THE BLACK SWAN

The Impact of the Highly Improbable

NASSIM NICHOLAS TALEB

CONTENTS

Tables and Figures 1

Glossary 14

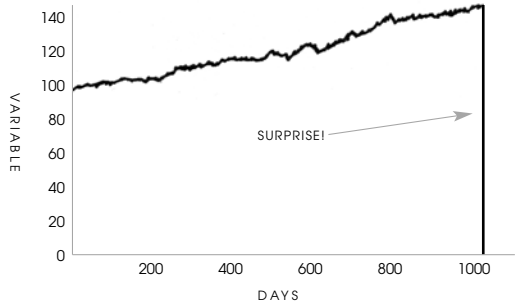
Notes. 18

TABLE 1

Mediocristan	*Extremistan*
Nonscalable	Scalable
Mild or type 1 randomness	Wild (even superwild) or type 2 randomness
The most typical member is mediocre	The most "typical" is either giant or dwarf, i.e., there is no typical member
Winners get a small segment of the total pie	Winner-take-almost-all effects
Example: audience of an opera singer before the gramophone	Today's audience for an artist
More likely to be found in our ancestral environment	More likely to be found in our modern environment
Impervious to the Black Swan	Vulnerable to the Black Swan
Subject to gravity	There are no physical constraints on what a number can be
Corresponds (generally) to physical quantities, i.e., height	Corresponds to numbers, say, wealth
As close to utopian equality as reality can spontaneously deliver	Dominated by extreme winner-take-all inequality
Total is not determined by a single instance or observation	Total will be determined by a small number of extreme events
When you observe for a while you can get to know what's going on	It takes a long time to know what's going on
Tyranny of the collective	Tyranny of the accidental
Easy to predict from what you see and extend to what you do not see	Hard to predict from past information
History crawls	History makes jumps
Events are distributed* according to the "bell curve" (the GIF) or its variations	The distribution is either Mandelbrotian "gray" Swans (tractable scientifically) or totally intractable Black Swans

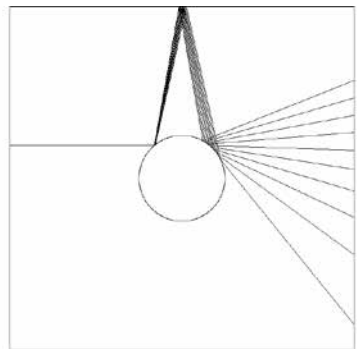
* What I call "probability distribution" here is the model used to calculate the odds of different events, how they are distributed. When I say that an event is distributed according to the "bell curve," I mean that the Gaussian bell curve (after C. F. Gauss; more on him later) can help provide probabilities of various occurrences.

FIGURE 1: ONE THOUSAND AND ONE DAYS OF HISTORY



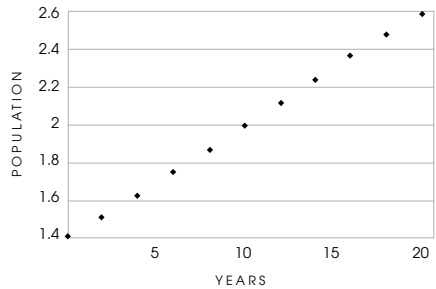
A turkey before and after Thanksgiving. The history of a process over a thousand days tells you nothing about what is to happen next. This naïve projection of the future from the past can be applied to anything.

FIGURE 2: PRECISION AND FORECASTING



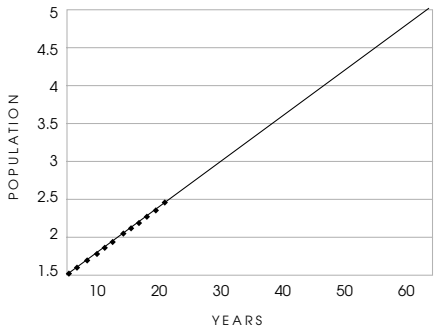
One of the readers of a draft of this book, David Cowan, gracefully drew this picture of scattering, which shows how, at the second bounce, variations in the initial conditions can lead to extremely divergent results. As the initial imprecision in the angle is multiplied, every additional bounce will be further magnified. This causes a severe multiplicative effect where the error grows out disproportionately.

FIGURE 3



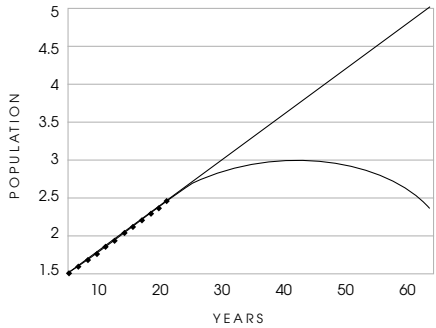
A series of a seemingly growing bacterial population (or of sales records, or of any variable observed through time—such as the total feeding of the turkey in Chapter 4).

