction). All several trillion particles in my coffee cup are not going to jump in the same direction; this is not going to happen in the lifetime of this universe. So I can safely put the coffee cup on the edge of my writing table and worry about more serious sources of uncertainty.

FIGURE 7: How the Law of Large Numbers Works



In Mediocristan, as your sample size increases, the observed average will present itself with less and less dispersion—as you can see, the distribution will be narrower and narrower. This, in a nutshell, is how everything in statistical theory works (or is supposed to work). Uncertainty in Mediocristan vanishes under averaging. This illustrates the hackneyed “law of large numbers.”

The safety of my coffee cup illustrates how the randomness of the Gaussian is tamable by averaging. If my cup were one large particle, or acted as one, then its jumping would be a problem. But my cup is the sum of trillions of very small particles.

Casino operators understand this well, which is why they never (if they do things right) lose money. They simply do not let one gambler make a massive bet, instead preferring to have plenty of gamblers make series of bets of limited size. Gamblers may bet a total of \$20 million, but you needn’t worry about the casino’s health: the bets run, say, \$20 on average; the casino caps the bets at a maximum that will allow the casino owners to sleep at night. So the variations in

the casino's returns are going to be ridiculously small, no matter the total gambling activity. You will not see anyone leaving the casino with \$1 billion—in the lifetime of this universe.

The above is an application of the supreme law of Mediocristan: when you have plenty of gamblers, no single gambler will impact the total more than minutely.

The consequence of this is that variations around the average of the Gaussian, also called “errors,” are not truly worrisome. They are small and they wash out. They are domesticated fluctuations around the mean.

Love of Certainties

If you ever took a (dull) statistics class in college, did not understand much of what the professor was excited about, and wondered what “standard deviation” meant, there is nothing to worry about. The notion of standard deviation is meaningless outside of Mediocristan. Clearly it would have been more beneficial, and certainly more entertaining, to have taken classes in the neurobiology of aesthetics or postcolonial African dance, and this is easy to see empirically.

Standard deviations do not exist outside the Gaussian, or if they do exist they do not matter and do not explain much. But it gets worse. The Gaussian family (which includes various friends and relatives, such as the Poisson law) are the only class of distributions that the standard deviation (and the average) is sufficient to describe. You need nothing else. The bell curve satisfies the reductionism of the deluded.

There are other notions that have little or no significance outside of the Gaussian: *correlation* and, worse, *regression*. Yet they are deeply ingrained in our methods; it is hard to have a business conversation without hearing the word *correlation*.

To see how meaningless correlation can be outside of Mediocristan, take a historical series involving two variables that are patently from Extremistan, such as the bond and the stock markets, or two securities prices, or two variables like, say, changes in book sales of children's books in the United States, and fertilizer production in China; or real-estate prices in New York City and returns of the Mongolian stock market. Measure correlation between the pairs of variables in different subperiods, say, for 1994, 1995, 1996, *etc.* The correlation measure will

be likely to exhibit severe instability; it will depend on the period for which it was computed. Yet people talk about correlation as if it were something real, making it tangible, investing it with a physical property, reifying it.

The same illusion of concreteness affects what we call “standard” deviations. Take any series of historical prices or values. Break it up into subsegments and measure its “standard” deviation. Surprised? Every sample will yield a different “standard” deviation. Then why do people talk about standard deviations? Go figure.

Note here that, as with the narrative fallacy, when you look at past data and compute one single correlation or standard deviation, you do not notice such instability.

How to Cause Catastrophes

If you use the term *statistically significant*, beware of the illusions of certainties. Odds are that someone has looked at his observation errors and assumed that they were Gaussian, which necessitates a Gaussian context, namely, Mediocristan, for it to be acceptable.

To show how endemic the problem of misusing the Gaussian is, and how dangerous it can be, consider a (dull) book called *Catastrophe* by Judge Richard Posner, a prolific writer. Posner bemoans civil servants’ misunderstandings of randomness and recommends, among other things, that government policy makers learn statistics ... from economists. Judge Posner appears to be trying to foment catastrophes. Yet, in spite of being one of those people who should spend more time reading and less time writing, he can be an insightful, deep, and original thinker; like many people, he just isn’t aware of the distinction between Mediocristan and Extremistan, and he believes that statistics is a “science,” never a fraud. If you run into him, please make him aware of these things.

QUÉTELET'S AVERAGE MONSTER

This monstrosity called the Gaussian bell curve is not Gauss's doing. Although he worked on it, he was a mathematician dealing with a theoretical point, not making claims about the structure of reality like statistical-minded scientists. G. H. Hardy wrote in "A Mathematician's Apology":

The "real" mathematics of the "real" mathematicians, the mathematics of Fermat and Euler and Gauss and Abel and Riemann, is almost wholly "useless" (and this is as true of "applied" as of "pure" mathematics).

As I mentioned earlier, the bell curve was mainly the concoction of a gambler, Abraham de Moivre (1667–1754), a French Calvinist refugee who spent much of his life in London, though speaking heavily accented English. But it is Quételet, not Gauss, who counts as one of the most destructive fellows in the history of thought, as we will see next.

Adolphe Quételet (1796–1874) came up with the notion of a physically average human, *l'homme moyen*. There was nothing *moyen* about Quételet, "a man of great creative passions, a creative man full of energy." He wrote poetry and even coauthored an opera. The basic problem with Quételet was that he was a mathematician, not an empirical scientist, but he did not know it. He found harmony in the bell curve.

The problem exists at two levels. *Primo*, Quételet had a normative idea, to make the world fit his average, in the sense that the average, to him, was the "normal." It would be wonderful to be able to ignore the contribution of the unusual, the "nonnormal," the Black Swan, to the total. But let us leave that dream for utopia.

Secondo, there was a serious associated empirical problem. Quételet saw bell curves everywhere. He was blinded by bell curves and, I have learned, again, once you get a bell curve in your head it is hard to get it out. Later, Frank Ysidro Edgeworth would refer to Quételesmus as the grave mistake of seeing bell curves everywhere.

Golden Mediocrity

Quételet provided a much needed product for the ideological appetites of his day. As he lived between 1796 and 1874, so consider the roster of his

contemporaries: Saint-Simon (1760–1825), Pierre-Joseph Proudhon (1809–1865), and Karl Marx (1818–1883), each the source of a different version of socialism. Everyone in this post-Enlightenment moment was longing for the *aurea mediocritas*, the golden mean: in wealth, height, weight, and so on. This longing contains some element of wishful thinking mixed with a great deal of harmony and ... Platoncity.

I always remember my father's injunction that *in medio stat virtus*, "virtue lies in moderation." Well, for a long time that was the ideal; mediocrity, in that sense, was even deemed golden. All-embracing mediocrity.

But Quételet took the idea to a different level. Collecting statistics, he started creating standards of "means." Chest size, height, the weight of babies at birth, very little escaped his *standards*. Deviations from the norm, he found, became exponentially more rare as the magnitude of the deviation increased. Then, having conceived of this idea of the physical characteristics of *l'homme moyen*, Monsieur Quételet switched to social matters. *L'homme moyen* had his habits, his consumption, his methods.

Through his construct of *l'homme moyen physique* and *l'homme moyen moral*, the physically and morally average man, Quételet created a range of deviance from the average that positions all people either to the left or right of center and, truly, punishes those who find themselves occupying the extreme left or right of the statistical bell curve. They became *abnormal*. How this inspired Marx, who cites Quételet regarding this concept of an average or normal man, is obvious: "Societal deviations in terms of the distribution of wealth for example, must be minimized," he wrote in *Das Kapital*.

One has to give some credit to the scientific establishment of Quételet's day. They did not buy his arguments at once. The philosopher/mathematician/economist Augustin Cournot, for starters, did not believe that one could establish a standard human on purely quantitative grounds. Such a standard would be dependent on the attribute under consideration. A measurement in one province may differ from that in another province. Which one should be the standard? *L'homme moyen* would be a monster, said Cournot. I will explain his point as follows.

Assuming there is something desirable in being an average man, he must have an unspecified specialty in which he would be more gifted than other people—he cannot be average in everything. A pianist would be better on average at playing the piano, but worse than the norm at, say, horseback riding. A draftsman would

have better drafting skills, and so on. *The notion of a man deemed average is different from that of a man who is average in everything he does.* In fact, an exactly average human would have to be half male and half female. Quételet completely missed that point.

God's Error

A much more worrisome aspect of the discussion is that in Quételet's day, the name of the Gaussian distribution was *la loi des erreurs*, the law of errors, since one of its earliest applications was the distribution of errors in astronomic measurements. Are you as worried as I am? Divergence from the mean (here the median as well) was treated precisely as an error! No wonder Marx fell for Quételet's ideas.

This concept took off very quickly. The *ought* was confused with the *is*, and this with the imprimatur of science. The notion of the average man is steeped in the culture attending the birth of the European middle class, the nascent post-Napoleonic shopkeeper's culture, chary of excessive wealth and intellectual brilliance. In fact, the dream of a society with compressed outcomes is assumed to correspond to the aspirations of a rational human being facing a genetic lottery. If you had to pick a society to be born into for your next life, but could not know which outcome awaited you, it is assumed you would probably take no gamble; you would like to belong to a society without divergent outcomes.

One entertaining effect of the glorification of mediocrity was the creation of a political party in France called Poujadism, composed initially of a grocery-store movement. It was the warm huddling together of the semi-favored hoping to see the rest of the universe compress itself into their rank—a case of non-proletarian revolution. It had a grocery-store-owner mentality, down to the employment of the mathematical tools. Did Gauss provide the mathematics for the shopkeepers?

Poincaré to the Rescue

Poincaré himself was quite suspicious of the Gaussian. I suspect that he felt queasy when it and similar approaches to modeling uncertainty were presented to him. Just consider that the Gaussian was initially meant to measure astronomic errors, and that Poincaré's ideas of modeling celestial mechanics were fraught with a sense of deeper uncertainty.

Poincaré wrote that one of his friends, an unnamed “eminent physicist,” complained to him that physicists tended to use the Gaussian curve because they thought mathematicians believed it a mathematical necessity; mathematicians used it because they believed that physicists found it to be an empirical fact.

Eliminating Unfair Influence

Let me state here that, except for the grocery-store mentality, I truly believe in the value of middleness and mediocrity—what humanist does not want to minimize the discrepancy between humans? Nothing is more repugnant than the inconsiderate ideal of the *Übermensch*! My true problem is epistemological. Reality is not Mediocristan, so we should learn to live with it.

“The Greeks Would Have Deified It”

The list of people walking around with the bell curve stuck in their heads, thanks to its Platonic purity, is incredibly long.

Sir Francis Galton, Charles Darwin’s first cousin and Erasmus Darwin’s grandson, was perhaps, along with his cousin, one of the last independent gentlemen scientists—a category that also included Lord Cavendish, Lord Kelvin, Ludwig Wittgenstein (in his own way), and to some extent, our überphilosopher Bertrand Russell. Although John Maynard Keynes was not quite in that category, his thinking epitomizes it. Galton lived in the Victorian era when heirs and persons of leisure could, among other choices, such as horseback riding or hunting, become thinkers, scientists, or (for those less gifted) politicians. There is much to be wistful about in that era: the authenticity of someone doing science for science’s sake, without direct career motivations.

Unfortunately, doing science for the love of knowledge does not necessarily mean you will head in the right direction. Upon encountering and absorbing the “normal” distribution, Galton fell in love with it. He was said to have exclaimed that if the Greeks had known about it, they would have deified it. His enthusiasm may have contributed to the prevalence of the use of the Gaussian.

Galton was blessed with no mathematical baggage, but he had a rare obsession with measurement. He did not know about the law of large numbers, but rediscovered it from the data itself. He built the quincunx, a pinball machine that shows the development of the bell curve—on which, more in a few

paragraphs. True, Galton applied the bell curve to areas like genetics and heredity, in which its use was justified. But his enthusiasm helped thrust nascent statistical methods into social issues.

“Yes/No” Only Please

Let me discuss here the extent of the damage. If you’re dealing with qualitative inference, such as in psychology or medicine, looking for yes/no answers to which magnitudes don’t apply, then you can assume you’re in Mediocristan without serious problems. The impact of the improbable cannot be too large. You have cancer or you don’t, you are pregnant or you are not, et cetera. Degrees of deadness or pregnancy are not relevant (unless you are dealing with epidemics). But if you are dealing with aggregates, where magnitudes do matter, such as income, your wealth, return on a portfolio, or book sales, then you will have a problem and get the wrong distribution if you use the Gaussian, as it does not belong there. One single number can disrupt all your averages; one single loss can eradicate a century of profits. You can no longer say “this is an exception.” The statement “Well, I can lose money” is not informational unless you can attach a quantity to that loss. You can lose all your net worth or you can lose a fraction of your daily income; there is a difference.

This explains why empirical psychology and its insights on human nature, which I presented in the earlier parts of this book, are robust to the mistake of using the bell curve; they are also lucky, since most of their variables allow for the application of conventional Gaussian statistics. When measuring how many people in a sample have a bias, or make a mistake, these studies generally elicit a yes/no type of result. No single observation, by itself, can disrupt their overall findings.

I will next proceed to a sui generis presentation of the bell-curve idea from the ground up.

A (LITERARY) THOUGHT EXPERIMENT ON WHERE THE BELL CURVE COMES FROM

Consider a pinball machine like the one shown in [Figure 8](#). Launch 32 balls, assuming a well-balanced board so that the ball has equal odds of falling right or left at any juncture when hitting a pin. Your expected outcome is that many balls will land in the center columns and that the number of balls will decrease as you move to the columns away from the center.

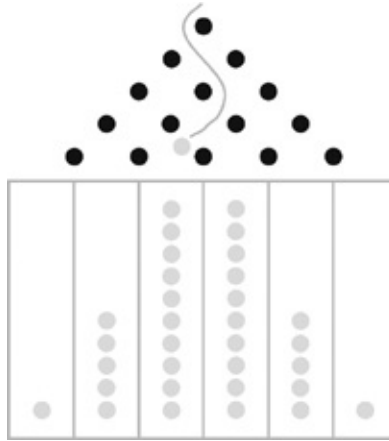
Next, consider a gedanken, a thought experiment. A man flips a coin and after each toss he takes a step to the left or a step to the right, depending on whether the coin came up heads or tails. This is called the random walk, but it does not necessarily concern itself with walking. You could identically say that instead of taking a step to the left or to the right, you would win or lose \$1 at every turn, and you will keep track of the cumulative amount that you have in your pocket.

Assume that I set you up in a (legal) wager where the odds are neither in your favor nor against you. Flip a coin. Heads, you make \$1, tails, you lose \$1.

At the first flip, you will either win or lose.

At the second flip, the number of possible outcomes doubles. Case one: win, win. Case two: win, lose. Case three: lose, win. Case four: lose, lose. Each of these cases has equivalent odds, the combination of a single win and a single loss has an incidence twice as high because cases two and three, win-lose and lose-win, amount to the same outcome. And that is the key for the Gaussian. So much in the middle washes out—and we will see that there is a lot in the middle. So, if you are playing for \$1 a round, after two rounds you have a 25 percent chance of making or losing \$2, but a 50 percent chance of breaking even.

FIGURE 8: THE QUINCUNX (SIMPLIFIED)—A PINBALL MACHINE



Drop balls that, at every pin, randomly fall right or left. Above is the most probable scenario, which greatly resembles the bell curve (a.k.a. Gaussian distribution). *Courtesy of Alexander Taleb.*

Let us do another round. The third flip again doubles the number of cases, so we face eight possible outcomes. Case 1 (it was win, win in the second flip) branches out into win, win, win and win, win, lose. We add a win or lose to the end of each of the previous results. Case 2 branches out into win, lose, win and win, lose, lose. Case 3 branches out into lose, win, win and lose, win, lose. Case 4 branches out into lose, lose, win and lose, lose, lose.

We now have eight cases, all equally likely. Note that again you can group the middling outcomes where a win cancels out a loss. (In Galton's quincunx, situations where the ball falls left and then falls right, or vice versa, dominate so you end up with plenty in the middle.) The net, or cumulative, is the following: 1) *three wins*; 2) two wins, one loss, net *one win*; 3) two wins, one loss, net *one win*; 4) one win, two losses, net *one loss*; 5) two wins, one loss, net *one win*; 6) two losses, one win, net *one loss*; 7) two losses, one win, net *one loss*; and, finally, 8) *three losses*.

Out of the eight cases, the case of three wins occurs once. The case of three losses occurs once. The case of one net loss (one win, two losses) occurs three times. The case of one net win (one loss, two wins) occurs three times.

Play one more round, the fourth. There will be sixteen equally likely outcomes. You will have one case of four wins, one case of four losses, four cases of two wins, four cases of two losses, and six break-even cases.

The quincunx (its name is derived from the Latin for five) in the pinball example shows the fifth round, with thirty-two possibilities, easy to track. Such was the concept behind the quincunx used by Francis Galton. Galton was both

insufficiently lazy and a bit too innocent of mathematics; instead of building the contraption, he could have worked with simpler algebra, or perhaps undertaken a thought experiment like this one.

Let's keep playing. Continue until you have forty flips. You can perform them in minutes, but we will need a calculator to work out the number of outcomes, which are taxing to our simple thought method. You will have about 1,099,511,627,776 possible combinations—more than one thousand billion. Don't bother doing the calculation manually, it is two multiplied by itself forty times, since each branch doubles at every juncture. (Recall that we added a win and a lose at the end of the alternatives of the third round to go to the fourth round, thus doubling the number of alternatives.) Of these combinations, only one will be up forty, and only one will be down forty. The rest will hover around the middle, here zero.

We can already see that in this type of randomness extremes are exceedingly rare. One in 1,099,511,627,776 is up forty out of forty tosses. If you perform the exercise of forty flips once per hour, the odds of getting 40 ups in a row are so small that it would take quite a bit of forty-flip trials to see it. Assuming you take a few breaks to eat, argue with your friends and roommates, have a beer, and sleep, you can expect to wait close to four million lifetimes to get a 40-up outcome (or a 40-down outcome) just once. And consider the following. Assume you play one additional round, for a total of 41; to get 41 straight heads would take eight million lifetimes! Going from 40 to 41 halves the odds. This is a key attribute of the nonscalable framework to analyzing randomness: extreme deviations decrease at an increasing rate. You can expect to toss 50 heads in a row once in four billion lifetimes!

FIGURE 9: NUMBERS OF WINS TOSSED



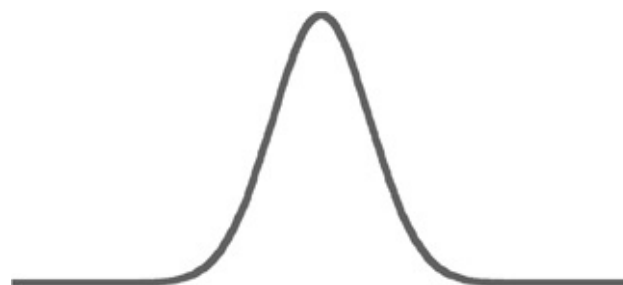
Result of forty tosses. We see the proto-bell curve emerging.

We are not yet fully in a Gaussian bell curve, but we are getting dangerously close. This is still proto-Gaussian, but you can see the gist. (Actually, you will never encounter a Gaussian in its purity since it is a Platonic form—you just get closer but cannot attain it.) However, as you can see in [Figure 9](#), the familiar bell shape is starting to emerge.

[illegible]

Continue the process for a while. We go from 40 tosses for \$1 each to 4,000 tosses for 10 cents, to 400,000 tosses for 1 cent, getting close and closer to a Gaussian. [Figure 10](#) shows results spread between -40 and 40 , namely eighty plot points. The next one would bring that up to 8,000 points.

FIGURE 10: A MORE ABSTRACT VERSION: PLATO'S CURVE



An infinite number of tosses.

Let's keep going. We can flip 4,000 times staking a tenth of a penny. How about 400,000 times at 1/1000 of a penny? As a Platonic form, the pure Gaussian curve is principally what happens when he have an infinity of tosses per round.

with each bet infinitesimally small. Do not bother trying to visualize the results, or even make sense out of them. We can no longer talk about an “infinitesimal” bet size (since we have an infinity of these, and we are in what mathematicians call a continuous framework). The good news is that there is a substitute.

We have moved from a simple bet to something completely abstract. We have moved from observations into the realm of mathematics. In mathematics things have a purity to them.

Now, something completely abstract is not supposed to exist, so *please do not even make an attempt to understand Figure 10*. Just be aware of its use. Think of it as a thermometer: you are not supposed to understand what the temperature *means* in order to talk about it. You just need to know the correspondence between temperature and comfort (or some other empirical consideration). Sixty degrees corresponds to pleasant weather; ten below is not something to look forward to. You don’t necessarily care about the actual speed of the collisions among particles that more technically explains temperature. Degrees are, in a way, a means for your mind to translate some external phenomena into a number. Likewise, the Gaussian bell curve is set so that 68.2 percent of the observations fall between minus one and plus one standard deviations away from the average. I repeat: do not even try to understand whether *standard deviation* is *average deviation*—it is not, and a large (too large) number of people using the word *standard deviation* do not understand this point. Standard deviation is just a number that you scale things to, a matter of mere correspondence *if phenomena were Gaussian*.

These standard deviations are often nicknamed “sigma.” People also talk about “variance” (same thing: variance is the square of the sigma, i.e., of the standard deviation).

Note the symmetry in the curve. You get the same results whether the sigma is positive or negative. The odds of falling below -4 sigmas are the same as those of exceeding 4 sigmas, here 1 in 32,000 times.

As the reader can see, the main point of the Gaussian bell curve is, as I have been saying, that most observations hover around the mediocre, the mean, while the odds of a deviation decline faster and faster (exponentially) as you move away from the mean. If you need to retain one single piece of information, just remember this dramatic speed of decrease in the odds as you move away from the average. Outliers are increasingly unlikely. You can safely ignore them.

This property also generates the supreme law of Mediocristan: given the

paucity of large deviations, their contribution to the total will be vanishingly small.

In the height example earlier in this chapter, I used units of deviations of ten centimeters, showing how the incidence declined as the height increased. These were one sigma deviations; the height table also provides an example of the operation of “scaling to a sigma” by using the sigma as a unit of measurement.

Those Comforting Assumptions

Note the central assumptions we made in the coin-flip game that led to the proto-Gaussian, or mild randomness.

First central assumption: the flips are independent of one another. The coin has no memory. The fact that you got heads or tails on the previous flip does not change the odds of your getting heads or tails on the next one. You do not become a “better” coin flipper over time. If you introduce memory, or skills in flipping, the entire Gaussian business becomes shaky.