FIGURE 4



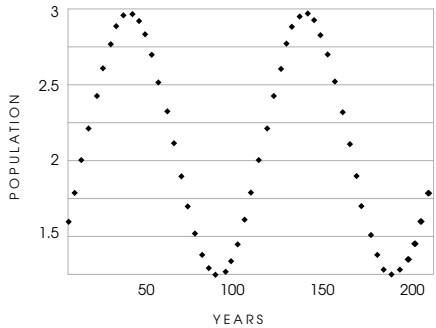
Easy to fit the trend—there is one and only one linear model that fits the data. You can project a continuation into the future.

FIGURE 5



We look at a broader scale. Hey, other models also fit it rather well.

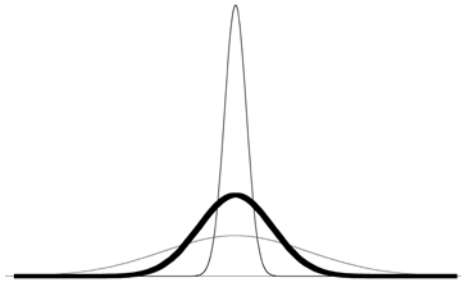
FIGURE 6



And the real "generating process" is extremely simple but it had nothing to do with a linear model! Some parts of it appear to be linear and we are fooled by extrapolating in a direct line.*

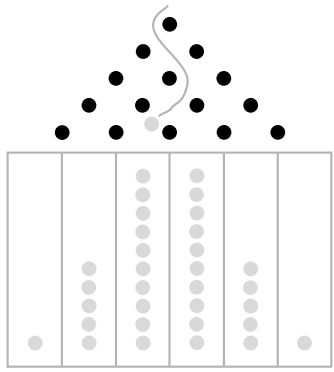
* These graphs also illustrate a statistical version of the narrative fallacy—you find a model that fits the past. "Linear regression" or "R-square" can ultimately fool you beyond measure, to the point where it is no longer funny. You can fit the linear part of the curve and claim a high R-square, meaning that your model fits the data very well and has high predictive powers. All that off hot air: you only fit the linear segment of the series. Always remember that "R-square" is unfit for Extremistan; it is only good for academic promotion.

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."

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

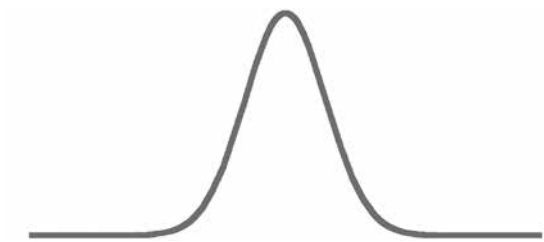


Drop balls that, at every pin, randomly fall right or left. Above is the most probable scenario, which greatly resembles the bell curve (a.k.a. Gaussian disribution). *Courtesy of Alexander Taleb.*

FIGURE 9: NUMBERS OF WINS TOSSED

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

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

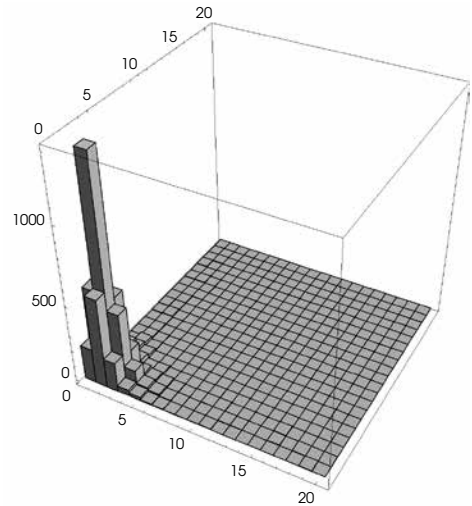


An infinite number of tosses.

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

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.*

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*

Phenomenon	*Assumed Exponent (vague approximation)*
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)

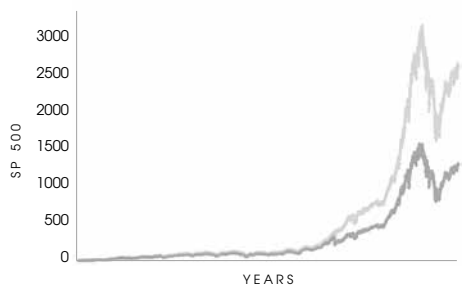
** Source: M.E.J. Newman (2005) and the author's own calculations.*

TABLE 3: THE MEANING OF THE EXPONENT

Exponent	*Share of the top 1%*	*Share of the top 20%*
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%

** Clearly, you do not observe 100 percent in a finite sample.*

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.)