Recall our discussions in [Chapter 14](#) on preferential attachment and cumulative advantage. Both theories assert that winning today makes you more likely to win in the future. Therefore, probabilities are dependent on history, and the first central assumption leading to the Gaussian bell curve fails in reality. In games, of course, past winnings are not supposed to translate into an increased probability of future gains—but not so in real life, which is why I worry about teaching probability from games. But when winning leads to more winning, you are far more likely to see forty wins in a row than with a proto-Gaussian.

Second central assumption: no “wild” jump. The step size in the building block of the basic random walk is always known, namely one step. There is no uncertainty as to the size of the step. We did not encounter situations in which the move varied wildly.

Remember that if either of these two central assumptions is not met, your moves (or coin tosses) will not cumulatively lead to the bell curve. Depending on what happens, they can lead to the wild Mandelbrotian-style scale-invariant randomness.

“The Ubiquity of the Gaussian”

One of the problems I face in life is that whenever I tell people that the Gaussian

bell curve is not ubiquitous in real life, only in the minds of statisticians, they require me to “prove it”—which is easy to do, as we will see in the next two chapters, yet nobody has managed to prove the opposite. Whenever I suggest a process that is not Gaussian, I am asked to justify my suggestion and to, beyond the phenomena, “give them the theory behind it.” We saw in [Chapter 14](#) the rich-get-richer models that were proposed in order to justify not using a Gaussian. Modelers were forced to spend their time writing theories on possible models that generate the scalable—as if they needed to be apologetic about it. Theory shmeory! I have an epistemological problem with that, with the need to justify the world’s failure to resemble an idealized model that someone blind to reality has managed to promote.

My technique, instead of studying the possible models generating non-bell curve randomness, hence making the same errors of blind theorizing, is to do the opposite: to know the bell curve as intimately as I can and identify where it can and cannot hold. I know where Mediocristan is. To me it is frequently (nay, almost always) the users of the bell curve who do not understand it well, and have to justify it, and not the opposite.

This ubiquity of the Gaussian is not a property of the world, but a problem in our minds, stemming from the way we look at it.

• • •

The next chapter will address the scale invariance of nature and address the properties of the fractal. The chapter after that will probe the misuse of the Gaussian in socioeconomic life and “the need to produce theories.”

I sometimes get a little emotional because I’ve spent a large part of my life thinking about this problem. Since I started thinking about it, and conducting a variety of thought experiments as I have above, I have not for the life of me been able to find anyone around me in the business and statistical world who was intellectually consistent in that he both accepted the Black Swan and rejected the Gaussian and Gaussian tools. Many people accepted my Black Swan idea but could not take it to its logical conclusion, which is that you cannot use one single measure for randomness called standard deviation (and call it “risk”); you cannot expect a *simple* answer to characterize uncertainty. To go the extra step requires courage, commitment, an ability to connect the dots, a desire to understand randomness fully. It also means not accepting other people’s wisdom as gospel.

Then I started finding physicists who had rejected the Gaussian tools but fell for another sin: gullibility about precise predictive models, mostly elaborations around the preferential attachment of [Chapter 14](#)—another form of Platonicity. I could not find anyone with depth and scientific technique who looked at the world of randomness and understood its nature, who looked at calculations as an aid, not a principal aim. It took me close to a decade and a half to find that thinker, the man who made many swans gray: Mandelbrot—the great Benoît Mandelbrot.

* The nontechnical (or intuitive) reader can skip this chapter, as it goes into some details about the bell curve. Also, you can skip it if you belong to the category of fortunate people who do not know about the bell curve.

* I have fudged the numbers a bit for simplicity's sake.

* One of the most misunderstood aspects of a Gaussian is its fragility and vulnerability in the estimation of tail events. The odds of a 4 sigma move are twice that of a 4.15 sigma. The odds of a 20 sigma are a trillion times higher than those of a 21 sigma! It means that a small measurement error of the sigma will lead to a massive underestimation of the probability. We can be a trillion times wrong about some events.

† My main point, which I repeat in some form or another throughout Part Three, is as follows. Everything is made easy, conceptually, when you consider that there are two, and only two, possible paradigms: nonscalable (like the Gaussian) and *other* (such as Mandebrotian randomness). The rejection of the application of the nonscalable is sufficient, as we will see later, *to eliminate a certain vision of the world*. This is like negative empiricism: I know a lot by determining what is wrong.

* Note that variables may not be infinitely scalable; there could be a very, very remote upper limit—but we do not know where it is so we treat a given situation as if it were infinitely scalable. Technically, you cannot sell more of one book than there are denizens of the planet—but that upper limit is large enough to be treated as if it didn't exist. Furthermore, who knows, by repackaging the book, you might be able to sell it to a person twice, or get that person to watch the same movie several times.

† As I was revising this draft, in August 2006, I stayed at a hotel in Dedham, Massachusetts, near one of my children's summer camps. There, I was a little intrigued by the abundance of weight-challenged people walking around the lobby and causing problems with elevator backups. It turned out that the annual convention of NAFA, the National Association for Fat Acceptance, was being held there. As most of the members were extremely overweight, I was not able to figure out which delegate was the heaviest: some form of equality prevailed among the very heavy (someone much heavier than the persons I saw would have been dead). I am sure that at the NARA convention, the National Association for Rich Acceptance, one person would dwarf the others, and, even among the superrich, a very small percentage would represent a large section of the total wealth.

THE AESTHETICS OF RANDOMNESS

Mandelbrot's library—Was Galileo blind?—Pearls to swine—Self-affinity—How the world can be complicated in a simple way, or, perhaps, simple in a very complicated way

THE POET OF RANDOMNESS

It was a melancholic afternoon when I smelled the old books in Benoît Mandelbrot's library. This was on a hot day in August 2005, and the heat exacerbated the musty odor of the glue of old French books bringing on powerful olfactory nostalgia. I usually succeed in repressing such nostalgic excursions, but not when they sneak up on me as music or smell. The odor of Mandelbrot's books was that of French literature, of my parents' library, of the hours spent in bookstores and libraries when I was a teenager when many books around me were (alas) in French, when I thought that Literature was above anything and everything. (I haven't been in contact with many French books since my teenage days.) However abstract I wanted it to be, Literature had a physical embodiment, it had a smell, and this was it.

The afternoon was also gloomy because Mandelbrot was moving away, exactly when I had become entitled to call him at crazy hours just because I had a question, such as why people didn't realize that the 80/20 could be 50/01. Mandelbrot had decided to move to the Boston area, not to retire, but to work for a research center sponsored by a national laboratory. Since he was moving to an apartment in Cambridge, and leaving his oversize house in the Westchester suburbs of New York, he had invited me to come take my pick of his books.

Even the titles of the books had a nostalgic ring. I filled up a box with French titles, such as a 1949 copy of Henri Bergson's *Matière et mémoire*, which it seemed Mandelbrot bought when he was a student (the smell!).

After having mentioned his name left and right throughout this book, I will finally introduce Mandelbrot, principally as the first person with an academic title with whom I ever spoke about randomness without feeling defrauded. Other mathematicians of probability would throw at me theorems with Russian names such as "Sobolev," "Kolmogorov," Wiener measure, without which they were lost; they had a hard time getting to the heart of the subject or exiting their little box long enough to consider its empirical flaws. With Mandelbrot, it was different: it was as if we both originated from the same country, meeting after years of frustrating exile, and were finally able to speak in our mother tongue without straining. He is the only flesh-and-bones teacher I ever had—my teachers are usually books in my library. I had way too little respect for mathematicians dealing with uncertainty and statistics to consider any of them my teachers—in my mind mathematicians, trained for certainties, had no

business dealing with randomness. Mandelbrot proved me wrong.

He speaks an unusually precise and formal French, much like that spoken by Levantines of my parents' generation or Old World aristocrats. This made it odd to hear, on occasion, his accented, but very standard, colloquial American English. He is tall, overweight, which makes him look baby-faced (although I've never seen him eat a large meal), and has a strong physical presence.

From the outside one would think that what Mandelbrot and I have in common is wild uncertainty, Black Swans, and dull (and sometimes less dull) statistical notions. But, although we are collaborators, this is not what our major conversations revolve around. It is mostly matters literary and aesthetic, or historical gossip about people of extraordinary intellectual refinement. I mean refinement, not achievement. Mandelbrot could tell stories about the phenomenal array of hotshots he has worked with over the past century, but somehow I am programmed to consider scientists' personae far less interesting than those of colorful erudites. Like me, Mandelbrot takes an interest in urbane individuals who combine traits generally thought not to coexist together. One person he often mentions is Baron Pierre Jean de Menasce, whom he met at Princeton in the 1950s, where de Menasce was the roommate of the physicist Oppenheimer. De Menasce was exactly the kind of person I am interested in, the embodiment of a Black Swan. He came from an opulent Alexandrian Jewish merchant family, French and Italian-speaking like all sophisticated Levantines. His forebears had taken a Venetian spelling for their Arabic name, added a Hungarian noble title along the way, and socialized with royalty. De Menasce not only converted to Christianity, but became a Dominican priest and a great scholar of Semitic and Persian languages. Mandelbrot kept questioning me about Alexandria, since he was always looking for such characters.

True, intellectually sophisticated characters were exactly what I looked for in life. My erudite and polymathic father—who, were he still alive, would have only been two weeks older than Benoît M.—liked the company of extremely cultured Jesuit priests. I remember these Jesuit visitors occupying my chair at the dining table. I recall that one had a medical degree and a PhD in physics, yet taught Aramaic to locals in Beirut's Institute of Eastern Languages. His previous assignment could have been teaching high school physics, and the one before that was perhaps in the medical school. This kind of erudition impressed my father far more than scientific assembly-line work. I may have something in my genes driving me away from *bildungsphilisters*.

Although Mandelbrot often expressed amazement at the temperament of high-flying erudites and remarkable but not-so-famous scientists, such as his old friend Carleton Gajdusek, a man who impressed him with his ability to uncover the causes of tropical diseases, he did not seem eager to trumpet his association with those we consider great scientists. It took me a while to discover that he had worked with an impressive list of scientists in seemingly every field, something a name-dropper would have brought up continuously. Although I have been working with him for a few years now, only the other day, as I was chatting with his wife, did I discover that he spent two years as the mathematical collaborator of the psychologist Jean Piaget. Another shock came when I discovered that he had also worked with the great historian Fernand Braudel, but Mandelbrot did not seem to be interested in Braudel. He did not care to discuss John von Neuman with whom he had worked as a postdoctoral fellow. His scale was inverted. I asked him once about Charles Tresser, an unknown physicist I met at a party who wrote papers on chaos theory and supplemented his researcher's income by making pastry for a shop he ran near New York City. He was emphatic: "*un homme extraordinaire*," he called Tresser, and could not stop praising him. But when I asked him about a particular famous hotshot, he replied, "He is the prototypical *bon élève*, a student with good grades, no depth, and no vision." That hotshot was a Nobel laureate.

THE PLATONICITY OF TRIANGLES

Now, why am I calling this business Mandelbrotian, or fractal, randomness? Every single bit and piece of the puzzle has been previously mentioned by someone else, such as Pareto, Yule, and Zipf, but it was Mandelbrot who a) connected the dots, b) linked randomness to geometry (and a special brand at that), and c) took the subject to its natural conclusion. Indeed many mathematicians are famous today partly because he dug out their works to back up his claims—the strategy I am following here in this book. “I had to invent my predecessors, so people take me seriously,” he once told me, and he used the credibility of big guns as a rhetorical device. One can almost always ferret out predecessors for any thought. You can always find someone who worked on a part of your argument and use his contribution as your backup. The scientific association with a big idea, the “brand name,” goes to the one who connects the dots, not the one who makes a casual observation—even Charles Darwin, who uncultured scientists claim “invented” the survival of the fittest, was not the first to mention it. He wrote in the introduction of *The Origin of Species* that the facts he presented were not necessarily original; it was the consequences that he thought were “interesting” (as he put it with characteristic Victorian modesty). In the end it is those who derive consequences and seize the importance of the ideas, seeing their real value, who win the day. They are the ones who can talk about the subject.

So let me describe Mandelbrotian geometry.

The Geometry of Nature

Triangles, squares, circles, and the other geometric concepts that made many of us yawn in the classroom may be beautiful and pure notions, but they seem more present in the minds of architects, design artists, modern art buildings, and schoolteachers than in nature itself. That’s fine, except that most of us aren’t aware of this. Mountains are not triangles or pyramids; trees are not circles; straight lines are almost never seen anywhere. Mother Nature did not attend high school geometry courses or read the books of Euclid of Alexandria. Her geometry is jagged, but with a logic of its own and one that is easy to understand.

I have said that we seem naturally inclined to Platonify, and to think

exclusively in terms of studied material: nobody, whether a bricklayer or a natural philosopher, can easily escape the enslavement of such conditioning. Consider that the great Galileo, otherwise a debunker of falsehoods, wrote the following:

The great book of Nature lies ever open before our eyes and the true philosophy is written in it. ... But we cannot read it unless we have first learned the language and the characters in which it is written. ... It is written in mathematical language and the characters are triangles, circles and other geometric figures.

Was Galileo legally blind? Even the great Galileo, with all his alleged independence of mind, was not capable of taking a clean look at Mother Nature. I am confident that he had windows in his house and that he ventured outside from time to time: he should have known that triangles are not easily found in nature. We are so easily brainwashed.

We are either blind, or illiterate, or both. That nature's geometry is not Euclid's was so obvious, and nobody, almost nobody, saw it.

This (physical) blindness is identical to the ludic fallacy that makes us think casinos represent randomness.

Fractality

But first, a description of fractals. Then we will show how they link to what we call power laws, or scalable laws.

Fractal is a word Mandelbrot coined to describe the geometry of the rough and *broken*—from the Latin *fractus*, the origin of *fractured*. *Fractality* is the repetition of geometric patterns at different scales, revealing smaller and smaller versions of themselves. Small parts resemble, to some degree, the whole. I will try to show in this chapter how the fractal applies to the brand of uncertainty that should bear Mandelbrot's name: Mandelbrotian randomness.

The veins in leaves look like branches; branches look like trees; rocks look like small mountains. There is no qualitative change when an object changes size. If you look at the coast of Britain from an airplane, it resembles what you see when you look at it with a magnifying glass. This character of self-affinity implies that one deceptively short and simple rule of iteration can be used, either by a computer or, more randomly, by Mother Nature, to build shapes of seemingly great complexity. This can come in handy for computer graphics, but, more important, it is how nature works. Mandelbrot designed the mathematical

object now known as the Mandelbrot set, the most famous object in the history of mathematics. It became popular with followers of chaos theory because it generates pictures of ever increasing complexity by using a deceptively minuscule recursive rule; *recursive* means that something can be reapplied to itself infinitely. You can look at the set at smaller and smaller resolutions without *ever* reaching the limit; you will continue to see recognizable shapes. The shapes are never the same, yet they bear an affinity to one another, a strong family resemblance.

These objects play a role in aesthetics. Consider the following applications:

Visual arts: Most computer-generated objects are now based on some version of the Mandelbrotian fractal. We can also see fractals in architecture, paintings, and many works of visual art—of course, not consciously incorporated by the work's creator.

Music: Slowly hum the four-note opening of Beethoven's Fifth Symphony: *ta-ta-ta-ta*. Then replace each individual note with the same four-note opening, so that you end up with a measure of sixteen notes. You will see (or, rather, hear) that each smaller wave resembles the original larger one. Bach and Mahler, for instance, wrote submovements that resemble the larger movements of which they are a part.

Poetry: Emily Dickinson's poetry, for instance, is fractal: the large resembles the small. It has, according to a commentator, "a consciously made assemblage of dictions, metres, rhetorics, gestures, and tones."

Fractals initially made Benoît M. a pariah in the mathematical establishment. French mathematicians were horrified. What? Images? *Mon dieu!* It was like showing a porno movie to an assembly of devout Eastern Orthodox grandmothers in my ancestral village of Amioun. So Mandelbrot spent time as an intellectual refugee at an IBM research center in upstate New York. It was a *f*** you money* situation, as IBM let him do whatever he felt like doing.

But the general public (mostly computer geeks) got the point. Mandelbrot's book *The Fractal Geometry of Nature* made a splash when it came out a quarter century ago. It spread through artistic circles and led to studies in aesthetics, architectural design, even large industrial applications. Benoît M. was even offered a position as a professor of medicine! Supposedly the lungs are self-similar. His talks were invaded by all sorts of artists, earning him the nickname the Rock Star of Mathematics. The computer age 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 offers an unusual attribute: it is remarkably easy to understand.

A few words on his biography. Mandelbrot came to France from Warsaw in 1936, at the age of twelve. Owing to the vicissitudes of a clandestine life during Nazi-occupied France, he was spared some of the conventional Gallic education with its uninspiring algebraic drills, becoming largely self-taught. He was later deeply influenced by his uncle Szolem, a prominent member of the French mathematical establishment and holder of a chair at the Collège de France. Benoît M. later settled in the United States, working most of his life as an industrial scientist, with a few transitory and varied academic appointments.

The computer played two roles in the new science Mandelbrot helped conceive. First, fractal objects, as we have seen, can be generated with a simple rule applied to itself, which makes them ideal for the automatic activity of a computer (or Mother Nature). Second, in the generation of visual intuitions lies a dialectic between the mathematician and the objects generated.

Now let us see how this takes us to randomness. In fact, it is with probability that Mandelbrot started his career.

A Visual Approach to Extremistan/Mediocristan

I am looking at the rug in my study. If I examine it with a microscope, I will see a very rugged terrain. If I look at it with a magnifying glass, the terrain will be smoother but still highly uneven. But when I look at it from a standing position, it appears uniform—it is almost as smooth as a sheet of paper. The rug at eye level corresponds to Mediocristan and the law of large numbers: I am seeing the sum of undulations, and *these iron out*. This is like Gaussian randomness: the reason my cup of coffee does not jump is that the sum of all of its moving particles becomes smooth. Likewise, you reach certainties by adding up small Gaussian uncertainties: this is the law of large numbers.

The Gaussian is not self-similar, and that is why my coffee cup does not jump on my desk.

Now, consider a trip up a mountain. No matter how high you go on the surface of the earth, it will remain jagged. This is even true at a height of 30,000 feet. When you are flying above the Alps, you will still see jagged mountains in place of small stones. So some surfaces are not from Mediocristan, and changing

the resolution does not make them much smoother. (Note that this effect only disappears when you go up to more extreme heights. Our planet looks smooth to an observer from space, but this is because it is too small. If it were a bigger planet, then it would have mountains that would dwarf the Himalayas, and it would require observation from a greater distance for it to look smooth. Likewise, if the planet had a larger population, even maintaining the same average wealth, we would be likely to find someone whose net worth would vastly surpass that of Bill Gates.)

Figures 11 and 12 illustrate the above point: an observer looking at the first picture might think that a lens cap has fallen on the ground.

Recall our brief discussion of the coast of Britain. If you look at it from an airplane, its contours are not so different from the contours you see on the shore. The change in scaling does not alter the shapes or their degree of smoothness.

Pearls to Swine

What does fractal geometry have to do with the distribution of wealth, the size of cities, returns in the financial markets, the number of casualties in war, or the size of planets? Let us connect the dots.

The key here is that *the fractal has numerical or statistical measures that are (somewhat) preserved across scales*—the ratio is the same, unlike the Gaussian. Another view of such self-similarity is presented in Figure 13. As we saw in Chapter 15, the superrich are similar to the rich, only richer—wealth is scale independent, or, more precisely, of unknown scale dependence.

In the 1960s Mandelbrot presented his ideas on the prices of commodities and financial securities to the economics establishment, and the financial economists got all excited. In 1963 the then dean of the University of Chicago Graduate School of Business, George Shultz, offered him a professorship. This is the same George Shultz who later became Ronald Reagan's secretary of state.



FIGURE 11: Apparently, a lens cap has been dropped on the ground. Now turn the page.

Shultz called him one evening to rescind the offer.

At the time of writing, forty-four years later, nothing has happened in economics and social science statistics—except for some cosmetic fiddling that treats the world as if we were subject only to mild randomness—and yet Nobel medals were being distributed. Some papers were written offering “evidence” that Mandelbrot was wrong by people who do not get the central argument of this book—you can always produce data “corroborating” that the underlying process is Gaussian by finding periods that do not have rare events, just like you can find an afternoon during which no one killed anyone and use it as “evidence” of honest behavior. I will repeat that, because of the asymmetry with induction, just as it is easier to reject innocence than accept it, it is easier to reject a bell curve than accept it; conversely, it is more difficult to reject a fractal than to accept it. Why? Because a single event can destroy the argument that we face a Gaussian bell curve.

In sum, four decades ago, Mandelbrot gave pearls to economists and résumé-building philistines, which they rejected because the ideas were too good for them. It was, as the saying goes, *margaritas ante porcos*, pearls before swine.



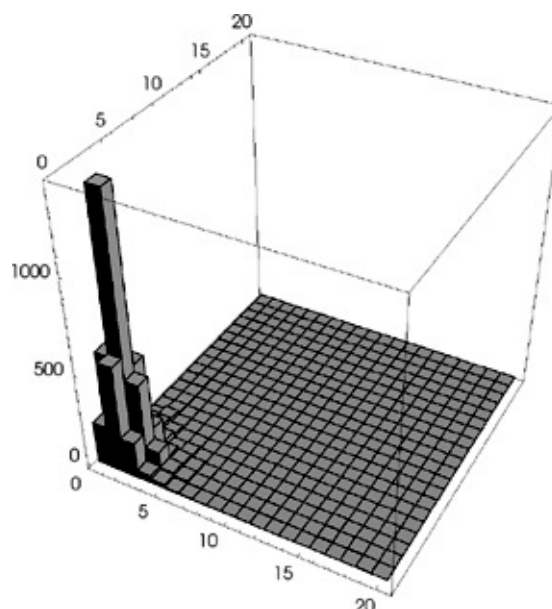
FIGURE 12: The object is not in fact a lens cap. These two photos illustrate scale invariance: the terrain is fractal. Compare it to man-made objects such as a car or a house. *Source: Professor Stephen W. Wheatcraft, University of Nevada, Reno.*

In the rest of this chapter I will explain how I can endorse Mandelbrotian fractals as a representation of much of randomness without necessarily accepting their precise use. Fractals should be the default, the approximation, the framework. They do not solve the Black Swan problem and do not turn all Black Swans into predictable events, but they significantly mitigate the Black Swan problem by making such large events conceivable. (It makes them gray. Why gray? Because only the Gaussian give you certainties. More on that, later.)