TABLE 4: TWO WAYS TO APPROACH RANDOMNESS

Skeptical Empiricism and the a-Platonic School	***The Platonic Approach***
Interested in what lies outside the Platonic fold	Focuses on the inside of the Platonic fold
Respect for those who have the guts to say "I don't know"	"You keep criticizing these models. These models are *all* we have."
Fat Tony	Dr. John
Thinks of Black Swans as a dominant source of randomness	Thinks of ordinary fluctuations as a dominant source of randomness, with jumps as an afterthought
Bottom-up	Top-down
Would ordinarily not wear suits (except to funerals)	Wears dark suits, white shirts; speaks in a boring tone
Prefers to be broadly right	Precisely wrong
Minimal theory, considers theorizing as a disease to resist	Everything needs to fit some grand, general socioeconomic model and "the rigor of economic theory"; frowns on the "descriptive"
Does not believe that we can easily compute probabilities	Built their entire apparatus on the assumptions that we can compute probabilities
Model: Sextus Empiricus and the school of evidence-based, minimum-theory empirical medicine	Model: Laplacian mechanics, the world and the economy like a clock
Develops intuitions from practice, goes from observations to books	Relies on scientific papers, goes from books to practice
Not inspired by any science, uses messy mathematics and computational methods	Inspired by physics, relies on abstract mathematics
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

TABLE 1: TABLEAU OF DECISIONS BY PAYOFF

M0 **"True/False"**	***M1*** **Expectations**
Medical results for one person (health, not epidemics)	Epidemics (number of persons infected)
Psychology experiments (yes/no answers)	Intellectual and artistic success (defined as book sales, citations, etc.)
Life/Death (for a single person, not for *n* persons)	Climate effects (any quantitative metric)
Symmetric bets in roulette	War damage (number of casualties)
Prediction markets	Security, terrorism, natural catastrophes (number of victims)
	General risk management
	Finance: performance of a nonleveraged investment (say, a retirement account)
	Insurance (measures of expected losses)
	Economics (policy)
	Casinos

TABLE 2: THE FOUR QUADRANTS

	I Simple Payoffs	**II** Complex Payoffs
A Mediocristan	First Quadrant ***Extremely Safe***	Second Quadrant ***(Sort of) Safe***
B Extremistan	Third Quadrant ***Safe***	Fourth Quadrant ***Black Swan Domain***

* This is a true philosophical a priori since when you assume events belong to Extremistan (because of the lack of structure to the randomness), no additional empirical observations can possibly change your mind, since the property of Extremistan is to hide the possibility of Black Swan events—what I called earlier the masquerade problem.

GLOSSARY

Academic libertarian: someone (like myself) who considers that knowledge is subjected to strict rules but not institutional authority, as the interest of organized knowledge is self-perpetuation, not necessarily truth (as with governments). Academia can suffer from an acute **expert problem** (q.v.), producing cosmetic but fake knowledge, particularly in **narrative disciplines** (q.v.), and can be a main source of Black Swans.

Apelles-style strategy: A strategy of seeking gains by collecting positive accidents from maximizing exposure to "good Black Swans."

Barbell strategy: a method that consists of taking both a defensive attitude and an excessively aggressive one at the same time, by protecting assets from all sources of uncertainty while allocating a small portion for high-risk strategies.

Bildungsphilister: a philistine with cosmetic, nongenuine culture. Nietzsche used this term to refer to the dogma-prone newspaper reader and opera lover with cosmetic exposure to culture and shallow depth. I extend it to the buzzword-using researcher in nonexperimental fields who lacks in imagination, curiosity, erudition, and culture and is closely centered on his ideas, on his "discipline." This prevents him from seeing the conflicts between his ideas and the texture of the world.

Black Swan blindness: the underestimation of the role of the Black Swan, and occasional overestimation of a specific one.

Black Swan ethical problem: Owing to the nonrepeatable aspect of the Black Swan, there is an asymmetry between the rewards of those who prevent and those who cure.

Confirmation error (or Platonic confirmation): You look for instances that confirm your beliefs, your construction (or model)—and find them.

Empty-suit problem (or "expert problem"): Some professionals have no differential abilities from the rest of the population, but for some reason, and against their empirical records, are believed to be experts: clinical psychologists, academic economists, risk "experts," statisticians, political analysts, financial "experts," military analysts, CEOs, et cetera. They dress up their expertise in beautiful language, jargon, mathematics, and often wear expensive suits.