THE LOGIC OF FRACTAL RANDOMNESS (WITH A WARNING)*

I have shown in the wealth lists in [Chapter 15](#) the logic of a fractal distribution: if wealth doubles from 1 million to 2 million, the incidence of people with at least that much money is cut in four, which is an exponent of two. If the exponent were one, then the incidence of that wealth or more would be cut in two. The exponent is called the “power” (which is why some people use the term *power law*). Let us call the number of occurrences higher than a certain level an “exceedance”—an exceedance of two million is the number of persons with wealth more than two million. One main property of these fractals (or another way to express their main property, scalability) is that the ratio of two exceedances* is going to be the ratio of the two numbers to the negative power of the power exponent. Let us illustrate this. Say that you “think” that only 96 books a year will sell more than 250,000 copies (which is what happened last year), and that you “think” that the exponent is around 1.5. You can extrapolate to estimate that around 34 books will sell more than 500,000 copies—simply 96 times $(500,000/250,000)^{-1.5}$. We can continue, and note that around 12 books should sell more than a million copies, here 96 times $(1,000,000/250,000)^{-1.5}$.

FIGURE 13: THE PURE FRACTAL STATISTICAL MOUNTAIN



The degree of inequality will be the same in all sixteen subsections of the graph. In the Gaussian world, disparities in wealth (or any other quantity) decrease when you look at the upper end—so billionaires

should be more equal in relation to one another than millionaires are, and millionaires more equal in relation to one another than the middle class. This lack of equality at all wealth levels, in a nutshell, is statistical self-similarity.

TABLE 2: ASSUMED EXPONENTS FOR VARIOUS PHENOMENA*

<i>Phenomenon</i>	<i>Assumed Exponent (vague approximation)</i>
Frequency of use of words	1.2
Number of hits on websites	1.4
Number of books sold in the U.S.	1.5
Telephone calls received	1.22
Magnitude of earthquakes	2.8
Diameter of moon craters	2.14
Intensity of solar flares	0.8
Intensity of wars	0.8
Net worth of Americans	1.1
Number of persons per family name	1
Population of U.S. cities	1.3
Market moves	3 (or lower)
Company size	1.5
People killed in terrorist attacks	2 (but possibly a much lower exponent)

* Source: M.E.J. Newman (2005) and the author's own calculations.

Let me show the different measured exponents for a variety of phenomena.

Let me tell you upfront that these exponents mean very little in terms of numerical precision. We will see why in a minute, but just note for now that we do not *observe* these parameters; we simply guess them, or infer them for statistical information, which makes it hard at times to know the true parameters—if it in fact exists. Let us first examine the practical consequences of an exponent.

TABLE 3: THE MEANING OF THE EXPONENT

<i>Exponent</i>	<i>Share of the top 1%</i>	<i>Share of the top 20%</i>
1	99.99%*	99.99%
1.1	66%	86%
1.2	47%	76%
1.3	34%	69%
1.4	27%	63%
1.5	22%	58%
2	10%	45%
2.5	6%	38%
3	4.6%	34%

Table 3 illustrates the impact of the highly improbable. It shows the contributions of the top 1 percent and 20 percent to the total. The lower the exponent, the higher those contributions. But look how sensitive the process is: between 1.1 and 1.3 you go from 66 percent of the total to 34 percent. Just a 0.2 difference in the exponent changes the result dramatically—and such a difference can come from a simple measurement error. This difference is not trivial: just consider that we have no precise idea what the exponent is because we cannot measure it directly. All we do is estimate from past data or rely on theories that allow for the building of some model that would give us some idea—but these models may have hidden weaknesses that prevent us from blindly applying them to reality.

So keep in mind that the 1.5 exponent is an approximation, that it is hard to compute, that you do not get it from the gods, at least not easily, and that you will have a monstrous sampling error. You will observe that the number of books selling above a million copies is not always going to be 12—It could be as high as 20, or as low as 2.

More significantly, this exponent begins to apply at some number called “crossover,” and addresses numbers larger than this crossover. It may start at 200,000 books, or perhaps only 400,000 books. Likewise, wealth has different properties before, say, \$600 million, when inequality grows, than it does below such a number. How do you know where the crossover point is? This is a problem. My colleagues and I worked with around 20 million pieces of financial data. We all had the same data set, yet we never agreed on exactly what the exponent was in our sets. We knew the data revealed a fractal power law, but we

learned that one could not produce a precise number. But what we did know—that *the distribution is scalable and fractal*—was sufficient for us to operate and make decisions.

The Problem of the Upper Bound

Some people have researched and accepted the fractal “up to a point.” They argue that wealth, book sales, and market returns all have a certain level when things stop being fractal. “Truncation” is what they propose. I agree that there is a level where fractality *might* stop, but where? Saying that there is an upper limit *but I don’t know how high it is*, and saying *there is no limit* carry the same consequences in practice. Proposing an upper limit is highly unsafe. You may say, Let us cap wealth at \$150 billion in our analyses. Then someone else might say, Why not \$151 billion? Or why not \$152 billion? We might as well consider that the variable is unlimited.

Beware the Precision

I have learned a few tricks from experience: whichever exponent I try to measure will be likely to be overestimated (recall that a higher exponent implies a smaller role for large deviations)—what you see is likely to be less Black Swannish than what you do not see. I call this the masquerade problem.

Let’s say I generate a process that has an exponent of 1.7. You do not see what is inside the engine, only the data coming out. If I ask you what the exponent is, odds are that you will compute something like 2.4. You would do so even if you had a million data points. The reason is that it takes a long time for some fractal processes to reveal their properties, and you underestimate the severity of the shock.

Sometimes a fractal can make you believe that it is Gaussian, particularly when the cutpoint starts at a high number. With fractal distributions, extreme deviations of that kind are rare enough to smoke you: you don’t recognize the distribution as fractal.

The Water Puddle Revisited

As you have seen, we have trouble knowing the parameters of whichever model

we assume runs the world. So with Extremistan, the problem of induction pops up again, this time even more significantly than at any previous time in this book. Simply, if a mechanism is fractal it can deliver large values; therefore the incidence of large deviations is possible, but how possible, how often they should occur, will be hard to know with any precision. This is similar to the water puddle problem: plenty of ice cubes could have generated it. As someone who goes from reality to possible explanatory models, I face a completely different spate of problems from those who do the opposite.

I have just read three “popular science” books that summarize the research in complex systems: Mark Buchanan’s *Ubiquity*, Philip Ball’s *Critical Mass*, and Paul Ormerod’s *Why Most Things Fail*. These three authors present the world of social science as full of power laws, a view with which I most certainly agree. They also claim that there is *universality* of many of these phenomena, that there is a wonderful similarity between various processes in nature and the behavior of social groups, which I agree with. They back their studies with the various theories on networks and show the wonderful correspondence between the so-called critical phenomena in natural science and the self-organization of social groups. They bring together processes that generate avalanches, social contagions, and what they call informational cascades, which I agree with.

Universality is one of the reasons physicists find power laws associated with critical points particularly interesting. There are many situations, both in dynamical systems theory and statistical mechanics, where many of the properties of the dynamics around critical points are independent of the details of the underlying dynamical system. The exponent at the critical point may be the same for many systems in the same group, even though many other aspects of the system are different. I almost agree with this notion of universality. Finally, all three authors encourage us to apply techniques from statistical physics, avoiding econometrics and Gaussian-style nonscalable distributions like the plague, and I couldn’t agree more.

But all three authors, by producing, or promoting precision, fall into the trap of not differentiating between the forward and the backward processes (between the problem and the inverse problem)—to me, the greatest scientific and epistemological sin. They are not alone; nearly everyone who works with data but doesn’t make decisions on the basis of these data tends to be guilty of the same sin, a variation of the narrative fallacy. In the absence of a feedback process you look at models and think that they confirm reality. I believe in the

ideas of these three books, but not in the way they are being used—and certainly not with the precision the authors ascribe to them. As a matter of fact, complexity theory should make us *more* suspicious of scientific claims of precise models of reality. It does not make all the swans white; that is predictable: it makes them gray, and only gray.*

As I have said earlier, the world, epistemologically, is literally a different place to a bottom-up empiricist. We don't have the luxury of sitting down to read the equation that governs the universe; we just observe data and make an assumption about what the real process might be, and "calibrate" by adjusting our equation in accordance with additional information. As events present themselves to us, we compare what we see to what we expected to see. It is usually a humbling process, particularly for someone aware of the narrative fallacy, to discover that history runs forward, not backward. As much as one thinks that businessmen have big egos, these people are often humbled by reminders of the differences between decision and results, between precise models and reality.

What I am talking about is opacity, incompleteness of information, the invisibility of the generator of the world. History does not reveal its mind to us—we need to guess what's inside of it.

From Representation to Reality

The above idea links all the parts of this book. While many study psychology, mathematics, or evolutionary theory and look for ways to take it to the bank by applying their ideas to business, I suggest the exact opposite: study the intense, uncharted, humbling uncertainty in the markets as a means to get insights about the nature of randomness that is applicable to psychology, probability, mathematics, decision theory, and even statistical physics. You will see the sneaky manifestations of the narrative fallacy, the ludic fallacy, and the great errors of Platonicity, of going from representation to reality.

When I first met Mandelbrot I asked him why an established scientist like him who should have more valuable things to do with his life would take an interest in such a vulgar topic as finance. I thought that finance and economics were just a place where one learned from various empirical phenomena and filled up one's bank account with *f*** you* cash before leaving for bigger and better things. Mandelbrot's answer was, "*Data*, a gold mine of data." Indeed, everyone forgets

that he started in economics before moving on to physics and the geometry of nature. Working with such abundant data humbles us; it provides the intuition of the following error: traveling the road between representation and reality in the wrong direction.

The problem of the *circularity of statistics* (which we can also call the statistical regress argument) is as follows. Say you need past data to discover whether a probability distribution is Gaussian, fractal, or something else. You will need to establish whether you have enough data to back up your claim. How do we know if we have enough data? From the probability distribution—a distribution does tell you whether you have enough data to “build confidence” about what you are inferring. If it is a Gaussian bell curve, then a few points will suffice (the law of large numbers once again). And how do you know if the distribution is Gaussian? Well, from the data. So we need the data to tell us what the probability distribution is, and a probability distribution to tell us how much data we need. This causes a severe regress argument.

This regress does not occur if you *assume beforehand* that the distribution is Gaussian. It happens that, for some reason, the Gaussian yields its properties rather easily. Extremistan distributions do not do so. So selecting the Gaussian while invoking some general law appears to be convenient. The Gaussian is used as a default distribution for that very reason. As I keep repeating, assuming its application beforehand may work with a small number of fields such as crime statistics, mortality rates, matters from Mediocristan. But not for historical data of unknown attributes and not for matters from Extremistan.

Now, why aren't statisticians who work with historical data aware of this problem? First, they do not like to hear that their entire business has been canceled by the problem of induction. Second, they are not confronted with the results of their predictions in rigorous ways. As we saw with the Makridakis competition, they are grounded in the narrative fallacy, and they do not want to hear it.

ONCE AGAIN, BEWARE THE FORECASTERS

Let me take the problem one step higher up. As I mentioned earlier, plenty of fashionable models attempt to explain the genesis of Extremistan. In fact, they are grouped into two broad classes, but there are occasionally more approaches. The first class includes the simple rich-get-richer (or big-get-bigger) style model that is used to explain the lumping of people around cities, the market domination of Microsoft and VHS (instead of Apple and Betamax), the dynamics of academic reputations, *etc.* The second class concerns what are generally called “percolation models,” which address not the behavior of the individual, but rather the terrain in which he operates. When you pour water on a porous surface, the structure of that surface matters more than does the liquid. When a grain of sand hits a pile of other grains of sand, how the terrain is organized is what determines whether there will be an avalanche.

Most models, of course, attempt to be precisely predictive, not just descriptive; I find this infuriating. They are nice tools for illustrating the genesis of Extremistan, but I insist that the “generator” of reality does not appear to obey them closely enough to make them helpful in precise forecasting. At least to judge by anything you find in the current literature on the subject of Extremistan. Once again we face grave calibration problems, so it would be a great idea to avoid the common mistakes made while calibrating a nonlinear process. Recall that nonlinear processes have greater degrees of freedom than linear ones (as we saw in [Chapter 11](#)), with the implication that you run a great risk of using the wrong model. Yet once in a while you run into a book or articles advocating the application of models from statistical physics to reality. Beautiful books like Philip Ball’s illustrate and inform, but they should not lead to precise quantitative models. Do not take them at face value.

But let us see what we *can* take home from these models.

Once Again, a Happy Solution

First, in assuming a scalable, I accept that an arbitrarily large number is possible. In other words, inequalities should not stop above some *known* maximum bound.

Say that the book *The Da Vinci Code* sold around 60 million copies. (The Bible sold about a billion copies but let’s ignore it and limit our analysis to lay books written by individual authors.) Although we have never known a lay book

to sell 200 million copies, we can consider that the possibility is not zero. It's small, but it's not zero. For every three *Da Vinci Code*-style bestsellers, there might be one superbestseller, and though one has not happened so far, we cannot rule it out. And for every fifteen *Da Vinci Codes* there will be one superbestseller selling, say, 500 million copies.

Apply the same logic to wealth. Say the richest person on earth is worth \$50 billion. There is a nonnegligible probability that next year someone with \$100 billion or more will pop out of nowhere. For every three people with more than \$50 billion, there could be one with \$100 billion or more. There is a much smaller probability of there being someone with more than \$200 billion—one third of the previous probability, but nevertheless not zero. There is even a minute, but not zero probability of there being someone worth more than \$500 billion.

This tells me the following: I can make inferences about things that I do not see in my data, but these things should still belong to the realm of possibilities. There is an invisible bestseller out there, one that is absent from the past data but that you need to account for. Recall my point in [Chapter 13](#): it makes investment in a book or a drug better than statistics on past data might suggest. But it can make stock market losses worse than what the past shows.

Wars are fractal in nature. A war that kills more people than the devastating Second World War is possible—not likely, but not a zero probability, although such a war has never happened in the past.

Second, I will introduce an illustration from nature that will help to make the point about precision. A mountain is somewhat similar to a stone: it has an affinity with a stone, a family resemblance, but it is not identical. The word to describe such resemblances is *self-affine*, not the precise *self-similar*, but Mandelbrot had trouble communicating the notion of affinity, and the term *self-similar* spread with its connotation of precise resemblance rather than family resemblance. As with the mountain and the stone, the distribution of wealth above \$1 billion is not exactly the same as that below \$1 billion, but the two distributions have “affinity.”

Third, I said earlier that there have been plenty of papers in the world of econophysics (the application of statistical physics to social and economic phenomena) aiming at such calibration, at pulling numbers from the world of phenomena. Many try to be predictive. Alas, we are not able to predict “transitions” into crises or contagions. My friend Didier Sornette attempts to

build predictive models, which I love, except that I can not use them to make predictions—but please don't tell him; he might stop building them. That I can't use them as he intends does not invalidate his work, it just makes the interpretations require broad-minded thinking, unlike models in conventional economics that are fundamentally flawed. We may be able to do well with some of Sornette's phenomena, but not all.

WHERE IS THE GRAY SWAN?

I have written this entire book about the Black Swan. This is not because I am in love with the Black Swan; as a humanist, I hate it. I hate most of the unfairness and damage it causes. Thus I would like to eliminate many Black Swans, or at least to mitigate their effects and be protected from them. Fractal randomness is a way to reduce these surprises, to make some of the swans appear possible, so to speak, to make us aware of their consequences, to make them gray. *But fractal randomness does not yield precise answers.* The benefits are as follows. If you know that the stock market *can* crash, as it did in 1987, then such an event is not a Black Swan. The crash of 1987 is not an outlier if you use a fractal with an exponent of three. If you know that biotech companies can deliver a megablockbuster drug, bigger than all we've had so far, then it won't be a Black Swan, and you will not be surprised, should that drug appear.

Thus Mandelbrot's fractals allow us to account for a few Black Swans, but not all. I said earlier that some Black Swans arise because we ignore sources of randomness. Others arise when we overestimate the fractal exponent. A gray swan concerns modelable extreme events, a black swan is about unknown unknowns.

I sat down and discussed this with the great man, and it became, as usual, a linguistic game. In [Chapter 9](#) I presented the distinction economists make between Knightian uncertainty (incomputable) and Knightian risk (computable); this distinction cannot be so original an idea to be absent in our vocabulary, and so we looked for it in French. Mandelbrot mentioned one of his friends and prototypical heroes, the aristocratic mathematician Marcel-Paul Schützenberger, a fine erudite who (like this author) was easily bored and could not work on problems beyond their point of diminishing returns. Schützenberger insisted on the clear-cut distinction in the French language between *hasard* and *fortuit*. *Hasard*, from the Arabic *az-zahr*, implies, like *alea*, dice—tractable randomness; *fortuit* is my Black Swan—the purely accidental and unforeseen. We went to the *Petit Robert* dictionary; the distinction effectively exists there. *Fortuit* seems to correspond to my epistemic opacity, *l'imprévu et non quantifiable*; *hasard* to the more ludic type of uncertainty that was proposed by the Chevalier de Méré in the early gambling literature. Remarkably, the Arabs may have introduced another word to the business of uncertainty: *rizk*, meaning property.

I repeat: Mandelbrot deals with gray swans; I deal with the Black Swan. So

Mandelbrot domesticated many of my Black Swans, but not all of them, not completely. But he shows us a glimmer of hope with his method, a way to start thinking about the problems of uncertainty. You are indeed much safer if you know where the wild animals are.

-
- * The nontechnical reader can skip from here until the end of the chapter.
 - * By using symmetry we could also examine the incidences below the number.
 - * *Clearly, you do not observe 100 percent in a finite sample.*
 - * The defense mechanism when you question their work is to say that they “do science, not philosophy” and to berate my approach of worrying about model errors. This is a common trap: people think that science is about formulating predictions (even when wrong). To me, science is about how not to be a sucker.

Chapter Seventeen

LOCKE'S MADMEN, OR BELL CURVES IN THE WRONG PLACES*

What?—Anyone can become president—Alfred Nobel's legacy—Those medieval days

I have in my house two studies: one real, with interesting books and literary material; the other nonliterary, where I do not enjoy working, where I relegate matters prosaic and narrowly focused. In the nonliterary study is a wall full of books on statistics and the history of statistics, books I never had the fortitude to burn or throw away; though I find them largely useless outside of their academic applications (Carneades, Cicero, and Foucher know a lot more about probability than all these pseudosophisticated volumes). I cannot use them in class because I promised myself never to teach trash, even if dying of starvation. Why can't I use them? Not one of these books deals with Extremistan. Not one. The few books that do are not by statisticians but by statistical physicists. We are teaching people methods from Mediocristan and turning them loose in Extremistan. It is like developing a medicine for plants and applying it to humans. It is no wonder that we run the biggest risk of all: we handle *matters that belong to Extremistan, but treated as if they belonged to Mediocristan*, as an "approximation."

Several hundred thousand students in business schools and social science departments from Singapore to Urbana-Champaign, as well as people in the business world, continue to study "scientific" methods, all grounded in the Gaussian, all embedded in the ludic fallacy.

This chapter examines disasters stemming from the application of phony mathematics to social science. The real topic might be the dangers to our society brought about by the Swedish academy that awards the Nobel Prize.

Only Fifty Years

Let us return to the story of my business life. Look at the graph in [Figure 14](#). In the last fifty years, the ten most extreme days in the financial markets represent half the returns. Ten days in fifty years. Meanwhile, we are mired in chitchat.

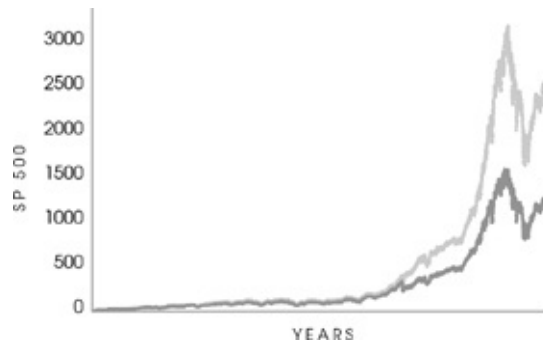
Clearly, anyone who wants more than the high number of six sigma as proof that markets are from Extremistan needs to have his head examined. Dozens of papers show the inadequacy of the Gaussian family of distributions and the scalable nature of markets. Recall that, over the years, I myself have run statistics backward and forward on 20 million pieces of data that made me despise anyone talking about markets in Gaussian terms. But people have a hard time making the leap to the consequences of this knowledge.

The strangest thing is that people in business usually agree with me when they listen to me talk or hear me make my case. But when they go to the office the next day they revert to the Gaussian tools so entrenched in their habits. Their minds are domain-dependent, so they can exercise critical thinking at a conference while not doing so in the office. Furthermore, the Gaussian tools give them numbers, which seem to be “better than nothing.” The resulting measure of future uncertainty satisfies our ingrained desire to simplify even if that means squeezing into one single number matters that are too rich to be described that way.

The Clerks’ Betrayal

I ended [Chapter 1](#) with the stock market crash of 1987, which allowed me to aggressively pursue my Black Swan idea. Right after the crash, when I stated that those using sigmas (i.e., standard deviations) as a measure of the degree of risk and randomness were charlatans, everyone agreed with me. If the world of finance were Gaussian, an episode such as the crash (more than twenty standard deviations) would take place every several billion lifetimes of the universe (look at the height example in [Chapter 15](#)). According to the circumstances of 1987, people accepted that rare events take place and are the main source of uncertainty. They were just unwilling to give up on the Gaussian as a central measurement tool—“Hey, we have nothing else.” People want a number to anchor on. Yet the two methods are logically incompatible.

FIGURE 14



By removing the ten biggest one-day moves from the U.S. stock market over the past fifty years, we see a huge difference in returns—and yet conventional finance sees these one-day jumps as mere anomalies. (This is only one of many such tests. While it is quite convincing on a casual read, there are many more-convincing ones from a mathematical standpoint, such as the incidence of 10 sigma events.)

Unbeknownst to me, 1987 was not the first time the idea of the Gaussian was shown to be lunacy. Mandelbrot proposed the scalable to the economics establishment around 1960, and showed them how the Gaussian curve did not fit prices *then*. But after they got over their excitement, they realized that they would have to relearn their trade. One of the influential economists of the day, the late Paul Cootner, wrote, “Mandelbrot, like Prime Minister Churchill before him, promised us not utopia, but blood, sweat, toil, and tears. If he is right, almost all our statistical tools are obsolete [or] meaningless.” I propose two corrections to Cootner’s statement. First, I would replace *almost all* with *all*. Second, I disagree with the blood and sweat business. I find Mandelbrot’s randomness considerably easier to understand than the conventional statistics. If you come fresh to the business, do not rely on the old theoretical tools, and do not have a high expectation of certainty.

Anyone Can Become President

And now a brief history of the “Nobel” Prize in economics, which was established by the Bank of Sweden in honor of Alfred Nobel, who may be, according to his family who wants the prize abolished, now rolling in his grave with disgust. An activist family member calls the prize a public relations coup by economists aiming to put their field on a higher footing than it deserves. True, the prize has gone to some valuable thinkers, such as the empirical psychologist Daniel Kahneman and the thinking economist Friedrich Hayek. But the committee has gotten into the habit of handing out Nobel Prizes to those who “bring rigor” to the process with pseudoscience and phony mathematics. *After*

the stock market crash, they rewarded two theoreticians, Harry Markowitz and William Sharpe, who built beautifully Platonic models on a Gaussian base, contributing to what is called Modern Portfolio Theory. Simply, if you remove their Gaussian assumptions and treat prices as scalable, you are left with hot air. The Nobel Committee could have tested the Sharpe and Markowitz models—they work like quack remedies sold on the Internet—but nobody in Stockholm seems to have thought of it. Nor did the committee come to us practitioners to ask us our opinions; instead it relied on an academic vetting process that, in some disciplines, can be corrupt all the way to the marrow. After that award I made a prediction: “In a world in which these two get the Nobel, anything can happen. Anyone can become president.”

So the Bank of Sweden and the Nobel Academy are largely responsible for giving credence to the use of the Gaussian Modern Portfolio Theory as institutions have found it a great cover-your-behind approach. Software vendors have sold “Nobel crowned” methods for millions of dollars. How could you go wrong using it? Oddly enough, everyone in the business world initially knew that the idea was a fraud, but people get used to such methods. Alan Greenspan, the chairman of the Federal Reserve bank, supposedly blurted out, “I’d rather have the opinion of a trader than a mathematician.” Meanwhile, the Modern Portfolio Theory started spreading. I will repeat the following until I am hoarse: it is contagion that determines the fate of a theory in social science, not its validity.

I only realized later that Gaussian-trained finance professors were taking over business schools, and therefore MBA programs, and producing close to a hundred thousand students a year in the United States alone, all brainwashed by a phony portfolio theory. No empirical observation could halt the epidemic. It seemed better to teach students a theory based on the Gaussian than to teach them no theory at all. It looked more “scientific” than giving them what Robert C. Merton (the son of the sociologist Robert K. Merton we discussed earlier) called the “anecdote.” Merton wrote that before portfolio theory, finance was “a collection of anecdotes, rules of thumb, and manipulation of accounting data.” Portfolio theory allowed “the subsequent evolution from this conceptual potpourri to a rigorous economic theory.” For a sense of the degree of intellectual seriousness involved, and to compare neoclassical economics to a more honest science, consider this statement from the nineteenth-century father of modern medicine, Claude Bernard: “Facts for now, but with scientific aspirations for later.” You should send economists to medical school.

So the Gaussian* pervaded our business and scientific cultures, and terms such as *sigma*, *variance*, *standard deviation*, *correlation*, *R square*, and the eponymous *Sharpe ratio*, all directly linked to it, pervaded the lingo. If you read a mutual fund prospectus, or a description of a hedge fund's exposure, odds are that it will supply you, among other information, with some quantitative summary claiming to measure "risk." That measure will be based on one of the above buzzwords derived from the bell curve and its kin. Today, for instance, pension funds' investment policy and choice of funds are vetted by "consultants" who rely on portfolio theory. If there is a problem, they can claim that they relied on standard scientific method.

More Horror

Things got a lot worse in 1997. The Swedish academy gave another round of Gaussian-based Nobel Prizes to Myron Scholes and Robert C. Merton, who had improved on an old mathematical formula and made it compatible with the existing grand Gaussian general financial equilibrium theories—hence acceptable to the economics establishment. The formula was now "useable." It had a list of long forgotten "precursors," among whom was the mathematician and gambler Ed Thorp, who had authored the bestselling *Beat the Dealer*, about how to get ahead in blackjack, but somehow people believe that Scholes and Merton invented it, when in fact they just made it acceptable. The formula was my bread and butter. Traders, bottom-up people, know its wrinkles better than academics by dint of spending their nights worrying about their risks, except that few of them could express their ideas in technical terms, so I felt I was representing them. Scholes and Merton made the formula dependent on the Gaussian, but their "precursors" subjected it to no such restriction.*

The postcrash years were entertaining for me, intellectually. I attended conferences in finance and mathematics of uncertainty; not once did I find a speaker, Nobel or no Nobel, who understood what he was talking about when it came to probability, so I could freak them out with my questions. They did "deep work in mathematics," but when you asked them where they got their probabilities, their explanations made it clear that they had fallen for the ludic fallacy—there was a strange cohabitation of technical skills and absence of understanding that you find in idiot savants. Not once did I get an intelligent answer or one that was not *ad hominem*. Since I was questioning their entire

business, it was understandable that I drew all manner of insults: “obsessive,” “commercial,” “philosophical,” “essayist,” “idle man of leisure,” “repetitive,” “practitioner” (this is an insult in academia), “academic” (this is an insult in business). Being on the receiving end of angry insults is not that bad; you can get quickly used to it and focus on what is *not* said. Pit traders are trained to handle angry rants. If you work in the chaotic pits, someone in a particularly bad mood from losing money might start cursing at you until he injures his vocal cords, then forget about it and, an hour later, invite you to his Christmas party. So you become numb to insults, particularly if you teach yourself to imagine that the person uttering them is a variant of a noisy ape with little personal control. Just keep your composure, smile, focus on analyzing the speaker not the message, and you’ll win the argument. An ad hominem attack against an intellectual, not against an idea, is highly flattering. It indicates that the person does not have anything intelligent to say about your message.

The psychologist Philip Tetlock (the expert buster in [Chapter 10](#)), after listening to one of my talks, reported that he was struck by the presence of an acute state of cognitive dissonance in the audience. But how people resolve this cognitive tension, as it strikes at the core of everything they have been taught and at the methods they practice, and realize that they will continue to practice, can vary a lot. It was symptomatic that almost all people who attacked my thinking attacked a deformed version of it, like “it is all random and unpredictable” rather than “it is largely random,” or got mixed up by showing me how the bell curve works in some physical domains. Some even had to change my biography. At a panel in Lugano, Myron Scholes once got in to a state of rage, and went after a transformed version of my ideas. I could see pain in his face. Once, in Paris, a prominent member of the mathematical establishment, who invested part of his life on some minute sub-sub-property of the Gaussian, blew a fuse—right when I showed empirical evidence of the role of Black Swans in markets. He turned red with anger, had difficulty breathing, and started hurling insults at me for having desecrated the institution, lacking *pudeur* (modesty); he shouted “I am a member of the Academy of Science!” to give more strength to his insults. (The French translation of my book was out of stock the next day.) My best episode was when Steve Ross, an economist perceived to be an intellectual far superior to Scholes and Merton, and deemed a formidable debater, gave a rebuttal to my ideas by signaling small errors or approximations in my presentation, such as “Markowitz was not the first to ...” thus certifying that he had no answer to my main point. Others who had invested

much of their lives in these ideas resorted to vandalism on the Web. Economists often invoke a strange argument by Milton Friedman that states that models do not have to have realistic assumptions to be acceptable—giving them license to produce severely defective mathematical representations of reality. The problem of course is that these Gaussianizations do not have realistic assumptions and do not produce reliable results. They are neither realistic nor predictive. Also note a mental bias I encounter on the occasion: people mistake an event with a small probability, say, one in twenty years for a periodically occurring one. They think that they are safe if they are only exposed to it for ten years.

I had trouble getting the message about the difference between Mediocristan and Extremistan through—many arguments presented to me were about how society has done well with the bell curve—just look at credit bureaus, *etc.*

The only comment I found unacceptable was, “You are right; we need you to remind us of the weakness of these methods, but you cannot throw the baby out with the bath water,” meaning that I needed to accept their reductive Gaussian distribution while also accepting that large deviations could occur—they didn’t realize the incompatibility of the two approaches. It was as if one could be half dead. Not one of these users of portfolio theory in twenty years of debates, explained *how* they could accept the Gaussian framework as well as large deviations. Not one.

Confirmation

Along the way I saw enough of the confirmation error to make Karl Popper stand up with rage. People would find data in which there were no jumps or extreme events, and show me a “proof” that one could use the Gaussian. This was exactly like my example of the “proof” that O. J. Simpson is not a killer in [Chapter 5](#). The entire statistical business confused absence of proof with proof of absence. Furthermore, people did not understand the elementary asymmetry involved: you need one single observation to reject the Gaussian, but millions of observations will not fully confirm the validity of its application. Why? Because the Gaussian bell curve disallows large deviations, but tools of Extremistan, the alternative, do not disallow long quiet stretches.

I did not know that Mandelbrot’s work mattered outside aesthetics and geometry. Unlike him, I was not ostracized: I got a lot of approval from practitioners and decision makers, though not from their research staffs.

But suddenly I got the most unexpected vindication.

IT WAS JUST A BLACK SWAN

Robert Merton, Jr., and Myron Scholes were founding partners in the large speculative trading firm called Long-Term Capital Management, or LTCM, which I mentioned in [Chapter 4](#). It was a collection of people with top-notch résumés, from the highest ranks of academia. They were considered geniuses. The ideas of portfolio theory inspired their risk management of possible outcomes—thanks to their sophisticated “calculations.” They managed to enlarge the ludic fallacy to industrial proportions.

Then, during the summer of 1998, a combination of large events, triggered by a Russian financial crisis, took place that lay outside their models. It was a Black Swan. LTCM went bust and almost took down the entire financial system with it, as the exposures were massive. Since their models ruled out the possibility of large deviations, they allowed themselves to take a monstrous amount of risk. The ideas of Merton and Scholes, as well as those of Modern Portfolio Theory, were starting to go bust. The magnitude of the losses was spectacular, too spectacular to allow us to ignore the intellectual comedy. Many friends and I thought that the portfolio theorists would suffer the fate of tobacco companies: they were endangering people’s savings and would soon be brought to account for the consequences of their Gaussian-inspired methods.

None of that happened.

Instead, MBAs in business schools went on learning portfolio theory. And the option formula went on bearing the name Black-Scholes-Merton, instead of reverting to its true owners, Louis Bachelier, Ed Thorp, and others.

How to “Prove” Things

Merton the younger is a representative of the school of neoclassical economics, which, as we have seen with LTCM, represents most powerfully the dangers of Platonified knowledge.* Looking at his methodology, I see the following pattern. He starts with rigidly Platonic assumptions, completely unrealistic—such as the Gaussian probabilities, along with many more equally disturbing ones. Then he generates “theorems” and “proofs” from these. The math is tight and elegant. The theorems are compatible with other theorems from Modern Portfolio Theory, themselves compatible with still other theorems, building a grand theory of how people consume, save, face uncertainty, spend, and project the future. He

assumes that we know the likelihood of events. The beastly word *equilibrium* is always present. But the whole edifice is like a game that is entirely closed, like Monopoly with all of its rules.

A scholar who applies such methodology resembles Locke's definition of a madman: someone "reasoning correctly from erroneous premises."

Now, elegant mathematics has this property: it is perfectly right, not 99 percent so. This property appeals to mechanistic minds who do not want to deal with ambiguities. Unfortunately you have to cheat somewhere to make the world fit perfect mathematics; and you have to fudge your assumptions somewhere. We have seen with the Hardy quote that professional "pure" mathematicians, however, are as honest as they come.

So where matters get confusing is when someone like Merton tries to be mathematical and airtight rather than focus on fitness to reality.

This is where you learn from the minds of military people and those who have responsibilities in security. They do not care about "perfect" ludic reasoning; they want realistic ecological assumptions. In the end, they care about lives.

I mentioned in [Chapter 11](#) how those who started the game of "formal thinking," by manufacturing phony premises in order to generate "rigorous" theories, were Paul Samuelson, Merton's tutor, and, in the United Kingdom, John Hicks. These two wrecked the ideas of John Maynard Keynes, which they tried to formalize (Keynes was interested in uncertainty, and complained about the mind-closing certainties induced by models). Other participants in the formal thinking venture were Kenneth Arrow and Gerard Debreu. All four were Nobeled. All four were in a delusional state under the effect of mathematics—what Dieudonné called "the music of reason," and what I call Locke's madness. All of them can be safely accused of having invented an imaginary world, one that lent itself to their mathematics. The insightful scholar Martin Shubik, who held that the degree of excessive abstraction of these models, a few steps beyond necessity, makes them totally unusable, found himself ostracized, a common fate for dissenters.*

If you question what they do, as I did with Merton Jr., they will ask for "tight proof." So they set the rules of the game, and you need to play by them. Coming from a practitioner background in which the principal asset is being able to work with messy, but empirically acceptable, mathematics, I cannot accept a pretense of science. I much prefer a sophisticated craft, focused on tricks, to a failed science looking for certainties. Or could these neoclassical model builders be

doing something worse? Could it be that they are involved in what Bishop Huet calls the manufacturing of certainties?

TABLE 4: TWO WAYS TO APPROACH RANDOMNESS

Skeptical Empiricism and the a-Platonic School

Interested in what lies outside the Platonic fold

Respect for those who have the guts to say “I don’t know”

Fat Tony

Thinks of Black Swans as a dominant source of randomness

Bottom-up

Would ordinarily not wear suits (except to funerals)

Prefers to be broadly right

Minimal theory, considers theorizing as a disease to resist

Does not believe that we can easily compute probabilities

Model: Sextus Empiricus and the school of evidence-based, minimum-theory empirical medicine

Develops intuitions from practice, goes from observations to books

Not inspired by any science, uses messy mathematics and computational methods

The Platonic Approach

Focuses on the inside of the Platonic fold

“You keep criticizing these models. These models are all we have.”

Dr. John

Thinks of ordinary fluctuations as a dominant source of randomness, with jumps as an afterthought

Top-down

Wears dark suits, white shirts; speaks in a boring tone

Precisely wrong

Everything needs to fit some grand, general socioeconomic model and “the rigor of economic theory;” frowns on the “descriptive”

Built their entire apparatus on the assumptions that we can compute probabilities

Model: Laplacian mechanics, the world and the economy like a clock

Relies on scientific papers, goes from books to practice

Inspired by physics, relies on abstract mathematics

These two approaches are the only ones that have been used in the history of economics.

ideas based on skepticism, on the unread books in the library	ideas based on beliefs, on what they think they know
Assumes Extremistan as a starting point	Assumes Mediocristan as a starting point
Sophisticated craft	Poor science
Seeks to be approximately right across a broad set of eventualities	Seeks to be perfectly right in a narrow model, under precise assumptions

Let us see.

Skeptical empiricism advocates the opposite method. I care about the premises more than the theories, and I want to minimize reliance on theories, stay light on my feet, and reduce my surprises. I want to be broadly right rather than precisely wrong. Elegance in the theories is often indicative of Platoncity and weakness—it invites you to seek elegance for elegance’s sake. A theory is like medicine (or government): often useless, sometimes necessary, always self-serving, and on occasion lethal. So it needs to be used with care, moderation, and close adult supervision.

The distinction in the above table between my model modern, skeptical empiricist and what Samuelson’s puppies represent can be generalized across disciplines.

I’ve presented my ideas in finance because that’s where I refined them. Let us now examine a category of people expected to be more thoughtful: the philosophers.

-
- * This is a simple illustration of the general point of this book in finance and economics. If you do not believe in applying the bell curve to social variables, and if, like many professionals, you are already convinced that “modern” financial theory is dangerous junk science, you can safely skip this chapter.
 - * Granted, the Gaussian has been tinkered with, using such methods as complementary “jumps,” stress testing, regime switching, or the elaborate methods known as GARCH, but while these methods represent a good effort, they fail to address the bell curve’s fundamental flaws. Such methods are not scale-invariant. This, in my opinion, can explain the failures of sophisticated methods in real life as shown by the Makridakis competition.
 - * More technically, remember my career as an option professional. Not only does an option on a very long shot benefit from Black Swans, but it benefits disproportionately from them—something Scholes and Merton’s “formula” misses. The option payoff is so powerful that you do not have to be right on the odds: you can be wrong on the probability, but get a monstrously large payoff. I’ve called this the “double bubble”: the mispricing of the probability and that of the payoff.

- * I am selecting Merton because I found him very illustrative of academically stamped obscurantism. I discovered Merton's shortcomings from an angry and threatening seven-page letter he sent me that gave me the impression that he was not too familiar with how we trade options, his very subject matter. He seemed to be under the impression that traders rely on "rigorous" economic theory—as if birds had to study (bad) engineering in order to fly.
- * Medieval medicine was also based on equilibrium ideas when it was top-down and similar to theology. Luckily its practitioners went out of business, as they could not compete with the bottom-up surgeons, ecologically driven former barbers who gained clinical experience, and after whom a-Platonic clinical science was born. If I am alive, today, it is because scholastic top-down medicine went out of business a few centuries ago.

Chapter Eighteen

THE UNCERTAINTY OF THE PHONY

*Philosophers in the wrong places—Uncertainty about (mostly) lunch—What I don't care about—
Education and intelligence*

This final chapter of Part Three focuses on a major ramification of the ludic fallacy: how those whose job it is to make us aware of uncertainty fail us and divert us into bogus certainties through the back door.

LUDIC FALLACY REDUX

I have explained the ludic fallacy with the casino story, and have insisted that the sterilized randomness of games does not resemble randomness in real life. Look again at [Figure 7](#) in [Chapter 15](#). The dice average out so quickly that I can say with certainty that the casino will beat me in the very near long run at, say, roulette, as the noise will cancel out, though not the skills (here, the casino's advantage). The more you extend the period (or reduce the size of the bets) the more randomness, by virtue of averaging, drops out of these gambling constructs.

The ludic fallacy is present in the following chance setups: random walk, dice throwing, coin flipping, the infamous digital “heads or tails” expressed as 0 or 1, Brownian motion (which corresponds to the movement of pollen particles in water), and similar examples. These setups generate a quality of randomness that does not even qualify as randomness—*protorandomness* would be a more appropriate designation. At their core, all theories built around the ludic fallacy ignore a layer of uncertainty. Worse, their proponents do not know it!

One severe application of such focus on small, as opposed to large, uncertainty concerns the hackneyed *greater uncertainty principle*.

Find the Phony

The greater uncertainty principle states that in quantum physics, one cannot measure certain pairs of values (with arbitrary precision), such as the position and momentum of particles. You will hit a lower bound of measurement: what you gain in the precision of one, you lose in the other. So there is an incompressible uncertainty that, in theory, will defy science and forever remain an uncertainty. This minimum uncertainty was discovered by Werner Heisenberg in 1927. I find it ludicrous to present the uncertainty principle as having anything to do with uncertainty. Why? First, this uncertainty is Gaussian. On average, it will disappear—recall that no one person's weight will significantly change the total weight of a thousand people. We may always remain uncertain about the future positions of small particles, but these uncertainties are very small and very numerous, and they average out—for Pluto's sake, they average out! They obey the law of large numbers we discussed in [Chapter 15](#). Most other types of randomness do not average out! If there is

one thing on this planet that is not so uncertain, it is the behavior of a collection of subatomic particles! Why? Because, as I have said earlier, when you look at an object, composed of a collection of particles, the fluctuations of the particles tend to balance out.

But political, social, and weather events do not have this handy property, and we patently cannot predict them, so when you hear “experts” presenting the problems of uncertainty in terms of subatomic particles, odds are that the expert is a phony. As a matter of fact, this may be the best way to spot a phony.

I often hear people say, “Of course there are limits to our knowledge,” then invoke the greater uncertainty principle as they try to explain that “we cannot model everything”—I have heard such types as the economist Myron Scholes say this at conferences. But I am sitting here in New York, in August 2006, trying to go to my ancestral village of Amioun, Lebanon. Beirut’s airport is closed owing to the conflict between Israel and the Shiite militia Hezbollah. There is no published airline schedule that will inform me when the war will end, if it ends. I can’t figure out if my house will be standing, if Amioun will still be on the map—recall that the family house was destroyed once before. I can’t figure out whether the war is going to degenerate into something even more severe. Looking into the outcome of the war, with all my relatives, friends, and property exposed to it, I face *true* limits of knowledge. Can someone explain to me why I should care about subatomic particles that, anyway, converge to a Gaussian? People can’t predict how long they will be happy with recently acquired objects, how long their marriages will last, how their new jobs will turn out, yet it’s subatomic particles that they cite as “limits of prediction.” They’re ignoring a mammoth standing in front of them in favor of matter even a microscope would not allow them to see.

Can Philosophers Be Dangerous to Society?

I will go further: people who worry about pennies instead of dollars can be dangerous to society. They mean well, but, invoking my Bastiat argument of [Chapter 8](#), they are a threat to us. They are wasting our studies of uncertainty by focusing on the insignificant. Our resources (both cognitive and scientific) are limited, perhaps too limited. Those who distract us increase the risk of Black Swans.

This commoditization of the notion of uncertainty as symptomatic of Black

Swan blindness is worth discussing further here.

Given that people in finance and economics are seeped in the Gaussian to the point of choking on it, I looked for financial economists with philosophical bents to see how their critical thinking allows them to handle this problem. I found a few. One such person got a PhD in philosophy, then, four years later, another in finance; he published papers in both fields, as well as numerous textbooks in finance. But I was disheartened by him: he seemed to have compartmentalized his ideas on uncertainty so that he had two distinct professions: philosophy and quantitative finance. The problem of induction, Mediocristan, epistemic opacity, or the offensive assumption of the Gaussian—these did not hit him as true problems. His numerous textbooks drilled Gaussian methods into students' heads, as though their author had forgotten that he was a philosopher. Then he promptly remembered that he was when writing philosophy texts on seemingly scholarly matters.

The same context specificity leads people to take the escalator to the StairMasters, but the philosopher's case is far, far more dangerous since he uses up our storage for critical thinking in a sterile occupation. Philosophers like to practice philosophical thinking on me-too subjects that other philosophers call philosophy, and they leave their minds at the door when they are outside of these subjects.

The Problem of Practice

As much as I rail against the bell curve, Platonicity, and the ludic fallacy, my principal problem is not so much with statisticians—after all, these are computing people, not thinkers. We should be far less tolerant of philosophers, with their bureaucratic apparatchiks closing our minds. Philosophers, the watchdogs of critical thinking, have duties beyond those of other professions.

HOW MANY WITTGENSTEINS CAN DANCE ON THE HEAD OF A PIN?

A number of semishabbily dressed (but thoughtful-looking) people gather in a room, silently looking at a guest speaker. They are all professional philosophers attending the prestigious weekly colloquium at a New York–area university. The speaker sits with his nose drowned in a set of typewritten pages, from which he reads in a monotone voice. He is hard to follow, so I daydream a bit and lose his thread. I can vaguely tell that the discussion revolves around some “philosophical” debate about Martians invading your head and controlling your will, all the while preventing you from knowing it. There seem to be several theories concerning this idea, but the speaker’s opinion differs from those of other writers on the subject. He spends some time showing where his research on these head-hijacking Martians is unique. After his monologue (fifty-five minutes of relentless reading of the typewritten material) there is a short break, then another fifty-five minutes of discussion about Martians planting chips and other outlandish conjectures. Wittgenstein is occasionally mentioned (you can always mention Wittgenstein since he is vague enough to always seem relevant).