Epilogism: A theory-free method of looking at history by accumulating facts with minimal generalization and being conscious of the side effects of making causal claims.

Epistemic arrogance: Measure the difference between what someone actually knows and how much he thinks he knows. An excess will imply arrogance, a deficit humility. An epistemocrat is someone of epistemic humility, who holds his own knowledge in greatest suspicion.

Epistemic opacity: Randomness is the result of incomplete information at some layer. It is functionally indistinguishable from "true" or "physical" randomness.

Extremistan: the province where the total can be conceivably impacted by a single observation.

Fallacy of silent evidence: Looking at history, we do not see the full story, only the rosier parts of the process.

Fooled by randomness: the general confusion between luck and determinism, which leads to a variety of superstitions with practical consequences, such as the belief that higher earnings in some professions are generated by skills when there is a significant component of luck in them.

Future blindness: our natural inability to take into account the properties of the future—like autism, which prevents one from taking into account the existence of the minds of others.

Locke's madman: someone who makes impeccable and rigorous reasoning from faulty premises—such as Paul Samuelson, Robert Merton the minor, and Gerard Debreu—thus producing phony models of uncertainty that make us vulnerable to Black Swans.

Lottery-ticket fallacy: the naïve analogy equating an investment in collect-

ing positive Black Swans to the accumulation of lottery tickets. Lottery tickets are not scalable.

Ludic fallacy (or uncertainty of the nerd): the manifestation of the Platonic fallacy in the study of uncertainty; basing studies of chance on the narrow world of games and dice. A-Platonic randomness has an additional layer of uncertainty concerning the rules of the game in real life. The bell curve (Gaussian), or GIF (Great Intellectual Fraud), is the application of the ludic fallacy to randomness.

Mandelbrotian Gray Swan: Black Swans that we can somewhat take into account—earthquakes, blockbuster books, stock market crashes—but for which it is not possible to completely figure out the properties and produce precise calculations.

Mediocristan: the province dominated by the mediocre, with few extreme successes or failures. No single observation can meaningfully affect the aggregate. The bell curve is grounded in Mediocristan. There is a qualitative difference between Gaussians and scalable laws, much like gas and water.

Narrative discipline: the discipline that consists in fitting a convincing and well-sounding story to the past. Opposed to experimental discipline.

Narrative fallacy: our need to fit a story or pattern to a series of connected or disconnected facts. The statistical application is data mining.

Nerd knowledge: the belief that what cannot be Platonized and studied does not exist at all, or is not worth considering. There even exists a form of skepticism practiced by the nerd.

Platonic fold: the place where our Platonic representation enters into contact with reality and you can see the side effects of models.

Platonicity: the focus on those pure, well-defined, and easily discernible objects like triangles, or more social notions like friendship or love, at the cost of ignoring those objects of seemingly messier and less tractable structures.

Probability distribution: the model used to calculate the odds of different events, how they are "distributed." When we say that an event is distributed according to the bell curve, we mean that the Gaussian bell curve can help provide probabilities of various occurrences.

Problem of induction: the logical-philosophical extension of the Black Swan problem.

Randomness as incomplete information: simply, what I cannot guess is random because my knowledge about the causes is incomplete, not necessarily because the process has truly unpredictable properties.

Retrospective distortion: examining past events without adjusting for the forward passage of time. It leads to the illusion of posterior predictability.

Reverse-engineering problem: It is easier to predict how an ice cube would melt into a puddle than, looking at a puddle, to guess the shape of the ice cube that may have caused it. This "inverse problem" makes narrative disciplines and accounts (such as histories) suspicious.

Round-trip fallacy: the confusion of absence of evidence of Black Swans (or something else) for evidence of absence of Black Swans (or something else). It affects statisticians and other people who have lost part of their reasoning by solving too many equations.

Scandal of prediction: the poor prediction record in some forecasting entities (particularly narrative disciplines) mixed with verbose commentary and a lack of awareness of their own dire past record.

Scorn of the abstract: favoring contextualized thinking over more abstract, though more relevant, matters. "The death of one child is a tragedy; the death of a million is a statistic."

Statistical regress argument (or the problem of the circularity of statistics): We need data to discover a probability distribution. How do we know if we have enough? From the probability distribution. If it is a Gaussian, then a few points of data will suffice. How do we know it is a Gaussian? From the data. So we need the data to tell us what probability distribution to assume, and we need a probability distribution to tell us how much data we need. This causes a severe regress argument, which is somewhat shamelessly circumvented by resorting to the Gaussian and its kin.