Every Friday, at four P.M., the paychecks of these philosophers will hit their respective bank accounts. A fixed proportion of their earnings, about 16 percent on average, will go into the stock market in the form of an automatic investment into the university’s pension plan. These people are professionally employed in the business of questioning what we take for granted; they are trained to argue about the existence of god(s), the definition of truth, the redness of red, the meaning of meaning, the difference between the semantic theories of truth, conceptual and nonconceptual representations ... Yet they believe blindly in the stock market, and in the abilities of their pension plan manager. Why do they do so? Because they accept that this is what people should do with their savings, because “experts” tell them so. They doubt their own senses, but not for a second do they doubt their automatic purchases in the stock market. This domain dependence of skepticism is no different from that of medical doctors (as we saw in [Chapter 8](#)).

Beyond this, they may believe without question that we can predict societal events, that the Gulag will toughen you a bit, that politicians know more about what is going on than their drivers, that the chairman of the Federal Reserve saved the economy, and so many such things. They may also believe that nationality matters (they always stick “French,” “German,” or “American” in

front of a philosopher's name, as if this has something to do with anything he has to say). Spending time with these people, whose curiosity is focused on regimented on-the-shelf topics, feels stifling.

Where Is Popper When You Need Him?

I hope I've sufficiently drilled home the notion that, as a practitioner, my thinking is rooted in the belief that you cannot go from books to problems, but the reverse, from problems to books. This approach incapacitates much of that career-building verbiage. A scholar should not be a library's tool for making another library, as in the joke by Daniel Dennett.

Of course, what I am saying here has been said by philosophers before, at least by the real ones. The following remark is one reason I have inordinate respect for Karl Popper; it is one of the few quotations in this book that I am not attacking.

The degeneration of philosophical schools in its turn is the consequence of the mistaken belief that one can philosophize without having been compelled to philosophize by problems outside philosophy. ... *Genuine philosophical problems are always rooted outside philosophy and they die if these roots decay.* ... [emphasis mine] These roots are easily forgotten by philosophers who "study" philosophy instead of being forced into philosophy by the pressure of nonphilosophical problems.

Such thinking may explain Popper's success outside philosophy, particularly with scientists, traders, and decision makers, as well as his relative failure inside of it. (He is rarely studied by his fellow philosophers; they prefer to write essays on Wittgenstein.)

Also note that I do not want to be drawn into philosophical debates with my Black Swan idea. What I mean by Platonicity is not so metaphysical. Plenty of people have argued with me about whether I am against "essentialism" (i.e., things that I hold don't have a Platonic essence), if I believe that mathematics would work in an alternative universe, or some such thing. Let me set the record straight. I am a no-nonsense practitioner; I am not saying that mathematics does not correspond to an objective structure of reality; my entire point is that we are, epistemologically speaking, putting the cart before the horse and, of the space of possible mathematics, risk using the wrong one and being blinded by it. I truly believe that there are some mathematics that work, but that these are not as easily within our reach as it seems to the "confirmators."

The Bishop and the Analyst

I am most often irritated by those who attack the bishop but somehow fall for the securities analyst—those who exercise their skepticism against religion but not against economists, social scientists, and phony statisticians. Using the confirmation bias, these people will tell you that religion was horrible for mankind by counting deaths from the Inquisition and various religious wars. But they will not show you how many people were killed by nationalism, social science, and political theory under Stalinism or during the Vietnam War. Even priests don't go to bishops when they feel ill: their first stop is the doctor's. But we stop by the offices of many pseudo-scientists and "experts" without alternative. We no longer believe in papal infallibility; we seem to believe in the infallibility of the Nobel, though, as we saw in [Chapter 17](#).

Easier Than You Think: The Problem of Decision Under Skepticism

I have said all along that there is a problem with induction and the Black Swan. In fact, matters are far worse: we may have no less of a problem with phony skepticism.

- a. I can't do anything to stop the sun from nonrising tomorrow (no matter how hard I try),
- b. I can't do anything about whether or not there is an afterlife,
- c. I can't do anything about Martians or demons taking hold of my brain.

But I have plenty of ways to avoid being a sucker. It is not much more difficult than that.

I conclude Part Three by reiterating that my antidote to Black Swans is precisely to be noncommoditized in my thinking. But beyond avoiding being a sucker, this attitude lends itself to a protocol of how to act—not how to think, but how to convert knowledge into action and figure out what knowledge is worth. Let us examine what to do or not do with this in the concluding section of this book.



Chapter Nineteen

HALF AND HALF, OR HOW TO GET EVEN WITH THE BLACK SWAN

The other half—Remember Apelles—When missing a train can be painful

It is now time for a few last words.

Half the time I am a hyperskeptic; the other half I hold certainties and can be intransigent about them, with a very stubborn disposition. Of course I am hyperskeptic where others, particularly those I call *bildungsphilisters*, are gullible, and gullible where others seem skeptical. I am skeptical about confirmation—though only when errors are costly—not about disconfirmation. Having plenty of data will not provide confirmation, but a single instance can disconfirm. I am skeptical when I suspect wild randomness, gullible when I believe that randomness is mild.

Half the time I hate Black Swans, the other half I love them. I like the randomness that produces the texture of life, the positive accidents, the success of Apelles the painter, the potential gifts you do not have to pay for. Few understand the beauty in the story of Apelles; in fact, most people exercise their error avoidance by repressing the Apelles in them.

Half the time I am hyperconservative in the conduct of my own affairs; the other half I am hyperaggressive. This may not seem exceptional, except that my conservatism applies to what others call risk taking, and my aggressiveness to areas where others recommend caution.

I worry less about small failures, more about large, potentially terminal ones. I worry far more about the “promising” stock market, particularly the “safe” blue chip stocks, than I do about speculative ventures—the former present invisible risks, the latter offer no surprises since you know how volatile they are and can limit your downside by investing smaller amounts.

I worry less about advertised and sensational risks, more about the more vicious hidden ones. I worry less about terrorism than about diabetes, less about matters people usually worry about because they are obvious worries, and more about matters that lie outside our consciousness and common discourse (I also have to confess that I do not worry a lot—I try to worry about matters I can do something about). I worry less about embarrassment than about missing an opportunity.

In the end this is a trivial decision making rule: I am very aggressive when I can gain exposure to positive Black Swans—when a failure would be of small moment—and very conservative when I am under threat from a negative Black Swan. I am very aggressive when an error in a model can benefit me, and paranoid when the error can hurt. This may not be too interesting except that it is exactly what other people do not do. In finance, for instance, people use flimsy theories to manage their risks and put wild ideas under “rational” scrutiny.

Half the time I am intellectual, the other half I am a no-nonsense practitioner. I am no-nonsense and practical in academic matters, and intellectual when it comes to practice.

Half the time I am shallow, the other half I want to avoid shallowness. I am shallow when it comes to aesthetics; I avoid shallowness in the context of risks and returns. My aestheticism makes me put poetry before prose, Greeks before Romans, dignity before elegance, elegance before culture, culture before erudition, erudition before knowledge, knowledge before intellect, and intellect before truth. But only for matters that are Black Swan free. Our tendency is to be very rational, except when it comes to the Black Swan.

Half the people I know call me irreverent (you have read my comments about your local Platonified professors), half call me fawning (you have seen my slavish devotion to Huet, Bayle, Popper, Poincaré, Montaigne, Hayek, and others).

Half the time I hate Nietzsche, the other half I like his prose.

WHEN MISSING A TRAIN IS PAINLESS

I once received another piece of life-changing advice, which, unlike the advice I got from a friend in [Chapter 3](#), I find applicable, wise, and empirically valid. My classmate in Paris, the novelist-to-be Jean-Olivier Tedesco, pronounced, as he prevented me from running to catch a subway, “I don’t run for trains.”

Snub your destiny. I have taught myself to resist running to keep on schedule. This may seem a very small piece of advice, but it registered. In refusing to run to catch trains, I have felt the true value of *elegance* and aesthetics in behavior, a sense of being in control of my time, my schedule, and my life. *Missing a train is only painful if you run after it!* Likewise, not matching the idea of success others expect from you is only painful if that’s what you are seeking.

You stand *above* the rat race and the pecking order, not *outside* of it, if you do so by choice.

Quitting a high-paying position, if it is *your* decision, will seem a better payoff than the utility of the money involved (this may seem crazy, but I’ve tried it and it works). This is the first step toward the stoic’s throwing a four-letter word at fate. You have far more control over your life if you decide on your criterion by yourself.

Mother Nature has given us some defense mechanisms: as in Aesop’s fable, one of these is our ability to consider that the grapes we cannot (or did not) reach are sour. But an aggressively stoic *prior* disdain and rejection of the grapes is even more rewarding. Be aggressive; be the one to resign, if you have the guts.

It is more difficult to be a loser in a game you set up yourself.

In Black Swan terms, this means that you are exposed to the improbable only if you let it control you. You always control what *you* do; so make this your end.

THE END

But all these ideas, all this philosophy of induction, all these problems with knowledge, all these wild opportunities and scary possible losses, everything palls in front of the following metaphysical consideration.

I am sometimes taken aback by how people can have a miserable day or get angry because they feel cheated by a bad meal, cold coffee, a social rebuff, or a rude reception. Recall my discussion in [Chapter 8](#) on the difficulty in seeing the true odds of the events that run your own life. We are quick to forget that just being alive is an extraordinary piece of good luck, a remote event, a chance occurrence of monstrous proportions.

Imagine a speck of dust next to a planet a billion times the size of the earth. The speck of dust represents the odds in favor of your being born; the huge planet would be the odds against it. So stop sweating the small stuff. Don't be like the ingrate who got a castle as a present and worried about the mildew in the bathroom. Stop looking the gift horse in the mouth—remember that you are a Black Swan. And thank you for reading my book.

YEVGENIA'S WHITE SWANS

Yevgenia Krasnova went into the long hibernation that was necessary for producing a new book. She stayed in New York City, where she found it easiest to find tranquillity, alone with her text. It was easiest to concentrate after long periods during which she was surrounded by crowds, hoping to run into Nero so she could make a snide remark to him, perhaps humiliate him, possibly win him back. She canceled her e-mail account, switched to writing longhand, since she found it soothing, and hired a secretary to type her text. She spent eight years writing, erasing, correcting, venting her occasional anger at the secretary, interviewing new secretaries, and quietly rewriting. Her apartment was full of smoke, with papers strewn on every surface. Like all artists she remained dissatisfied with the state of completion of her work, yet she felt that she had gone far deeper than with her first book. She laughed at the public who extolled her earlier work, for she now found it shallow, hurriedly completed, and undistilled.

When the new book, which was aptly called *The Loop*, came out, Yevgenia was wise enough to avoid the press and ignore her reviews, and stayed insulated from the external world. As expected by her publisher, the reviews were laudatory. But, strangely, few were buying. People must be talking about the book without reading it, he thought. Her fans had been waiting for it and talking about it for years. The publisher, who now owned a very large collection of pink glasses and led a flamboyant lifestyle, was presently betting the farm on Yevgenia. He had no other hits and none in sight. He needed to score big to pay for his villa in Carpentras in Provence and his dues on the financial settlement with his estranged wife, as well as to buy a new convertible Jaguar (pink). He had been certain that he had a good shot with Yevgenia's long-awaited book, and he could not figure out why almost everyone called it a masterpiece yet no one was buying it. A year and a half later, *The Loop* was effectively out of print. The publisher, now in severe financial distress, thought he knew the reason: the book was "too f***ing long!"—Yevgenia should have written a shorter one. After a long but soothing

lachrymal episode, Yevgenia thought of the characters in the rainy novels of Georges Simenon and Graham Greene. They lived in a state of numbing and secure mediocrity. Second-rateness had charm, Yevgenia thought, and she had always preferred charm over beauty.

So Yevgenia's second book too was a Black Swan.

Postscript Essay

**ON ROBUSTNESS AND FRAGILITY,
DEEPER PHILOSOPHICAL AND
EMPIRICAL REFLECTIONS**

LEARNING FROM MOTHER NATURE, THE OLDEST AND THE WISEST

How to make friends among walking people—On becoming a grandmother—The charms of eco-Extremistan—Never small enough—Harvard-Soviet chic

I am writing this essay three years after the completion of *The Black Swan*—which I have kept intact except for a few clarifying footnotes. Since then, I’ve written a dozen “scholarly” papers around some aspects of the Black Swan idea. These are very, very boring to read, since almost all academic papers are made to bore, impress, provide credibility, intimidate even, be presented at meetings, but not to be read except by suckers (or detractors) or, even worse, graduate students. Also, I am making the “what to do next” more salient here—you can take a horse to water and, in addition, you may have to make it drink. So this essay will allow me to go deeper into some points. Like the main text itself, the beginning will be what is called literary, and progressively turn technical.

I owe the idea of this book-length essay to Danny Kahneman, toward whom I (and my ideas) have more debt than toward anyone else on this planet. He convinced me that I had obligations to try to make the horse drink.

ON SLOW BUT LONG WALKS

Over the past three years, my life experienced a bit of change, mostly for the better. Like parties, a book puts you on the envelope of serendipity; it even gets you invited to more parties. During my dark days, I was called a trader in Paris (something extremely *vulgaire*), a philosopher in London (meaning too theoretical), a prophet in New York (dissingly, because of my then false prophecy), and an economist in Jerusalem (something very materialistic). I now saw myself dealing with the stress of having to live up to the wholly undeserved designations of a prophet in Israel (a very, very ambitious project), a *philosophe* in France, an economist in London, and a trader in New York (where it is respectable).

Such exposure brought hate mail, at least one death threat (by former employees of the bankrupt firm Lehman Brothers*), which I found extremely flattering, and, worse than any threat of violence, hourly requests for interviews by Turkish and Brazilian journalists. I had to spend a lot of time writing personalized and courteous notes declining invitations to dinner with suit-wearing current hotshots, suit-wearing archeo-hotshots, suit-wearing proto-hotshots, and the nasty brand of suit-wearing namedroppers. But it also brought some benefits. I was contacted by like-minded persons, people I would have never dreamed of meeting in the past, or those I did not think existed before, in disciplines completely outside my normal circles, who helped me further my quest with the most unexpected of ideas. I was often reached by people I admired and whose work I knew well, and who became natural collaborators and critics; I will always remember the thrill of getting an unexpected e-mail from Spyros Makridakis of the M-Competition described in [Chapter 10](#), the great debunker of misforecasting, or another one from Jon Elster, the scholar of rare erudition and insights who integrated the wisdom of the ancients into modern social science thinking. I've met novelists and philosophical thinkers whose works I had read and admired, like Louis de Bernières, Will Self, John Gray (the philosopher, not the pop psychologist), or Lord Martin Rees; in all four cases I had the peculiar need to pinch myself upon hearing them talking to me about *my own* book.

Then, through a chain of friends of friends, cappuccinos, dessert wines, and security lines at airports, I got to partake of and understand the potency of oral knowledge, as discussions are vastly more powerful than just correspondence.

People say things in person they would never put in print. I met Nouriel Roubini (to my knowledge the only professional economist who *really* predicted the crisis of 2008, and perhaps the only independent thinker in that business). I also found a variety of people I did not know existed, *good* economists (i.e., with scientific standards), like Michael Spence and Barkley Rosser. Also Peter Bevelin and Yechezkel Zilber kept feeding me the papers I was looking for without knowing it, the first in biology, the second in cognitive science—thus they nudged my thinking in the appropriate direction.

So I have been dialoguing with many people. My problem is that I found only two persons who can have a conversation during a long walk (and walk slowly): Spyros Makridakis and Yechezkel Zilber. Most people, alas, walk too fast, mistaking walking for exercise, not understanding that walking is to be done slowly, at such a pace that one forgets one is walking—so I need to keep going to Athens (where Spyros lives) in order to indulge in my favorite activity, being a flâneur.

My Mistakes

And of course people will scrutinize the text. After examining messages and reports, I do not feel I need to retract anything in the initial version, or to correct any error (outside of typos and minor factual mistakes), except for two related matters. The first fault was pointed out to me by Jon Elster. I had written that the narrative fallacy pervades historical analyses, since I believed that there was no such thing as a test of a historical statement by forecasting and falsification. Elster explained to me that there are situations in which historical theory can escape the narrative fallacy and be subjected to empirical rejection—areas in which we are discovering documents or archeological sites yielding information capable of countering a certain narrative.

So, in relation to his point, I realized that the history of Arabic thought was not so definitive and that I had fallen into the trap of ignoring the continuous changes in *past* history, that the past too was largely a prediction. I (accidentally) discovered that I had fallen for conventional wisdom in textbook scholarship on Arabic philosophy, a wisdom that was contradicted by existing documents. I had exaggerated the import of the debate between Averroës and Algazel. Like everyone I thought that 1) it was a big deal and, 2) it killed Arabic *falsafah*. It turned out to be one of the mis-conceptions being

recently debunked by researchers (such as Dimitri Gutas and George Saliba). Most of those who theorized about Arabic philosophy did not know Arabic, so they left many things to their imagination (like Leo Strauss, for example). I am a bit ashamed, because Arabic is one of my native languages, and here I was reporting from tenth-hand sources developed by scholars illiterate in Arabic (and sufficiently overconfident and lacking in erudition to not realize it). I fell for the confirmation bias seen by Gutas: “It seems that one always starts with a preconception of what Arabic philosophy should be saying, and then concentrating only on those passages which seem to be supporting such a bias, thereby appearing to corroborate the preconception on the basis of the texts themselves.”

Once again, beware of history.

ROBUSTNESS AND FRAGILITY

Upon the completion of *The Black Swan*, I spent some time meditating on the items I raised in [Chapter 14](#) on the fragility of some systems with large concentration and illusions of stability—which had left me convinced that the banking system was the mother of all accidents waiting to happen. I explained in [Chapter 6](#), with the story of the old elephants, that the best teachers of wisdom are naturally the eldest, simply because they may have picked up invisible tricks and heuristics that escape our epistemic landscape, tricks that helped them survive in a world more complex than the one we think we can understand. So being old implies a higher degree of resistance to Black Swans, though, as we saw with the turkey story, it is not a guaranteed proof—older is almost always more solid, but older is not necessarily perfect. But a few billion years is vastly more proof than a thousand days of survival, and the oldest system around is clearly Mother Nature.

That was, in a way, the reasoning behind the *epilogism* argument of the medical empiricists of the post-classical Levant (like Menodotus of Nicomedia), who were the only practitioners to merge skepticism and decision-making in the real world. They are also the only group of people to use philosophy for anything useful. They proposed *historia*: maximal recording of facts with minimal interpretation and theorizing, describing of facts without the *why*, and resisting universals. Their form of nontheoretical knowledge was degraded by the medieval Scholastics, who favored more explicit learning. *Historia*, just the recording of facts, was inferior to *philosophia* or *scientia*. Even philosophy, until then, had more to do with decision-making wisdom than it does today, not with impressing a tenure committee, and medicine was where such wisdom was practiced (and learned): *Medicina soror philosophiae*: “Medicine, sister of Philosophy.”*

Giving an ancillary status to a field that prefers particulars to universals is what formalized knowledge since the Scholastics has been doing, which necessarily gives short shrift to experience and age (too much accumulation of particulars), in favor of those who hold a PhD like Dr. John. This may work in classical physics, but not in the complex domain; it has killed a lot of patients in the history of medicine, particularly before clinical medicine was born, and is causing a lot of damage in the social domain, particularly at the time of writing.

The central things the old teachers communicate to you are, to use religious terms, dogmas (rules you need to execute without necessarily understanding them) not kerygmas (rules you can understand and that have a purpose clear to you).

Mother Nature is clearly a complex system, with webs of interdependence, nonlinearities, and a robust ecology (otherwise it would have blown up a long time ago). It is an old, very old person with an impeccable memory. Mother Nature does not develop Alzheimer's—actually there is evidence that even humans would not easily lose brain function with age if they followed a regimen of stochastic exercise and stochastic fasting, took long walks, avoided sugar, bread, white rice, and stock market investments, and refrained from taking economics classes or reading such things as *The New York Times*.

Let me summarize my ideas about how Mother Nature deals with the Black Swan, both positive and negative—it knows much better than humans how to take advantage of positive Black Swans.

Redundancy as Insurance

First, *Mother Nature likes redundancies*, three different types of redundancies. The first, the simplest to understand, is defensive redundancy, the insurance type of redundancy that allows you to survive under adversity, thanks to the availability of spare parts. Look at the human body. We have two eyes, two lungs, two kidneys, even two brains (with the possible exception of corporate executives)—and each has more capacity than needed in ordinary circumstances. So redundancy *equals* insurance, and the apparent inefficiencies are associated with the costs of maintaining these spare parts and the energy needed to keep them around in spite of their idleness.

The exact opposite of redundancy is naïve optimization. I tell everyone to avoid attending (orthodox) economics classes and say that economics will fail us and blow us up (and, as we will see, we have proofs that it failed us; but, as I kept saying in the original text, we did not need them; all we needed was to look at the lack of scientific rigor—and of ethics). The reason is the following: It is largely based on notions of naïve optimization, mathematized (poorly) by Paul Samuelson—and this mathematics contributed massively to the construction of an error-prone society. An economist would find it *inefficient* to maintain two lungs and two kidneys: consider the costs involved in

transporting these heavy items across the savannah. Such optimization would, eventually, kill you, after the first accident, the first “outlier.” Also, consider that if we gave Mother Nature to economists, it would dispense with individual kidneys: since we do not need them all the time, it would be more “efficient” if we sold ours and used a central kidney on a time-share basis. You could also lend your eyes at night since you do not need them to dream.

Almost every major idea in conventional economics (though a lesser number of minor ones) fails under the modification of some assumption, or what is called “perturbation,” when you change one parameter, or take a parameter heretofore assumed by the theory to be fixed and stable, and make it random. We call this “randomization” in the jargon. This is called the study of model error and examination of the consequences of such changes (my official academic specialty is now model error or “model risk”). For instance, if a model used for risk assumes that the type of randomness under consideration is from Mediocristan, it will ignore large deviations and encourage the building of a lot of risk that ignores large deviations; accordingly, risk management will be faulty. Hence the metaphor of “sitting on a barrel of dynamite” I used concerning Fannie Mae (now bust).

For another example of egregious model error, take the notion of comparative advantage supposedly discovered by Ricardo and behind the wheels of globalization. The idea is that countries should focus, as a consultant would say, on “what they do best” (more exactly, on where they are missing the smallest number of opportunities); so one country should specialize in wine and the other in clothes, although one of them might be better at both. But do some perturbations and alternative scenarios: consider what would happen to the country specializing in wine if the price of wine fluctuated. Just a simple perturbation around this assumption (say, considering that the price of wine is random, and can experience Extremistan-style variations) makes one reach a conclusion the opposite of Ricardo’s. Mother Nature does not like overspecialization, as it limits evolution and weakens the animals.

This also explains why I found current ideas on globalization (such as those promoted by the journalist Thomas Friedman) one step too naïve, and too dangerous for society—unless one takes into account side effects. Globalization might give the appearance of efficiency, but the operating leverage and the degrees of interaction between parts will cause small cracks

in one spot to percolate through the entire system. The result would be like a brain experiencing an epileptic seizure from too many cells firing at the same time. Consider that our brain, a well-functioning complex system, is not “globalized,” or, at least, not naïvely “globalized.”

The same idea applies to debt—it makes you fragile, very fragile under perturbations, particularly when we switch from the assumption of Mediocristan to that of Extremistan. We currently learn in business schools to engage in borrowing (by the same professors who teach the Gaussian bell curve, that Great Intellectual Fraud, among other pseudosciences), against all historical traditions, when all Mediterranean cultures developed through time a dogma against debt. *Felix qui nihil debet* goes the Roman proverb: “Happy is he who owes nothing.” Grandmothers who survived the Great Depression would have advised the exact opposite of debt: redundancy; they would urge us to have several years of income in cash before taking any personal risk—exactly my barbell idea of [Chapter 11](#), in which one keeps high cash reserves while taking more aggressive risks but with a small portion of the portfolio. Had banks done that, there would have been no bank crises in history.

We have documents since the Babylonians showing the ills of debt; Near Eastern religions banned debt. This tells me that one of the purposes of religion and tradition has been to enforce interdicts—simply to protect people against their own epistemic arrogance. Why? Debt implies a strong statement about the future, and a high degree of reliance on forecasts. If you borrow a hundred dollars and invest in a project, you still owe a hundred dollars even if you fail in the project (but you do a lot better in the event you succeed). So debt is dangerous if you have some overconfidence about the future and are Black Swan blind, which we all tend to be. And forecasting is harmful since people (especially governments) *borrow* in response to a forecast (or use the forecast as a cognitive excuse to borrow). My Scandal of Prediction (i.e., bogus predictions that seem to be there to satisfy psychological needs) is compounded by the Scandal of Debt: borrowing makes you more vulnerable to forecast errors.

Big is Ugly—and Fragile

Second, *Mother Nature does not like anything too big*. The largest land animal is the elephant, and there is a reason for that. If I went on a rampage and shot

an elephant, I might be put in jail, and get yelled at by my mother, but I would hardly disturb the ecology of Mother Nature. On the other hand, my point about banks in [Chapter 14](#)—that if you shot a large bank, I would “shiver at the consequences” and that “if one falls, they all fall”—was subsequently illustrated by events: one bank failure, that of Lehman Brothers, in September 2008, brought down the entire edifice. Mother Nature does not limit the interactions between entities; it just limits the size of its units. (Hence my idea is not to stop globalization and ban the Internet; as we will see, much more stability would be achieved by stopping governments from helping companies when they become large and by giving back advantages to the small guy.)

But there is another reason for man-made structures not to get too large. The notion of “economies of scale”—that companies save money when they become large, hence more efficient—is often, apparently behind company expansions and mergers. It is prevalent in the collective consciousness without evidence for it; in fact, the evidence would suggest the opposite. Yet, for obvious reasons, people keep doing these mergers—they are not good for companies, they are good for Wall Street bonuses; a company getting larger is good for the CEO. Well, I realized that as they become larger, companies appear to be more “efficient,” but they are also much more vulnerable to outside contingencies, those contingencies commonly known as “Black Swans” after a book of that name. All that under the illusion of more stability. Add the fact that when companies are large, they need to optimize so as to satisfy Wall Street analysts. Wall Street analysts (MBA types) will pressure companies to sell the extra kidney and ditch insurance to raise their “earnings per share” and “improve their bottom line”—hence eventually contributing to their bankruptcy.

Charles Tapiero and I have shown mathematically that a certain class of unforeseen errors and random shocks hurts large organisms vastly more than smaller ones. In another paper, we computed the costs to society of such size; don’t forget that companies, when they fall, cost us.

The problem with governments is that they will tend to support these fragile organisms “because they are large employers” and because they have lobbyists, the kind of phony but visible advertised contributions so decried by Bastiat. Large companies get government support and become progressively larger and more fragile, and, in a way, run government, another prophetic view of Karl Marx and Friedrich Engels. Hairdressers and small businesses on

the other hand, fail without anyone caring about them; they need to be efficient and to obey the laws of nature.

Climate Change and “Too Big” Polluters

I have been asked frequently on how to deal with climate change in connection with the Black Swan idea and my work on decision making under opacity. The position I suggest should be based both on ignorance and on deference to the wisdom of Mother Nature, since it is older than us, hence wiser than us, and has been proven much smarter than scientists. We do not understand enough about Mother Nature to mess with her—and I do not trust the models used to forecast climate change. Simply, we are facing nonlinearities and magnifications of errors coming from the so-called butterfly effects we saw in [Chapter 11](#), actually discovered by Lorenz using weather-forecasting models. Small changes in input, coming from measurement error, can lead to massively divergent projections—and that generously assumes that we have the right equations.

We have polluted for years, causing much damage to the environment, while the scientists currently making these complicated forecasting models were not sticking their necks out and trying to stop us from building these risks (they resemble those “risk experts” in the economic domain who fight the previous war)—these are the scientists now trying to impose the solutions on us. But the skepticism about models that I propose does not lead to the conclusions endorsed by anti-environmentalists and pro-market fundamentalists. Quite the contrary: we need to be hyper-conservationists ecologically, since we do not know what we are harming *with* now. That’s the sound policy under conditions of ignorance and epistemic opacity. To those who say “We have no proof that we are harming nature,” a sound response is “We have no proof that we are not harming nature, either;” the burden of the proof is not on the ecological conservationist, but on someone disrupting an old system. Furthermore we should not “try to correct” the harm done, as we may be creating another problem we do not know much about currently.

One practical solution I have come up with, based on the nonlinearities in the damage (under the assumption that harm increases disproportionately with the quantities released), and using the same mathematical reasoning that led to my opposing the “too big” concept, is to spread the damage across pollutants

—should we need to pollute, of course. Let us carry on a thought experiment.

Case 1: You give the patient a dose of cyanide, hemlock, or some poisonous substance, assuming they are equally harmful—and assuming, for the case of this experiment, the absence of super-additivity (that is, no synergetic effects).

Case 2: You give the patient a tenth of a dose of each of ten such substances, for the same total amount of poison.

Clearly we can see that Case 2, by spreading the poison ingested across substances, is at the worst equally harmful (if all the poisonous substances act in the same way), and at the best close to harmless to the patient.

Species Density

Mother Nature does not like too much connectivity and globalization—(biological, cultural, or economic). One of the privileges I got as a result of the book was meeting Nathan Myhrvold, the type of person I wish were cloned so I could have one copy here in New York, one in Europe, and one in Lebanon. I started meeting with him regularly; every single meeting has led to a big idea, or the rediscovery of my own ideas through the brain of a more intelligent person—he could easily claim co-authorship of my next book. The problem is that, unlike Spyros and those very few others, he does not have his conversations while walking (though I met him in excellent restaurants).

Myhrvold enlightened me about an additional way to interpret and prove how globalization takes us into Extremistan: the notion of species density. Simply, larger environments are more scalable than smaller ones—allowing the biggest to get even bigger, at the expense of the smallest, through the mechanism of preferential attachment we saw in [Chapter 14](#). We have evidence that small islands have many more species per square meter than larger ones, and, of course, than continents. As we travel more on this planet, epidemics will be more acute—we will have a germ population dominated by a few numbers, and the successful killer will spread vastly more effectively. Cultural life will be dominated by fewer persons: we have fewer books per reader in English than in Italian (this includes bad books). Companies will be more uneven in size. And fads will be more acute. So will runs on the banks, of course.

Once again, I am not saying that we need to stop globalization and prevent

travel. We just need to be aware of the side effects, the trade-offs—and few people are. I see the risks of a very strange acute virus spreading throughout the planet.

The Other Types of Redundancy

The other categories of redundancy, more complicated and subtle, explain how elements of nature exploit positive Black Swans (and have an additional toolkit for surviving negative ones). I will discuss this very briefly here, as it is mostly behind my next work on the exploitation of Black Swans, through *tinkering* or the domestication of uncertainty.

Functional redundancy, studied by biologists, is as follows: unlike organ redundancy—the availability of spare parts, where the same function can be performed by identical elements—very often the same function can be performed by two different structures. Sometimes the term *degeneracy* is used (by Gerald Edelman and Joseph Gally).

There is another redundancy: when an organ can be employed to perform a certain function that is not its current central one. My friend Peter Bevelin links this idea to the “spandrels of San Marco,” after an essay by Steven Jay Gould. There, the necessary space between arches in the Venetian cathedral of San Marco has led to art that is now central to our aesthetic experience while visiting the place. In what is now called the *spandrel effect*, an auxiliary offshoot of a certain adaptation leads to a new function. I can also see the adaptation as having a dormant potential function that could wake up in the right environment.

The best way to illustrate such redundancy is with an aspect of the life story of the colorful philosopher of science Paul Feyerabend. Feyerabend was permanently impotent from a war injury, yet he married four times, and was a womanizer to the point of leaving a trail of devastated boyfriends and husbands whose partners he snatched, and an equally long one of broken hearts, including those of many of his students (in his day, certain privileges were allowed to professors, particularly flamboyant professors of philosophy). This was a particular achievement given his impotence. So there were other parts of the body that came to satisfy whatever it was that made women attached to him.

Mother Nature initially created the mouth to eat, perhaps to breathe,

perhaps for some other function linked to the existence of the tongue. Then new functions emerged that were most probably not part of the initial plan. Some people use the mouth and tongue to kiss, or to do something more involved to which Feyerabend allegedly had recourse.

Over the past three years I became obsessed with the notion that, under epistemic limitations—some opacity concerning the future—progress (and survival) cannot take place without one of these types of redundancy. You don't know today what may be needed tomorrow. This conflicts very sharply with the notion of teleological design we all got from reading Aristotle, which has shaped medieval Arabic-western thought. For Aristotle, an object had a clear purpose set by its designer. An eye was there to see, a nose to smell. This is a rationalistic argument, another manifestation of what I call Platonicity. Yet anything that has a secondary use, and one you did not pay for, will present an extra opportunity should a heretofore unknown application emerge or a new environment appear. The organism with the largest number of secondary uses is the one that will gain the most from environmental randomness and epistemic opacity!

Take aspirin. Forty years ago, aspirin's *raison d'être* was its antipyretic (fever-reducing) effect. Later it was used for its analgesic (pain-reducing) effect. It has also been used for its anti-inflammatory properties. It is now used mostly as a blood thinner to avoid second (or first) heart attacks. The same thing applies to almost all drugs—many are used for secondary and tertiary properties.

I have just glanced at the desk in my business, nonliterary office (I separate the functional from the aesthetic). A laptop computer is propped up on a book, as I like to have some incline. The book is a French biography of the fiery Lou Andreas Salomé (Nietzsche's and Freud's friend) that I can very safely say I will never read; it was selected for its optimal thickness for the task. This makes me reflect on the foolishness of thinking that books are there to be read and could be replaced by electronic files. Think of the spate of functional redundancies provided by books. You cannot impress your neighbors with electronic files. You cannot prop up your ego with electronic files. Objects seem to have invisible but significant auxiliary functions that we are not aware of consciously, but that allow them to thrive—and on occasion, as with decorator books, the auxiliary function becomes the principal one.

So when you have a lot of functional redundancies, randomness helps on

balance, but under one condition—that you can benefit from the randomness more than you can be hurt by it (an argument I call more technically *convexity to uncertainty*). This is certainly the case with many engineering applications, in which tools emerge from other tools.

Also, I am currently absorbed in the study of the history of medicine, which struggled under this Aristotelian illusion of purpose, with Galen’s rationalistic methods that killed so many people while physicians thought they were curing them. Our psychology conspires: people like to go to a precise destination, rather than face some degree of uncertainty, even if beneficial. And research itself, the way it is designed and funded, seems to be teleological, aiming for precise results rather than looking for maximal exposures to forking avenues.

I have given more complicated names to this idea, in addition to *convexity*, like *optionality*—since you have the option of taking the freebie from randomness—but this is still work in progress for me. The progress coming from the second type of randomness is what I call *tinkering*, or *bricolage*, the subject of my next book.

Distinctions Without a Difference, Differences Without a Distinction

Another benefit of duplication. I have, throughout this book, focused on the absence of practical distinctions between the various notions of luck, uncertainty, randomness, incompleteness of information, and fortuitous occurrences using the simple criterion of predictability, which makes them all functionally equal. Probability can be degrees of belief, what one uses to make a bet, or something more physical associated with true randomness (called “ontic,” on which later). To paraphrase Gerd Gigerenzer, a “50 percent chance of rain tomorrow” in London might mean that it will rain half the day, while in Germany it will mean that half the experts think it will rain, and (I am adding), in Brooklyn, that the betting market at the bar is such that one would pay 50 cents to get a dollar if it rains.

For scientists, the treatment is the same. We use the same equation to describe a probability distribution, regardless of whether the probability is a degree of belief or something designed by Zeus, who, we believe, calls the shots. For us probabilists (persons who work with probability in a scientific context), the probability of an event, however it may be defined, is, simply, a weight between 0 and 1, called the measure of the set concerned. Giving

different names and symbols would be distracting and would prevent the transfer of analytical results from one domain to another.

For a philosopher, it is altogether another matter. I had two lunches with the (analytical) philosopher Paul Boghossian, three years apart, one upon the completion of the first edition of *The Black Swan*, the second upon the completion of this essay. During the first conversation he said that, from a philosophical point of view, it is a mistake to conflate probability as a measure of someone's rational degree of belief with probability as a property of events in the world. To me, this implied that we should not use the same mathematical language, say, the same symbol, p , and write down the same equation for the different types of probabilities. I spent three years wondering if he was right or wrong, whether this was a *good redundancy*. Then I had lunch with him again, though in a better (and even more friendly) restaurant.

He alerted me to a phrase philosophers use: "distinction without a difference." Then I realized the following: that there are distinctions philosophers use that make sense philosophically, but do not seem to make sense in practice, but that may be necessary if you go deeper into the idea, and may make sense in practice under a change of environment.

For consider the opposite: differences without a distinction. They can be brutally misleading. People use the same term, *measuring*, for measuring a table using a ruler, and for measuring risk—when the second is a forecast, or something of the sort. And the word *measuring* conveys an illusion of knowledge that can be severely distorting: we will see that we are psychologically very vulnerable to terms used and how things are framed. So if we used *measuring* for the table, and *forecasting* for risk, we would have fewer turkeys blowing up from Black Swans.

Mixing vocabulary has been very common in history. Let me take the idea of chance again. At some point in history the same Latin word, *felix* (from *felicitas*) was used to designate both someone lucky and someone happy. (The conflation of happiness and luck was explainable in an antique context: the goddess Felicitas represented both.) The English word *luck* comes from the Germanic *Glück*, happiness. An ancient would have seen the distinction between the two concepts as a waste, since all lucky people seem happy (not thinking that one could be happy without being lucky). But in a modern context we need to extricate luck from happiness—utility from probability—in order to perform any psychological analysis of decision making. (True, it is

hard to disentangle the two from observing people making decisions in a probabilistic environment. People may be so fearful of bad things that may happen to them that they tend to overpay for insurance, which in turn may make us mistakenly think that they believe the adverse event has a high probability.) So we can see now that the absence of such precision made the language of the ancients quite confusing to us; but to the ancients, the distinction would have been a redundancy.

A SOCIETY ROBUST TO ERROR

I will only very briefly discuss the crisis of 2008 (which took place after the publication of the book, and which was a lot of things, but *not* a Black Swan, only the result of fragility in systems built upon ignorance—and denial—of the notion of Black Swan events. You know with near certainty that a plane flown by an incompetent pilot will eventually crash).

Why briefly? *Primo*, this is not an economics book, but a book on the incompleteness of knowledge and the effects of high-impact uncertainty—it just so happens that economists are the most Black-Swan-blind species on the planet. *Secundo*, I prefer to talk about events *before* they take place, not *after*. But the general public confuses the prospective with the retrospective. The very same journalists, economists, and political experts who did not see the crisis coming provided abundant ex-post analyses about its inevitability. The other reason, the real one, is that the crisis of 2008 was not intellectually interesting enough to me—there is nothing in the developments that had not happened before, at a smaller scale (for example, banks losing in 1982 every penny they ever made). It was merely a financial opportunity for me, as I will discuss further down. Really, I reread my book and saw nothing to add to the text, nothing we had not already encountered at some point in history, like the earlier debacles, nothing I had learned from. Alas, nothing.

The corollary is obvious: since there is nothing new about the crisis of 2008, we will not learn from it and we will make the same mistake in the future. And the evidence is there at the time of writing: the IMF continues to issue forecasts (not realizing that previous ones did not work and that the poor suckers relying on them are—once again—going to get in trouble); economics professors still use the Gaussian; the current administration is populated with those who are bringing model error into industrial proportion, making us rely on models even more than ever before.*

But the crisis provides an illustration for the need for robustness, worth discussing here.

Over the past twenty-five hundred years of recorded ideas, only fools and Platonists (or, worse, the species called central bankers) have believed in engineered utopias. We will see in the section on the Fourth Quadrant that the idea is not to correct mistakes and eliminate randomness from social and

economic life through monetary policy, subsidies, and so on. *The idea is simply to let human mistakes and miscalculations remain confined*, and to prevent their spreading through the system, as Mother Nature does. Reducing volatility and ordinary randomness increases exposure to Black Swans—it creates an artificial quiet.

My dream is to have a true Epistemocracy—that is, a society robust to expert errors, forecasting errors, and hubris, one that can be resistant to the incompetence of politicians, regulators, economists, central bankers, bankers, policy wonks, and epidemiologists. We cannot make economists more scientific; we cannot make humans more rational (whatever that means); we cannot make fads disappear. The solution is somewhat simple, once we isolate harmful errors, as we will see with the Fourth Quadrant.

So I am currently torn between (a) my desire to spend time mulling my ideas in European cafés and in the tranquility of my study, or looking for someone who can have a conversation while walking slowly in a nice urban setting, and (b) the feeling of obligation to engage in activism to robustify society, by talking to uninteresting people and being immersed in the cacophony of the unaesthetic journalistic and media world, going to Washington to watch phonies in suits walking around the streets, having to defend my ideas while making an effort to be smooth and hide my disrespect. This proved to be very disruptive to my intellectual life. But there are tricks. One useful trick, I discovered, is to avoid listening to the question of the interviewer, and answer with whatever I have been thinking about recently. Remarkably, neither the interviewers nor the public notices the absence of correlation between question and answer.

I was once selected to be one of a group of a hundred who went to Washington to spend two days discussing how to solve the problems of the crisis that started in 2008. Almost all the biggies were included. After an hour of meeting, and during a speech by the prime minister of Australia, I walked out of the room because my pain became intolerable. My back would start hurting upon looking at the faces of these people. The center of the problem is that none of them knew the center of the problem.

This makes me convinced that there is a unique solution for the world, to be designed along very simple lines of robustness to Black Swans—it will explode otherwise.

So now I am disengaged. I am back in my library. I am not even experiencing any frustration, I don't even care about how forecasters can blow up society, and I am not even capable of being annoyed by fools of randomness (to the contrary), perhaps thanks to another discovery linked to a particular application of the study of complex systems, Extremistan, and that science of long walks.

-
- * Lehman Brothers was a financial institution with great-looking offices that abruptly went bust during the crisis of 2008.
 - * Empiricism is not about not having theories, beliefs, and causes and effects: it is about avoiding being a sucker, having a decided and preset bias about where you want your error to be—where the default is. An empiricist facing series of facts or data defaults to suspension of belief (hence the link between empiricism and the older skeptical Pyrrhonian tradition), while others prefer to default to a characterization or a theory. The entire idea is to avoid the *confirmation bias* (empiricists prefer to err on the side of the disconfirmation/falsification bias, which they discovered more than fifteen hundred years before Karl Popper).
 - * Clearly the entire economics establishment, with about a million people on the planet involved in some aspect of economic analysis, planning, risk management, and forecasting, turned out to be turkeys owing to the simple mistake of not understanding the structure of Extremistan, complex systems, and hidden risks, while relying on idiotic risk measures and forecasts—all this in spite of past experience, as these things have never worked before.

WHY I DO ALL THIS WALKING, OR HOW SYSTEMS BECOME FRAGILE

*Relearn to walk—Temperance, he knew not—Will I catch Bob Rubin? Extremistan and Air France
travel*

ANOTHER FEW BARBELLS

Again, thanks to the exposure the book has received, I was alerted to a new aspect of robustness in complex systems ... by the most unlikely of sources. The idea came from two fitness authors and practitioners who integrated the notions of randomness and Extremistan (though of the Gray Swan variety) into our understanding of human diet and exercise. Curiously, the first person, Art De Vany, is the same one who studied Extremistan in the movies (in [Chapter 3](#)). The second, Doug McGuff, is a physician. And both can talk about fitness, particularly Art, who, at seventy-two, looks like what a Greek god would like to look like at forty-two. Both were referring to the ideas of *The Black Swan* in their works and connecting to it; and I had no clue.

I then discovered to my great shame the following. I had spent my life thinking about randomness; I had written three books on dealing with randomness (one technical); I was prancing about as the expert in the subject of randomness from mathematics to psychology. And I had missed something central: living organisms (whether the human body or the economy) *need* variability and randomness. What's more, they need the Extremistan type of variability, certain extreme stressors. Otherwise they become fragile. That, I completely missed.*

Organisms need, to use the metaphor of Marcus Aurelius, to turn obstacles into fuel—just as fire does.

Brainwashed by the cultural environment and by my education, I was under the illusion that steady exercise and steady nutrition were a good thing for one's health. I did not realize that I was falling into evil rationalistic arguments, the Platonic projection of wishes into the world. Worse, I had been brainwashed though I had all the facts in my head.

From predator-prey models (the so-called Lotka-Volterra type of population dynamics), I knew that populations will experience Extremistan-style variability, hence predators will necessarily go through periods of feast and famine. That's us, humans—we had to have been designed to experience extreme hunger and extreme abundance. So our food intake had to have been fractal. Not a single one of those promoting the “three meals a day,” “eat in moderation” idea has tested it empirically to see whether it is healthier than intermittent fasts followed by large feasts.†

But Near Eastern religions (Judaism, Islam, and Orthodox Christianity) knew it, of course—just as they knew the need for debt avoidance—and so they had fast days.

I also knew that the size of stones and trees was, up to a point, fractal (I even wrote about that in [Chapter 16](#)). Our ancestors mostly had to face very light stones to lift, mild stressors; once or twice a decade, they encountered the need to lift a huge stone. So where on earth does this idea of “steady” exercise come from? Nobody in the Pleistocene jogged for forty-two minutes three days a week, lifted weights every Tuesday and Friday with a bullying (but otherwise nice) personal trainer, and played tennis at eleven on Saturday mornings. Not hunters. We swung between extremes: we sprinted when chased or when chasing (once in a while in an extremely exerting way), and walked about aimlessly the rest of the time. Marathon running is a modern abomination (particularly when done without emotional stimuli).