Uncertainty of the deluded: people who tunnel on sources of uncertainty by producing precise sources like the great uncertainty principle, or similar, less consequential matters, to real life; worrying about subatomic particles while forgetting that we can't predict tomorrow's crises.

NOTES

BEHIND THE CURTAIN: ADDITIONAL NOTES, TECHNICAL COMMENTS, REFERENCES, AND READING RECOMMENDATIONS

I separate topics thematically; so general references will mostly be found in the chapter in which they first occur. I prefer to use a logical sequence here rather than stick to chapter division.

PROLOGUE and CHAPTER 1

Bell curve: When I write *bell curve* I mean the Gaussian bell curve, a.k.a. normal distribution. All curves look like bells, so this is a nickname. Also, when I write *the Gaussian basin* I mean all distributions that are similar and for which the improbable is inconsequential and of low impact (more technically, nonscalable—all moments are finite). Note that the visual presentation of the bell curve in histogram form masks the contribution of the remote event, as such an event will be a point to the far right or far left of the center.

Diamonds: See Eco (2002).

Platonicity: I'm simply referring to incurring the risk of using a wrong form—not that forms don't exist. I am not against essentialisms; I am often skeptical of our reverse engineering and identification of the right form. It is an inverse problem!

Empiricist: If I call myself an empiricist, or an empirical philosopher, it is because I am just suspicious of confirmatory generalizations and hasty theorizing. Do not confuse this with the British empiricist tradition. Also, many statisticians, as we will see with the Makridakis competition, call themselves "empirical" researchers, but are in fact just the opposite—they fit theories to the past.

Mention of Christ: See Flavius Josephus's *The Jewish War.*

Great War and prediction: Ferguson (2006b).

Hindsight bias (retrospective distortion): See Fischhoff (1982b).

Historical fractures: Braudel (1985), p. 169, quotes a little known passage from Gautier. He writes, " 'This long history,' wrote Emile-Félix Gautier, 'lasted a dozen centuries, longer than the entire history of France. Encountering the first Arab sword, the Greek language and thought, all that heritage went up in smoke, as if it never happened.' " For discussions of discontinuity, see also Gurvitch (1957), Braudel (1953), Harris (2004).

Religions spread as bestsellers: Veyne (1971). See also Veyne (2005).
Clustering in political opinions: Pinker (2002).
Categories: Rosch (1973, 1978). See also Umberto Eco's *Kant and the Platypus*.
Historiography and philosophy of history: Bloch (1953), Carr (1961), Gaddis (2002), Braudel (1969, 1990), Bourdé and Martin (1989), Certeau (1975), *Muqaddamat* Ibn Khaldoun illustrate the search for causation, which we see already present in Herodotus. For philosophy of history, Aron (1961), Fukuyama (1992). For postmodern views, see Jenkins (1991). I show in Part Two how historiographers are unaware of the epistemological difference between forward and backward processes (i.e., between projection and reverse engineering).
Information and markets: See Shiller (1981, 1989), DeLong et al. (1991), and Cutler et al. (1989). The bulk of market moves does not have a "reason," just a contrived explanation.
Of descriptive value for crashes: See Galbraith (1997), Shiller (2000), and Kindleberger (2001).

CHAPTER 3

Movies: See De Vany (2002). See also Salganik et al. (2006) for the contagion in music buying.
Religion and domains of contagion: See Boyer (2001).
Wisdom (madness) of crowds: Collectively, we can both get wiser or far more foolish. We may collectively have intuitions for Mediocristan-related matters, such as the weight of an ox (see Surowiecki, 2004), but my conjecture is that we fail in more complicated predictions (economic variables for which crowds incur pathologies—two heads are worse than one). For decision errors and groups, see Sniezek and Buckley (1993). Classic: Charles Mackay's *Extraordinary Popular Delusions and the Madness of Crowds*.
Increase in the severity of events: Zajdenweber (2000).
Modern life: The nineteenth-century novelist Émile Zola welcomed the arrival of the market for culture in the late 1800s, of which he seemed to be one of the first beneficiaries. He predicted that the writers' and artists' ability to exploit the commercial system freed them from a dependence on patrons' whims. Alas, this was accompanied with more severe concentration—very few people benefited from the system. Lahire (2006) shows how most writers, throughout history, have starved. Remarkably, we have ample data from France about the literary tradition.