This is another application of the barbell strategy: plenty of idleness, some high intensity. The data shows that long, very long walks, combined with high-intensity exercise outperform just running.

I am not talking about “brisk walks” of the type you read about in the Health section of *The New York Times*. I mean walking without making any effort.

What’s more, consider the negative correlation between caloric expenditure and intake: we hunted in response to hunger; we did not eat breakfast to hunt, hunting accentuated our energy deficits.

If you deprive an organism of stressors, you affect its epigenetics and gene expression—some genes are up-regulated (or down-regulated) by contact with the environment. A person who does not face stressors will not survive should he encounter them. Just consider what happens to someone’s strength after he spends a year in bed, or someone who grows up in a sterile environment and then one day takes the Tokyo subway, where riders are squeezed like sardines.

Why am I using evolutionary arguments? Not because of the optimality of evolution, but entirely for epistemological reasons, how we should deal with a complex system with opaque causal links and complicated interactions. Mother Nature is not perfect, but has so far proven smarter than humans, certainly much smarter than biologists. So my approach is to combine evidence-based research (stripped of biological theory), with an a priori that Mother Nature has more authority than anyone.

After my “Aha!” flash, I embarked on an Extremistan barbell lifestyle under the guidance of Art De Vany: long, very long, slow, meditative (or conversational) walks in a stimulating urban setting, but with occasional (and random) very short sprints, during which I made myself angry imagining I was chasing the bankster Robert Rubin with a big stick, trying to catch him and bring him to human justice. I went to weight-lifting facilities in a random way for a completely stochastic workout—typically in hotels, when I was on the road. Like Gray Swan events, these were very, very rare, but highly consequential weight-lifting periods, after a day of semistarvation, leaving me completely exhausted. Then I would be totally sedentary for weeks and hang around cafés. Even the duration of the workouts remained random—but most often very short, less than fifteen minutes. I followed the path that minimized boredom, and remained very polite with gym employees who described my workouts as “erratic.” I put myself through thermal variability as well, exposed, on occasion, to extreme cold without a coat. Thanks to transcontinental travel and jet lag, I underwent periods of sleep deprivation followed by excessive rest. When I went to places with good restaurants, for instance Italy, I ate in quantities that would have impressed Fat Tony himself, then skipped meals for a while without suffering. Then, after two and a half years of this apparently “unhealthy” regimen, I saw significant changes in my own physique on every possible criterion—the absence of unnecessary adipose tissue, the blood pressure of a twenty-one-year-old, and so on. I also have a clearer, much more acute mind.

So the main idea is to trade duration for intensity—for a hedonic gain. Recall the reasoning I presented in [Chapter 6](#) about hedonic effects. Just as people prefer large but sudden losses to small but regular ones, just as one becomes dulled to pain beyond a certain threshold, so unpleasant experiences, like working out without external stimuli (say in a gym), or spending time in New Jersey, need to be as concentrated and as intense as possible.

Another way to view the connection to the Black Swan idea is as follows. Classical thermodynamics produces Gaussian variations, while informational variations are from Extremistan. Let me explain. If you consider your diet and exercise as simple energy deficits and excesses, with a straight calorie-in, calorie-burned equation, you will fall into the trap of misspecifying the system into simple causal and mechanical links. Your food intake becomes the equivalent of filling up the tank of your new BMW. If, on the other hand, you

look at food and exercise as activating metabolic signals, with potential metabolic cascades and nonlinearities from network effects, and with recursive links, then welcome to complexity, hence Extremistan. Both food and workouts provide your body with information about stressors in the environment. As I have been saying throughout, informational randomness is from Extremistan. Medicine fell into the trap of using simple thermodynamics, with the same physics envy, and the same mentality, and the same tools as economists did when they looked at the economy as a web of simple links.* And both humans and societies are complex systems.

But these lifestyle ideas do not come from mere self-experimentation or some quack theory. All the results were completely expected from the evidence-based, peer-reviewed research that is available. Hunger (or episodic energy deficit) strengthens the body and the immune system and helps rejuvenate brain cells, weaken cancer cells, and prevent diabetes. It was just that the current thinking—in a way similar to economics—was out of sync with the empirical research. I was able to re-create 90 percent of the benefits of the hunter-gatherer lifestyle with minimal effort, without compromising a modern lifestyle, in the aesthetics of an urban setting (I get extremely bored in nature and prefer walking around the Jewish quarter of Venice to spending time in Bora Bora).*

By the same argument we can lower 90 percent of Black Swan risks in economic life ... by just eliminating speculative debt.

The only thing currently missing from my life is panic, from, say, finding a gigantic snake in my library, or watching the economist Myron Scholes, armed to the teeth, walk into my bedroom in the middle of the night. I lack what the biologist Robert Sapolsky calls the beneficial aspect of acute stress, compared to the deleterious one of dull stress—another barbell, for no stress plus a little bit of extreme stress is vastly better than a little bit of stress (like mortgage worries) all the time.

Some have argued that my health benefits come from long walks, about ten to fifteen hours a week (though nobody has explained to me why they would count as workouts since I walk slowly), while others claim that they come from my few minutes of sprinting; I've had the same problem explaining the inseparability of the two extremes as I did explaining economic deviations. If you have acute stressors, then periods of rest, how can you separate the stressors from the recovery? Extremistan is characterized by both polar

extremes, a high share of low impact, a low share of high impact. Consider that the presence of concentration, here energy expenditure, necessitates that a high number of observations do not contribute to anything except to the dilution. Just as the condition that makes market volatility explained by bursts (say one day in five years represents half the variance) requires that most other days remain exceedingly quiet. If one in a million authors makes half the sales, you need a lot of authors to sell no books.

This is the turkey trap I will discuss later: philistines (and Federal Reserve chairpersons) mistake periods of low volatility (caused by stabilization policies) for periods of low risk, not for switches into Extremistan.

Welcome to Gray Extremistan. Do not tamper too much with the complex system Mother Nature gave you: your body.

Beware Manufactured Stability

By a variant of the same reasoning we can see how the fear of volatility I mentioned earlier, leading to interference with nature so as to impose “regularity,” makes us more fragile across so many domains. Preventing small forest fires sets the stage for more extreme ones; giving out antibiotics when it is not very necessary makes us more vulnerable to severe epidemics—and perhaps to that big one, the grand infection that will be resistant to known antibiotics and will travel on Air France.

Which brings me to another organism: economic life. Our aversion to variability and desire for order, and our acting on those feelings, have helped precipitate severe crises. Making something artificially bigger (instead of letting it die early if it cannot survive stressors) makes it more and more vulnerable to a very severe collapse—as I showed with the Black Swan vulnerability associated with an increase in size. Another thing we saw in the 2008 debacle: the U.S. government (or, rather, the Federal Reserve) had been trying for years to iron out the business cycle, leaving us exposed to a severe disintegration. This is my argument against “stabilization” policies and the manufacturing of a nonvolatile environment. More on that, later. Next, I will discuss a few things about the Black Swan idea that do not appear to easily penetrate consciousness. Predictably.

* There is a difference between stressors and toxic exposure that weakens organisms, like the radiation I

discussed in [Chapter 8](#) with the story of the rats.

- † There is a sociology-of-science dimension to the problem. The science writer Gary Taubes has convinced me that the majority of dietary recommendations (about lowering fats in diets) stand against the evidence. I can understand how one can harbor beliefs about natural things without justifying them empirically; I fail to understand beliefs that contravene both nature and scientific evidence.
- * The financial equations used by the villains for the “random walk” is based on heat diffusion.
- * The argument often heard about primitive people living *on average* less than thirty years ignores the distribution around that average; life expectancy needs to be analyzed conditionally. Plenty died early, from injuries; many lived very long—and healthy—lives. This is exactly the elementary “fooled by randomness” mistake, relying on the notion of “average” in the presence of variance, that makes people underestimate risks in the stock market.

MARGARITAS ANTE PORCOS*

How to not sell books in airports—Mineral water in the desert—How to denigrate other people's ideas and succeed at it

Let me start again. *The Black Swan* is about consequential epistemic limitations, both psychological (hubris and biases) and philosophical (mathematical) limits to knowledge, both individual and collective. I say “consequential” because the focus is on impactful rare events, as our knowledge, both empirical and theoretical, breaks down with those—the more remote the events, the less we can forecast them, yet they are the most impactful. So *The Black Swan* is about human error in some domains, swelled by a long tradition of scientism and a plethora of information that fuels confidence without increasing knowledge. It covers the expert problem—harm caused by reliance on scientific-looking charlatans, with or without equations, or regular noncharlatanic scientists with a bit more confidence about their methods than the evidence warrants. The focus is in not being the turkey in places where it matters, though there is nothing wrong in being a fool where that has no effect.

MAIN ERRORS IN UNDERSTANDING THE MESSAGE

I will briefly state some of the difficulties in understanding the message and the ideas of this book, typically perpetrated by professionals, though, surprisingly, less by the casual reader, the amateur, my friend. Here is a list.

1) Mistaking the Black Swan (capitalized) for the logical problem. (Mistake made by U.K. intellectuals—intellectuals in other countries do not know enough analytical philosophy to make that mistake.)*

2) Saying the maps we had were better than having no maps. (People who do not have experience in cartography, risk “experts,” or worse, employees of the Federal Reserve Bank of the United States.)

This is the strangest of errors. I know few people who would board a plane heading for La Guardia airport in New York City with a pilot who was using a map of Atlanta’s airport “because there is nothing else.” People with a functioning brain would rather drive, take the train, or stay home. Yet once they get involved in economics, they all prefer professionally to use in Extremistan the measures made for Mediocristan, on the ground that “we have nothing else.” The idea, well accepted by grandmothers, that one should pick a destination for which one has a good map, not travel and then find “the best” map, is foreign to PhDs in social science.

3) Thinking that a Black Swan should be a Black Swan to all observers. (Mistake made by people who have not spent a lot of time in Brooklyn and lack the street smarts and social intelligence to realize that *some* people are suckers.)

4) Not understanding the value of negative advice (“Don’t do”) and writing to me to ask me for something “constructive” or a “next step.” (Mistake usually made by chairmen of large companies and those who would like someday to become such chairmen.)†

5) Not understanding that doing nothing can be much more preferable to doing something potentially harmful. (Mistake made by most people who are not grandmothers.)

6) Applying to my ideas labels (*skepticism, fat tails, power laws*) off a supermarket shelf and equating those ideas with inadequate research traditions (or, worse, claiming it was dealt with by “modal logic,” “fuzzy logic,” or whatever the person has vaguely heard of). (Mistake made by those with

graduate degrees from both coasts.)

7) Thinking that *The Black Swan* is about the errors of using the bell curve, which supposedly everyone knew about, and that the errors can be remedied by substituting a number from the Mandelbrotian in place of another. (Mistake made by the pseudoscientific brand of tenured finance professors, like, say, Kenneth French.)

8) Claiming that “we knew all this” and “there is nothing new” in my idea during 2008, then, of course, going bust during the crisis. (Mistake made by the same type of tenured finance professors as before, but these went to work on Wall Street, and are now broke.)

9) Mistaking my idea for Popper’s notion of falsification—or taking any of my ideas and fitting them in a prepackaged category that sounds familiar. (Mistakes mostly made by sociologists, Columbia University political science professors, and others trying to be multidisciplinary intellectuals and learning buzzwords from Wikipedia.)

10) Treating probabilities (of future states) as measurable, like the temperature or your sister’s weight. (People who did a PhD at MIT or something like that, then went to work somewhere, and now spend time reading blogs.)

11) Spending energy on the difference between ontic and epistemic randomness—true randomness, and randomness that arises from incomplete information—instead of focusing on the more consequential difference between Mediocristan and Extremistan. (People with no hobby, no personal problems, no love, and too much free time.)

12) Thinking I am saying “Do not forecast” or “Do not use models,” rather than “Do not use sterile forecasts with huge error” and “Do not use models in the Fourth Quadrant.” (Mistake made by most people who forecast for a living.)

13) Mistaking what I say for “S**t happens” rather than “this is where s**t happens.” (Many former bonus earners.)*

Indeed, the intelligent, curious, and open-minded amateur is my friend. A pleasant surprise for me was to discover that the sophisticated amateur who uses books for his own edification, and the journalist (unless, of course, he was employed by *The New York Times*), could understand my idea much better than professionals. Professional readers, less genuine, either read too

quickly or have an agenda. When reading for “work” or for the purpose of establishing their status (say, to write a review), rather than to satisfy a genuine curiosity, readers who have too much baggage (or perhaps not enough) tend to read rapidly and efficiently, scanning jargon terms and rapidly making associations with prepackaged ideas. This resulted early on in the squeezing of the ideas expressed in *The Black Swan* into a commoditized well-known framework, as if my positions could be squeezed into standard skepticism, empiricism, essentialism, pragmatism, Popperian falsificationism, Knightian uncertainty, behavioral economics, power laws, chaos theory, *etc.* But the amateurs saved my ideas. Thank you, reader.

As I wrote, missing a train is only painful if you are running after it. I was not looking to have a bestseller (I thought I had already had one with my previous book, and just wanted to produce a real thing), so I had to deal with a spate of harrying side effects. I watched as the book was initially treated, owing to its bestseller status, like the nonfiction “idea books,” journalistic through and through, castrated by a thorough and “competent” copy editor, and sold in airports to “thinking” businessmen. Giving these enlightened *Bildungsphilisters*, commonly called idea-book readers, a real book is like giving vintage Bordeaux to drinkers of Diet Coke and listening to their comments about it. Their typical complaint is that they want diet-book-style “actionable steps” or “better forecasting tools,” satisfying the profile of the eventual Black Swan victim. We will see further that, in an ailment similar to confirmation bias, charlatans provide the much demanded positive advice (what to do), as people do not value negative advice (what not to do). Now, “how not to go bust” does not appear to be valid advice, yet, given that over time only a minority of companies do not go bust, avoiding death is the best possible—and most robust—advice. (It is particularly good advice after your competitors get in trouble and you can go on legal pillages of their businesses.)* Also, many readers (say, those who work in forecasting or banking) do not often understand that the “actionable step” for them is to simply quit their profession and do something more ethical.

In addition to playing into our mental biases, and telling people what they want to hear, these “idea books” often have an abhorrent definitive and investigative tone to their messages, like the reports of management consultants trying to make you believe that they told you more than they

actually did. I came up with a simple compression test using a version of what is called Kolmogorov complexity, a measure of how much a message can be reduced without losing its integrity: try to reduce a book to the shortest length possible without losing any of its intended message or aesthetic effects. A Swiss friend (he does not seem to like to walk slowly and drags me on hikes in the Alps), an owner of a firm that abstracts books and sells the summaries to busy businesspeople, convinced me that his firm has a lofty mission, as almost all business books can be reduced to a few pages without any loss of their message and essence; novels and philosophical treatments cannot be compressed.

So a philosophical essay is a beginning, not an end. To me the very same meditation continues from book to book, compared to the work of a nonfiction writer, who will, say, move to another distinct and journalistically confined topic. I want my contribution to be a new way of viewing knowledge, the very beginning of a long investigation, the start of something real. Indeed, I am glad at the time of writing, a few years into the life of the book, to see the idea spread among thoughtful readers, inspiring like-minded scholars to go beyond it and seeding research in epistemology, engineering, education, defense, operations research, statistics, political theory, sociology, climate studies, medicine, law, aesthetics, and insurance (though not so much in the area in which *The Black Swan* found Black Swan-style near instant vindication, economics).

I was lucky that it only took a couple of years (and a severe financial crisis) for the Republic of Letters to realize that *The Black Swan* was a philosophical tale.

How to Expunge One's Crimes

My ideas went through two distinctive stages after the release of the book. In the first, as the book hit the bestseller list in almost every single country where it was published, many social scientists and finance practitioners fell into the trap of refuting me with the sole argument that I was selling too many books and that my book was accessible to readers; hence it could not reflect original and systematic thought, it was just a “popularization,” not worth reading let alone commenting upon.

The first change of regime came with the release of my more difficult

mathematical, empirical, and scholarly work in a dozen articles in a variety of journals in an attempt to expiate my crime of having sold too many books.* Then, silence.

Still no refutation at the time of this writing; indeed, my paper on the Fourth Quadrant in the *International Journal of Forecasting* (which I simplify in this essay) produced incontrovertible evidence that most (perhaps all) “rigorous” papers in economics using fancy statistics are just hot air, partaking of a collective scam (with diffusion of responsibility), unusable for any form of risk management. Clearly, so far, in spite of a few smear campaigns, or, rather, attempts at a smear campaign (typically conducted by former Wall Street persons or Diet Coke drinkers), nobody has managed to present a formal (or even informal) refutation of the idea—neither of the logical-mathematical arguments nor of the empirical arguments.

But meanwhile I figured out something valuable in the packaging of the Black Swan idea. Just as in *Fooled by Randomness* I had argued (initially from personal experience) that a “70 percent chance of survival” is vastly different from a “30 percent chance of death,” I found out that telling researchers “This is where your methods work very well” is vastly better than telling them “This is what you guys don’t know.” So when I presented to what was until then the most hostile crowd in the world, members of the American Statistical Association, a map of the four quadrants, and told them: your knowledge works beautifully in these three quadrants, but beware of the fourth one, as this is where the Black Swans breed, I received instant approval, support, offers of permanent friendship, refreshments (Diet Coke), invitations to come present at their sessions, even hugs. Indeed, that is how a series of research papers started using my work on where the Fourth Quadrant is located, *etc.* They tried to convince me that statisticians were not responsible for these aberrations, which come from people in the social sciences who apply statistical methods without understanding them (something I verified later, in formal experiments, to my great horror, as we will see further down).

The second change of regime came with the crisis of 2008. I kept getting invited to debates, but I stopped obliging, as it became hard for me to hear complicated arguments and restrain a smile, sometimes a smirk. Why a smile? Well, the vindication. Not the intellectual vindication of winning an argument, no: academia, I discovered, does not change its mind voluntarily, except

perhaps in some real sciences such as physics. It was a different feeling: it is hard to focus on a conversation, especially when it is mathematical, when you have just personally earned several hundreds of times the annual salary of the researcher trying to tell you that you are “wrong,” by betting against his representation of the world.

A Desert Crossing

For I had undergone a difficult psychological moment, after the publication of *The Black Swan*, what the French call *traversée du désert*, when you go through the demoralizing desiccation and disorientation of crossing a desert in search of an unknown destination, or a more or less promised land. I had a rough time, shouting “Fire! Fire! Fire!” about the hidden risks in the system, and hearing people ignore the content and instead just criticize the presentation, as if they were saying “your diction in shouting ‘Fire!’ is bad.” For example, the curator of a conference known as TED (a monstrosity that turns scientists and thinkers into low-level entertainers, like circus performers) complained that my presentation style did not conform to his taste in slickness and kept my lecture on Black Swans and fragility off the Web. Of course, he subsequently tried to claim credit for my warnings voiced before the crisis of 2008.*

Most of the arguments offered were that “times are different,” invoking “the great moderation” by one Ben Bernanke (chairman of the Federal Reserve at the time of writing) who fell for the turkey-before-Thanksgiving trap of not understanding that moving into Extremistan comes through a drop in daily volatility.

Also when I was railing against models, social scientists kept repeating that they knew it and that there is a saying, “all models are wrong, but some are useful”—not understanding that the real problem is that “some are harmful.” Very harmful. As Fat Tony would say, “Tawk is cheap.” So Mark Spitznagel and I restarted the business of “robustifying” clients against the Black Swan (helping people get closer to the barbell of [Chapter 11](#)). We were convinced that the banking system was going to collapse under the weight of hidden risks—that such an event would be a white swan. It was moving from gray to white in color as the system was accumulating risks. The longer we had to wait for it, the more severe it would be. The collapse took place about a year

and a half after the publication of the book. We had been expecting it and betting against the banking system for a long time (and protecting clients by making them Black Swan robust), but the reception of the Black Swan—and the absence of refutation that was not *ad hominem*—made us vastly more worried about the need for protection than ever before.

Like Antaeus, who lost strength when separated from contact with the earth, I needed connection to the real world, something real and applied, instead of focusing on winning arguments and trying to convince people of my point (people are almost always only convinced of what they already know). Sticking my neck out in the real world, lining up my life with my ideas by getting involved in trading, had a therapeutic effect, even apart from the vindication; just having a trade on the books gave me strength to not care. A few months before the onset of the crisis of 2008, I was attacked at a party by a Harvard psychologist who, in spite of his innocence of probability theory, seemed to have a vendetta against me and my book. (The most vicious and bitter detractors tend to be those with a competing product on the bookstore shelves.) Having a trade on allowed me to laugh at him—or, what is even worse, made me feel some complicity with him, thanks to his anger. I wonder what would have happened to the psychological state of another author, identical to me in all respects except that he had no involvement with trading and risk taking. When you walk the walk, whether successful or not, you feel more indifferent and robust to people's opinion, freer, more real.

Finally, I got something out of my debates: the evidence that Black Swan events are largely caused by people using measures way over their heads, instilling false confidence based on bogus results. In addition to my befuddlement concerning why people use measures from Mediocristan outside those measures' applicability, and believe in them, I had the inkling of a much larger problem: that almost none of the people who worked *professionally* with probabilistic measures knew what they were talking about, which was confirmed as I got into debates and panels with many hotshots, at least four with "Nobels" in economics. Really. And this problem was measurable, very easily testable. You could have finance "quants," academics, and students use and write papers and papers using the notion of "standard deviation," yet not understand intuitively what it meant, so you could trip them up by asking them elementary questions about the nonmathematical, real conceptual meaning of their numbers. And trip them up we did. Dan Goldstein and I ran experiments on professionals using probabilistic tools, and were shocked to

see that up to 97 percent of them failed elementary questions.* Emre Soyer and Robin Hogarth subsequently took the point and tested it in the use of an abhorrent field called econometrics (a field that, if any scientific scrutiny was applied to it, would not exist)—again, most researchers don’t understand the tools they are using.

Now that the book’s reception is off my chest, let us move into more analytical territory.

* In Latin: “pearls before swine.”

* Most intellectuals keep attributing the black swan expression to Popper or Mill, sometimes Hume, in spite of the quote by Juvenal. The Latin expression *niger cygnus* might even be more ancient, possibly of Etruscan origin.

† One frequent confusion: people believe that I am suggesting that agents should bet on Black Swans taking place, when I am saying that they should avoid blowing up should a Black Swan take place. As we will see in section IV, I am advocating omission, not commission. The difference is enormous, and I have been completely swamped by people wondering if one can “bleed to death” making bets on the occurrence of Black Swans (like Nero, Giovanni Drogo, or the poor scientist with a rich brother-in-law). These people have made their choice for existential reasons, not necessarily economic ones, although the economics of such a strategy makes sense for a collective.