CHAPTER 4

Titanic: The quote is from Dave Ingram's presentation at the Enterprise Risk Management Symposium in Chicago on May 2, 2005. For more on LTCM, see Lowenstein (2000), Dunbar (1999).
Hume's exposition: Hume (1748, 2000).
Sextus Empiricus: "It is easy, I think, to reject the method of induction (επαγωγη). For since by way of it they want to make universals convincing on the basis of particulars, they will do this surveying all the particulars or some of them. But if some, the induction will be infirm, it being that some of the particulars omitted in the induction should be contrary to the universal; and if all, they will labor at an impossible task, since the particulars and infinite are indeterminate. Thus in either case it results, I think, that induction totters." *Outline of Pyrrhonism,* Book II, p. 204.
Bayle: The *Dictionnaire historique et critique* is long (twelve volumes, close to 6,000 pages) and heavy (40 pounds), yet it was an intellectual bestseller in its day, before being supplanted by the *philosophes*. It can be downloaded from the French Bibliothèque Nationale at www.bn.fr.

Hume's inspiration from Bayle: See Popkin (1951, 1955). Any reading of Bishop Huet (further down) would reveal the similarities with Hume.

Pre-Bayle thinkers*: Dissertation sur la recherche de la vérité*, Simon Foucher, from around 1673. It is a delight to read. It makes the heuristics and biases tradition look like the continuation of the pre-Enlightenment prescientific revolution atmosphere.

Bishop Huet and the problem of induction: "Things cannot be known with perfect certainty because their causes are infinite," wrote Pierre-Daniel Huet in his *Philosophical Treatise on the Weaknesses of the Human Mind*. Huet, former bishop of Avranches, wrote this under the name Théocrite de Pluvignac, Seigneur de la Roche, Gentilhomme de Périgord. The chapter has another exact presentation of what became later known as "Hume's problem." That was in 1690, when the future David Home (later Hume) was minus twenty-two, so of no possible influence on Monseigneur Huet.

Brochard's work: I first encountered the mention of Brochard's work (1888) in Nietzsche's *Ecce Homo,* in a comment where he also describes the skeptics as straight talkers. "An excellent study by Victor Brochard, *Les sceptiques grecs,* in which my Laertiana are also employed. The skeptics! the only *honourable* type among the two and five fold ambiguous philosopher crowd!" *More* trivia: Brochard taught Proust (see Kristeva, 1998).

Brochard seems to have understood Popper's problem (a few decades before Popper's birth). He presents the views of the negative empiricism of Menodotus of Nicomedia in similar terms to what we would call today "Popperian" empiricism. I wonder if Popper knew anything about Menodotus. He does not seem to quote him anywhere. Brochard published his doctoral thesis, *De l'erreur,* in 1878 at the University of Paris, on the subject of error—wonderfully modern.

Epilogism: We know very little about Menodotus except for attacks on his beliefs by his detractor Galen in the extant Latin version of the *Outline of Empiricism* (*Subfiguratio empirica),* hard to translate:

> *Memoriam et sensum et vocans* epilogismum *hoc tertium, multotiens autem et preter memoriam nihil aliud ponens quam* epilogismum. (In addition to perception and recollection, the third method is *epilogism sensum,* as the practitioner has, besides memory, nothing other than *epilogism senses;* Perilli's correction.

But there is hope. Perilli (2004) reports that, according to a letter by the translator Is-haq Bin Hunain, there may be a "transcription" of Menodotus's work in Arabic somewhere for a scholar to find.

Pascal: Pascal too had an idea of the confirmation problem and the asymmetry of inference. In his preface to the *Traité du vide,* Pascal writes (and I translate):

> In the judgment they made that nature did not tolerate a vacuum, they only meant nature in the state in which they knew it, since, so claim so in general, it would not be sufficient to witness it in a hundred different encounters, nor in a thousand, not in any other number no matter how large, since it would be a single case that would deny the general definition, and if one was contrary, a single one . . .

Hume's biographer: Mossner (1970). For a history of skepticism, Victor Cousin's lectures *Leçons d'histoire de la philosophie à la Sorbonne* (1828) and Hippolyte Taine's *Les philosophes classiques,* 9th edition (1868, 1905). Popkin (2003) is a modern account. Also see Heckman (2003) and Bevan (1913). I have seen nothing in the modern philosophy of probability linking it to skeptical inquiry.

Sextus: See Popkin (2003), Sextus, House (1980), Bayle, Huet, Annas and Barnes (1985), and Julia Anna and Barnes's introduction in Sextus Empiricus (2000). Favier (1906)

is hard to find; the only copy I located, thanks to Gur Huberman's efforts, was rotten—it seems that it has not been consulted in the past hundred years.

Menodotus of Nicomedia and the marriage between empiricism and skepticism: According to Brochard (1887), Menodotus is responsible for the mixing of empiricism and Pyrrhonism. See also Favier (1906). See skepticism about this idea in Dye (2004), and Perilli (2004).