* If most of the people mixed up about the message appear to be involved in economics and social science, while a much smaller share of readers come from those segments, it is because other members of society without such baggage get the book’s message almost immediately.

* For instance, one anecdote that helps explain the crisis of 2008. One Matthew Barrett, former Chairman of Barclays Bank and Bank of Montreal (both of which underwent blowups from exposures to Extremistan using risk management methods for Mediocristan) complained, after all the events of 2008 and 2009, that *The Black Swan* did not tell him “what should I do about that?” and he “can’t run a business” worrying about Black Swan risks. The person has never heard of the notion of fragility and robustness to extreme deviations—which illustrates my idea that evolution does not work by teaching, but destroying.

* So far, about fourteen scholarly (but very, very boring) articles. (They are boring both to read and to write!) The number keeps growing, though, and they are being published at a pace of three a year. Taleb (2007), Taleb and Pilpel (2007), Goldstein and Taleb (2007), Taleb (2008), Taleb (2009), Taleb, Goldstein and Spitznagel (2009), Taleb and Pilpel (2009), Mandelbrot and Taleb (2010), Makridakis and Taleb (2010), Taleb (2010), Taleb and Tapiero (2010a), Taleb and Tapiero (2010b), Taleb and Douady (2010), and Goldstein and Taleb (2010).

* Although his is a bit extreme, this phoniness is not uncommon at all. Many intellectually honest people I had warned, and who had read my book, later blamed me for not telling them about the crisis—they just could not remember it. It is hard for a newly enlightened pig to recall that he has seen a pearl in the past but did not know what it was.

* Dan Goldstein and I have been collaborating and running experiments about human intuitions with respect to different classes of randomness. He does not walk slowly.

ASPERGER AND THE ONTOLOGICAL BLACK SWAN

Are nerds more blind to swans? Social skills in Extremistan—On the immortality of Dr. Greenspan

If *The Black Swan* is about epistemic limitations, then, from this definition, we can see that it is not about some objectively defined phenomenon, like rain or a car crash—it is simply something that was not expected by a *particular* observer.

So I was wondering why so many otherwise intelligent people have casually questioned whether certain events, say the Great War, or the September 11, 2001, attack on the World Trade Center, were Black Swans, on the grounds that *some* predicted them. Of course the September 11 attack was a Black Swan to those victims who died in it; otherwise, they would not have exposed themselves to the risk. But it was certainly not a Black Swan to the terrorists who planned and carried out the attack. I have spent considerable time away from the weight-lifting room repeating that *a Black Swan for the turkey is not a Black Swan for the butcher*.

The same applies to the crisis of 2008, certainly a Black Swan to almost all economists, journalists, and financiers on this planet (including, predictably, Robert Merton and Myron Scholes, the turkeys of [Chapter 17](#)), but certainly not to this author. (Incidentally, as an illustration of another common mistake, almost none of those—very few—who seemed to have “predicted” the event predicted its depth. We will see that, because of the atypicality of events in Extremistan, the Black Swan is not just about the occurrence of some event but also about its depth and consequences.)

ASPERGER PROBABILITY

This consideration of an *objective* Black Swan, one that would be the same to all observers, aside from missing the point completely, seems dangerously related to the problem of underdevelopment of a human faculty called “theory of mind” or “folk psychology.” Some people, otherwise intelligent, have a deficiency of that human ability to impute to others knowledge that is different from their own. These, according to researchers, are the people you commonly see involved in engineering or populating physics departments. We saw one of them, Dr. John, in [Chapter 9](#).

You can test a child for underdevelopment of the theory of mind using a variant of the “false-belief task.” Two children are introduced. One child puts a toy under the bed and leaves the room. During his absence, the second child—the subject—removes it and hides it in a box. You ask the subject: Where, upon returning to the room, will the other child look for the toy? Those under, say, the age of four (when the theory of mind starts developing), choose the box, while older children correctly say that the other child will look under the bed. At around that age, children start realizing that another person can be deprived of some of the information they have, and can hold beliefs that are different from their own. Now, this test helps detect mild forms of autism: as high as one’s intelligence may be, it can be difficult for many to put themselves in other people’s shoes and imagine the world on the basis of other people’s information. There is actually a name for the condition of a person who can be functional but suffers from a mild form of autism: Asperger syndrome.

The psychologist Simon Baron-Cohen has produced much research distinguishing between polar extremes in people’s temperament with respect to two faculties: ability to systematize, and ability to empathize and understand others. According to his research, purely systematizing persons suffer from a lack of theory of mind; they are drawn to engineering and similar occupations (and when they fail, to, say, mathematical economics); empathizing minds are drawn to more social (or literary) professions. Fat Tony, of course, would fall in the more social category. Males are overrepresented in the systematizing category; females dominate the other extreme.

Note the unsurprising, but very consequential fact that people with

Asperger syndrome are highly averse to ambiguity.

Research shows that academics are overrepresented in the systematizing, Black-Swan-blind category; these are the people I called “Locke’s madmen” in [Chapter 17](#). I haven’t seen any formal direct test of Black Swan foolishness and the systematizing mind, except for a calculation George Martin and I made in 1998, in which we found evidence that all the finance and quantitative economics professors from major universities whom we tracked and who got involved in hedge fund trading ended up making bets *against* Black Swans, exposing themselves to blowups. This preference was nonrandom, since between one third and one half of the nonprofessors had that investment style at the time. The best known such academics were, once again, the “Nobel”-crowned Myron Scholes and Robert C. Merton, whom God created so that I could illustrate my point about Black Swan blindness.* They all experienced problems during the crisis, discussed in that chapter, that brought down their firm Long Term Capital Management. Note that the very same people who make a fuss about discussions of Asperger as a condition not compatible with risk-bearing and the analysis of nonexplicit off-model risks, with its corresponding dangers to society, would be opposed to using a person with highly impaired eyesight as the driver of a school bus. All I am saying is that just as I read Milton, Homer, Taha Husain, and Borges (who were blind) but would prefer not to have them drive me on the Nice–Marseilles motorway, I elect to use tools made by engineers but prefer to have society’s risky decisions managed by someone who is not affected with risk-blindness.

FUTURE BLINDNESS REDUX

Now recall the condition, described in [Chapter 12](#), of not properly transferring between past and future, an autism-like condition in which people do not see second-order relations—the subject does not use the relation between the past’s past and the past’s future to project the connection between today’s past and today’s future. Well, a gentleman called Alan Greenspan, the former chairman of the U.S. Federal Reserve Bank, went to Congress to explain that the banking crisis, which he and his successor Bernanke helped cause, could not have been foreseen because it “had never happened before.” Not a single member of congress was intelligent enough to shout, “Alan Greenspan, you have never died before, not in eighty years, not even once; does that make you immortal?” The abject Robert Rubin, the bankster I was chasing in Section II, a former secretary of the Treasury, used the same argument—but the fellow had written a long book on uncertainty (with, ironically, my publisher and the same staff used for *The Black Swan*).*

I discovered (but by then I was not even surprised) that no researcher has tested whether large deviations in economics can be predicted from past large deviations—whether large deviations have predecessors, that is. This is one of the elementary tests missing in the field, as elementary as checking whether a patient is breathing or whether a lightbulb is screwed in, but characteristically nobody seems to have tried to do it. It does not take a lot of introspection to figure out that big events don’t have big parents: the Great War did not have a predecessor; the crash of 1987, in which the market went down close to 23 percent in a single day, could not have been guessed from its worst predecessor, a one-day loss of around 10 percent—and this applies to almost all such events, of course. My results were that regular events can predict regular events, but that extreme events, perhaps because they are more acute when people are unprepared, are almost never predicted from narrow reliance on the past.

The fact that this notion is not obvious to people is shocking to me. It is particularly shocking that people do what are called “stress tests” by taking the worst possible *past* deviation as an anchor event to project the worst possible future deviation, not thinking that they would have failed to account for that past deviation had they used the same method on the day before the occurrence of that past anchor event.*

These people have PhDs in economics; some are professors—one of them is the chairman of the Federal Reserve (at the time of writing). Do advanced degrees make people blind to these elementary notions?

Indeed, the Latin poet Lucretius, who did not attend business school, wrote that we consider the biggest object of any kind that we have seen in our lives as the largest possible item: *et omnia de genere omni / Maxima quae vivit quisque, haec ingentia fingit.*

PROBABILITY HAS TO BE SUBJECTIVE[†]

This raises a problem that is worth probing in some depth. The fact that many researchers do not realize immediately that the Black Swan corresponds mainly to an incomplete map of the world, or that some researchers have to stress this subjective quality (Jochen Runde, for instance, wrote an insightful essay on the Black Swan idea, but one in which he felt he needed to go out of his way to stress its subjective aspect), takes us to the historical problem in the very definition of probability. Historically, there have been many approaches to the philosophy of probability. The notion that two people can have two different views of the world, then express them as different probabilities remained foreign to the research. So it took a while for scientific researchers to accept the non-Asperger notion that different people can, while being rational, assign different probabilities to different future states of the world. This is called “subjective probability.”

Subjective probability was formulated by Frank Plumpton Ramsey in 1925 and Bruno de Finetti in 1937. The take on probability by these two intellectual giants is that it can be represented as a quantification of the degree of belief (you set a number between 0 and 1 that corresponds to the strength of your belief in the occurrence of a given event), subjective to the observer, who expresses it as rationally as he wishes under some constraints. These constraints of consistency in decision making are obvious: you cannot bet there is a 60 percent chance of snow tomorrow *and* a 50 percent chance that there will be no snow. The agent needs to avoid violating something called the Dutch book constraint: that is, you cannot express your probabilities inconsistently by engaging in a series of bets that lock in a certain loss, for example, by acting as if the probabilities of separable contingencies can add up to more than 100 percent.

There is another difference here, between “true” randomness (say the equivalent of God throwing a die) and randomness that results from what I call epistemic limitations, that is, lack of knowledge. What is called ontological (or ontic) uncertainty, as opposed to epistemic, is the type of randomness where the future is not implied by the past (or not even implied by anything). It is created every minute by the complexity of our actions, which makes the uncertainty much more fundamental than the epistemic one coming from imperfections in knowledge.

It means that there is no such thing as a long run for such systems, called “nonergodic” systems—as opposed to the “ergodic” ones. In an ergodic system, the probabilities of what may happen in the long run are not impacted by events that may take place, say, next year. Someone playing roulette in the casino can become very rich, but, if he keeps playing, given that the house has an advantage, he will eventually go bust. Someone rather unskilled will eventually fail. So ergodic systems are invariant, on average, to paths, taken in the intermediate term—what researchers call absence of path dependency. A nonergodic system has no real long-term properties—it is prone to path dependency.

I believe that the distinction between epistemic and ontic uncertainty is important philosophically, but entirely irrelevant in the real world. Epistemic uncertainty is so hard to disentangle from the more fundamental one. This is the case of a “distinction without a difference” that (unlike the ones mentioned earlier) can mislead because it distracts from the real problems: practitioners make a big deal out of it instead of focusing on epistemic constraints. Recall that skepticism is costly, and should be available when needed.

There is no such thing as a “long run” in practice; what matters is what happens before the long run. The problem of using the notion of “long run,” or what mathematicians call the asymptotic property (what happens when you extend something to infinity), is that it usually makes us blind to what happens before the long run, which I will discuss later as *preasymptotics*. Different functions have different preasymptotics, according to speed of convergence to that asymptote. But, unfortunately, as I keep repeating to students, *life takes place in the preasymptote*, not in some Platonic long run, and some properties that hold in the preasymptote (or the short run) can be markedly divergent from those that take place in the long run. So theory, even if it works, meets a short-term reality that has more texture. Few understand that there is generally no such thing as a reachable *long run* except as a mathematical construct to solve equations; to assume a long run in a complex system, you need to also assume that nothing new will emerge. In addition, you may have a perfect model of the world, stripped of any uncertainty concerning the analytics of the representation, but have a small imprecision in one of the parameters to input in it. Recall Lorenz’s butterfly effect of [Chapter 11](#). Such minutely small uncertainty, at the level of the slightest parameter, might, because of nonlinearities, percolate to a huge uncertainty at the level of the output of the model. Climate models, for instance, suffer from such nonlinearities, and even

if we had the right model (which we, of course, don't), a small change in one of the parameters, called calibration, can entirely reverse the conclusions.

We will discuss preasymptotics further when we look at the distinctions between different classes of probability distributions. I will say for now that many of these mathematical and philosophical distinctions are entirely overblown, Soviet-Harvard-style, top-down, as people start with a model and then impose it on reality and start categorizing, rather than start with reality and look at what fits it, in a bottom-up way.

Probability on a Thermometer

This distinction, misused in practice, resembles another deficient separation discussed earlier, between what economists call Knightian risk (computable) and Knightian uncertainty (uncomputable). This assumes that something is computable, when really everything is more or less incomputable (and rare events more so). One has to have a mental problem to think that probabilities of future events are “measurable” in the same sense that the temperature is measurable by a thermometer. We will see in the following section that small probabilities are less computable, and that this matters when the associated payoffs are consequential.

Another deficiency I need to point out concerns a strangely unrealistic and unrigorous research tradition in social science, “rational expectations,” in which observers are shown to rationally converge on the same inference when supplied with the same data, even if their initial hypotheses were markedly different (by a mechanism of updating called Bayesian inference). Why unrigorous? Because one needs a very quick check to see that people do not converge to the same opinions in reality. This is partly, as we saw in [Chapter 6](#), because of psychological distortions such as the confirmation bias, which cause divergent interpretation of the data. But there is a mathematical reason why people do not converge to the same opinion: if you are using a probability distribution from Extremistan, and I am using a distribution from Mediocristan (or a different one from Extremistan), then we will never converge, simply because if you suppose Extremistan you do not update (or change your mind) that quickly. For instance, if you assume Mediocristan and do not witness Black Swans, you may eventually rule them out. Not if you assume we are in Extremistan.

To conclude, assuming that “randomness” is not epistemic and subjective, or making a big deal about the distinction between “ontological randomness” and “epistemic randomness,” implies some scientific autism, that desire to systematize, and a fundamental lack of understanding of randomness itself. It assumes that an observer can reach omniscience and can compute odds with perfect realism and without violating consistency rules. What is left becomes “randomness,” or something by another name that arises from aleatory forces that cannot be reduced by knowledge and analysis.

There is an angle worth exploring: why on earth do adults accept these Soviet-Harvard-style top-down methods without laughing, and actually go to build policies in Washington based on them, against the record, except perhaps to make readers of history laugh at them and diagnose new psychiatric conditions? And, likewise, why do we default to the assumption that events are experienced by people in the same manner? Why did we ever take notions of “objective” probability seriously?

After this foray into the psychology of the perception of the dynamics of time and events, let us move to our central point, the very core of our program, into what I have aggressively called the most useful problem in philosophy. The most useful, sadly.

-
- * Robert Merton, the villain of [Chapter 17](#), a person said to be of a highly mechanistic mind (down to his interest in machinery and his use of mechanical metaphors to represent uncertainty), seems to have been created for the sole purpose of providing an illustration of dangerous Black Swan foolishness. After the crisis of 2008, he defended the risk taking caused by economists, giving the argument that “it was a Black Swan” simply because he did not see it coming, therefore, he said, the theories were fine. He did not make the leap that, since we do not see these events coming, we need to be robust to them. Normally, such people exit the gene pool; academic tenure holds them a bit longer.
 - * The argument can actually be used to satisfy moral hazard and dishonest (probabilistically disguised) profiteering. Rubin had pocketed more than \$100 million from Citigroup’s earning of profits from hidden risks that blow up only occasionally. After he blew up, he had an excuse—“It never happened before.” He kept his money; we, the taxpayers, who include schoolteachers and hairdressers, had to bail the company out and pay for the losses. This I call the moral hazard element in paying bonuses to people who are not robust to Black Swans, and who we knew *beforehand* were not robust to the Black Swan. This *beforehand* is what makes me angry.
 - * It is indeed the absence of higher order representation—the inability to accept statements like “Is my method for assessing what is right or wrong right or wrong?”—that, we will see in the next section, is central when we deal with probability, that causes Dr. Johns to be suckers for measures and believe in them without doubting their beliefs. They fail to understand the metaprobability, the higher order probability—that is, the probability that the probability they are using may not be True.

† The nontechnical reader should skip the rest of this section.

(PERHAPS) THE MOST USEFUL PROBLEM IN THE HISTORY OF MODERN PHILOSOPHY

Small may not be the idea, after all—Where to find the powder room—Predict and perish—On school buses and intelligent textbooks

I am going to be blunt. Before *The Black Swan* (and associated papers) most of epistemology and decision theory was, to an actor in the real world, just sterile mind games and foreplay. Almost all the history of thought is about what we know, or think we know. *The Black Swan* is the *very first attempt* (that I know of) in the history of thought to provide a map of where we get hurt by what we don't know, to set systematic limits to the fragility of knowledge—and to provide exact locations where these maps no longer work.

To answer the most common “criticism” by economists and (now bankrupt) bankers I mentioned in Section III, I am not saying “S**t happens,” I am saying “S**t happens in the Fourth Quadrant,” which is as different as mistaking prudence and caution for paranoia.

Furthermore, to be more aggressive, while limits like those attributed to Gödel bear massive philosophical consequences, but we can't do much about them, I believe that the limits to empirical and statistical knowledge I have shown have sensible (if not vital) importance *and* we can do a lot with them in terms of solutions, by categorizing decisions based on the severity of the potential estimation error of the pair probability times consequence. For instance, we can use it to build a safer society—to robustify what lies in the Fourth Quadrant.

LIVING IN TWO DIMENSIONS

A vexing problem in the history of human thought is finding one's position on the boundary between skepticism and gullibility, or how to believe and how to *not* believe. And how to make decisions based on these beliefs, since beliefs without decisions are just sterile. So this is not an epistemological problem (i.e., focusing on what is true or false); it is one of decision, action, and commitment.

Clearly, you cannot doubt everything and function; you cannot believe everything and survive. Yet the philosophical treatment of the problem has been highly incomplete, and, worse, has not improved much over the centuries, if it has improved at all. One class of thinkers, say the Cartesians, or the academic skeptics some eighteen centuries before them, in their own way, started with the rejection of everything upfront, with some even more radical, such as the Pyrrhonians, rejecting so much that they even reject skepticism as too dogmatic. The other class, say the medieval Scholastics or the modern-day pragmatists, starts with the fixation of beliefs, or some beliefs. While the medieval thinkers stop there, in an Aristotelian way, the early pragmatists, with the great thinker Charles Sanders Peirce, provided a ray of hope. They proposed to update and correct beliefs as a continuous work in progress (albeit under a known structure of probability, as Peirce believed in the existence and attainability of an ergodic, long-run, reachable state of convergence to truth). That brand of pragmatism (initially called *pragmaticism*) viewed knowledge as a rigorous interplay between anti-skepticism and fallibilism, i.e., between the two categories of what to doubt and what to accept. The application to my field, probability, and perhaps the most sophisticated version of the program, lies in the dense, difficult, deep, and brilliant forays of Isaac Levi into decision theory with the notion of corpus of belief, doxastic commitment, distance from expectation, and credal probabilities.

A ray of hope, perhaps, but still not even close. Not even remotely close to anything useful.

Think of living in a three-dimensional space while under the illusion of being in two dimensions. It may work well if you are a worm, certainly not if you happen to be a bird. Of course, you will not be aware of the truncation—and will be confronted with many mysteries, mysteries that you cannot possibly clear up without adding a dimension, no matter how sophisticated

you may get. And, of course, you will feel helpless at times. Such was the fate of knowledge all these centuries, when it was locked in two dimensions too simplistic to be of any use outside of classrooms. Since Plato only philosophers have spent time discussing what Truth was, and for a reason: it is unusable in practice. By focusing on the True/False distinction, epistemology remained, with very few exceptions, prisoner of an inconsequential, and highly incomplete, 2-D framework. The third missing dimension is, of course, the consequence of the True, and the severity of the False, the expectation. In other words, *the payoff from decisions*, the impact and magnitude of the result of such a decision. Sometimes one may be wrong and the mistake may turn out to be inconsequential. Or one may be right, say, on such a subject as the sex of angels, and it may turn out to be of no use beyond intellectual stamp collecting.

The simplified, philistinified, academified, and glorified notion of “evidence” becomes useless. With respect to Black Swans, you act to protect yourself from negative ones (or expose yourself to positive ones) even though you may have *no evidence* that they can take place, just as we check people for weapons before they board a plane even though we have *no evidence* that they are terrorists. This focus on off-the-shelf commoditized notions such as “evidence,” is a problem with people who claim to use “rigor” yet go bust on occasion.

A probabilistic world has trouble with “proof” as it is, but in a Black Swan world things are a lot worse.

Indeed, I know of almost no decision that is based on notions of True/False.

Once you start examining the payoff, the result of decisions, you will see clearly that the consequences of some errors may be benign, those of others may be severe. And you pretty much know which is which beforehand. You know which errors are consequential and which ones are not so much.

But first let us look at a severe problem in the derivation of knowledge about probabilities.

THE DEPENDENCE ON THEORY FOR RARE EVENTS

During my *deserto* period, when I was getting severe but entertaining insults, I found myself debating a gentleman then employed by a firm called Lehman Brothers. That gentleman had made a statement in *The Wall Street Journal* saying that events we saw in August 2007 should have happened once every ten thousand years. Sure enough, we had three such events three days in a row. *The Wall Street Journal* ran his picture and if you look at it, you can safely say, “He does not look ten thousand years old.” So where is he getting his “once in ten thousand years” probability? Certainly not from personal experience; certainly not from the records of Lehman Brothers—his firm had not been around for ten thousand years, and of course it didn’t stay around for another ten thousand years, as it went under right after our debate. So, you know that he’s getting his small probabilities from a theory. *The more remote the event, the less we can get empirical data (assuming generously that the future will resemble the past) and the more we need to rely on theory.*

Consider that the frequency of rare events cannot be estimated from empirical observation for the very reason that *they are rare*. We thus need a prior model representation for that; the rarer the event, the higher the error in estimation from standard inductive methods (say, frequency sampling from counting past occurrences), hence the higher the dependence on an a priori representation that extrapolates into the space of low-probability events (which necessarily are not seen often).*

But even outside of small probabilities, the a priori problem is always present. It seems salient with respect to rare events, but it pervades probabilistic knowledge. I will present two versions I have been working on with two collaborators, Avital Pilpel, a philosopher of science (he walks fast), and Raphael Douady, a mathematician (he is sometimes a good walker, when he is not busy).

Epimenides the Cretan

Avital Pilpel and I expressed the regress argument as follows, as the epistemic problem of risk management, but the argument can be generalized to any form of probabilistic knowledge. It is a problem of *self-reference* by probability measures.