Function not structure; empirical tripod: There are three sources, and three only, for experience to rely upon: observation, history (i.e., recorded observation), and judgment by analogy.

Algazel: See his *Tahafut al falasifah,* which is rebutted by Averroës, a.k.a. Ibn-Rushd, in *Tahafut Attahafut*.

Religious skeptics: There is also a medieval Jewish tradition, with the Arabic-speaking poet Yehuda Halevi. See Floridi (2002).

Algazel and the ultimate/proximate causation: ". . . their determining, from the sole observation, of the nature of the necessary relationship between the cause and the effect, as if one could not witness the effect without the attributed cause of the cause without the same effect." (*Tahafut*)

At the core of Algazel's idea is the notion that if you drink because you are thirsty, thirst should not be seen as a *direct* cause. There may be a greater scheme being played out; in fact, there *is,* but it can only be understood by those familiar with evolutionary thinking. See Tinbergen (1963, 1968) for a modern account of the proximate. In a way, Algazel builds on Aristotle to attack him. In his *Physics,* Aristotle had already seen the distinction between the different layers of cause (formal, efficient, final, and material).

Modern discussions on causality: See Reichenbach (1938), Granger (1999), and Pearl (2000).

Children and natural induction: See Gelman and Coley (1990), Gelman and Hirschfeld (1999), and Sloman (1993).

Natural induction: See Hespos (2006), Clark and Boyer (2006), Inagaki and Hatano (2006), Reboul (2006). See summary of earlier works in Plotkin (1998).

CHAPTERS 5–7

"Economists": What I mean by "economists" are most members of the mainstream, neoclassical economics and finance establishment in universities—not fringe groups such as the Austrian or the Post-Keynesian schools.

Small numbers: Tversky and Kahneman (1971), Rabin (2000).

Domain specificity: Williams and Connolly (2006). We can see it in the usually overinterpreted Wason Selection Test: Wason (1960, 1968). See also Shaklee and Fischhoff (1982), Barron Beaty, and Hearshly (1988). Kahneman's "They knew better" in Gilovich et al. (2002).

Updike: The blurb is from Jaynes (1976).

Brain hemispheric specialization: Gazzaniga and LeDoux (1978), Gazzaniga et al. (2005). Furthermore, Wolford, Miller, and Gazzaniga (2000) show probability matching by the left brain. When you supply the right brain with, say, a lever that produces desirable goods 60% of the time, and another lever 40%, the right brain will correctly push the first lever as the optimal policy. If, on the other hand, you supply the left brain with the same options, it will push the first lever 60 percent of the time and the other one 40—it will refuse to accept randomness. Goldberg (2005) argues that the specialty is along different lines: left-brain damage does not bear severe effects in children, unlike right-brain lesions, while this is the reverse for the elderly. I thank Elkhonon Goldberg for referring me to Snyder's work; Snyder (2001). The experiment is from Snyder et al. (2003).

Sock selection and retrofit explanation: The experiment of the socks is presented in Carter (1999); the original paper appears to be Nisbett and Wilson (1977). See also Montier (2007).

Astebro: Astebro (2003). See "Searching for the Invisible Man," *The Economist*, March 9, 2006. To see how the overconfidence of entrepreneurs can explain the high failure rate, see Camerer (1995).

Dopamine: Brugger and Graves (1997), among many other papers. See also Mohr et al. (2003) on dopamine asymmetry.

Entropy and information: I am purposely avoiding the notion of entropy because the way it is conventionally phrased makes it ill-adapted to the type of randomness we experience in real life. Tsallis entropy works better with fat tails.

Notes on George Perec: Eco (1994).

Narrativity and illusion of understanding: Wilson, Gilbert, and Centerbar (2003): "Helplessness theory has demonstrated that if people feel that they cannot control or predict their environments, they are at risk for severe motivational and cognitive deficits, such as depression." For the writing down of a diary, see Wilson (2002) or Wegner (2002).

E. M. Forster's example: reference in Margalit (2002).

National character: Terracciano et al. (2005) and Robins (2005) for the extent of individual variations. The illusion of nationality trait, which I usually call the "nationality heuristic," does connect to the halo effect: see Rosenzweig (2006) and Cialdini (2001). See Anderson (1983) for the ontology of nationality.

Consistency bias: What psychologists call the consistency bias is the effect of revising memories in such a way to make sense with respect to subsequent information. See Schacter (2001).

Memory not like storage on a computer: Rose (2003), Nader and LeDoux (1999).

The myth of repressed memory: Loftus and Ketcham (2004).

Chess players and disconfirmation: Cowley and Byrne (2004).

Quine's problem: Davidson (1983) argues in favor of local, but against total, skepticism.

Narrativity: Note that my discussion is not existential here, but merely practical, so my idea is to look at narrativity as an informational compression, nothing more involved philosophically (like whether a self is sequential or not). There is a literature on the "narrative self"—Bruner (2002) or whether it is necessary—see Strawson (1994) and his attack in Strawson (2004). The debate: Schechtman (1997), Taylor (1999), Phelan (2005). Synthesis in Turner (1996).

"Postmodernists" and the desirability of narratives: See McCloskey (1990) and Frankfurter and McGoun (1996).

Narrativity of sayings and proverbs: Psychologists have long examined the gullibility of people in social settings when faced with well-sounding proverbs. For instance, experiments have been made since the 1960s where people are asked whether they believe that a proverb is right, while another cohort is presented with the opposite meaning. For a presentation of the hilarious results, see Myers (2002).

Science as a narrative: Indeed scientific papers can succeed by the same narrativity bias that "makes a story." You need to get attention. Bushman and Wells (2001).

Discovering probabilities: Barron and Erev (2003) show how probabilities are underestimated when they are not explicitly presented. Also personal communication with Barron.

Risk and probability: See Slovic, Fischhoff, and Lichtenstein (1976), Slovic et al. (1977), and Slovic (1987). For risk as analysis and risk as feeling theory, see Slovic et al. (2002, 2003), and Taleb (2004c). See Bar-Hillel and Wagenaar (1991).

Link between narrative fallacy and clinical knowledge: Dawes (1999) has a message for economists: see here his work on interviews and the concoction of a narrative. See also Dawes (2001) on the retrospective effect.

Two systems of reasoning: See Sloman (1996, 2002), and the summary in Kahneman and Frederick (2002). Kahneman's Nobel lecture sums it all up; it can be found at www.nobel.se. See also Stanovich and West (2000).

Risk and emotions: Given the growing recent interest in the emotional role in behavior, there has been a growing literature on the role of emotions in both risk bearing and

risk avoidance: the "risk as feeling" theory. See Loewenstein et al. (2001) and Slovic et al. (2003a). For a survey see Slovic et al. (2003b) and see also Slovic (1987). For a discussion of the "affect heuristic," see Finucane et al. (2000). For modularity, see Bates (1994).

Emotions and cognition: For the effect of emotions on cognition, see LeDoux (2002). For risk, see Bechara et al. (1994).

Availability heuristic (how easily things come to mind): See Tversky and Kahneman (1973).

Real incidence of catastrophes: For an insightful discussion, see Albouy (2002), Zajdenweber (2000), or Sunstein (2002).

Terrorism exploitation of the sensational: See the essay in Taleb (2004c).

General books on psychology of decision making (heuristics and biases): Baron (2000) is simply the most comprehensive on the subject. Kunda (1999) is a summary from the standpoint of social psychology (sadly, the author died prematurely); shorter: Plous (1993). Also Dawes (1988) and Dawes (2001). Note that a chunk of the original papers are happily compiled in Kahneman et al. (1982), Kahneman and Tversky (2000), Gilovich et al. (2002), and Slovic (2001a and 2001b). See also Myers (2002) for an account on intuition and Gigerenzer et al. (2000) for an ecological presentation of the subject. The most complete account in economics and finance is Montier (2007), where his beautiful summary pieces that fed me for the last four years are compiled—not being an academic, he gets straight to the point. See also Camerer, Loewenstein, and Rabin (2004) for a selection of technical papers. A recommended review article on clinical "expert" knowledge is Dawes (2001).

More general psychology of decision presentations: Klein (1998) proposes an alternative model of intuition. See Cialdini (2001) for social manipulation. A more specialized work, Camerer (2003), focuses on game theory.

General review essays and comprehensive books in cognitive science: Newell and Simon (1972), Varela (1988), Fodor (1983), Marr (1982), Eysenck and Keane (2000), Lakoff and Johnson (1980). The *MIT Encyclopedia of Cognitive Science* has review articles by main thinkers.

Evolutionary theory and domains of adaptation: See the original Wilson (2000), Kreps and Davies (1993), and Burnham (1997, 2003). Very readable: Burnham and Phelan (2000). The compilation of Robert Trivers's work is in Trivers (2002). See also Wrangham (1999) on wars.

Politics: "The Political Brain: A Recent Brain-imaging Study Shows That Our Political Predilections Are a Product of Unconscious Confirmation Bias," by Michael Shermer, *Scientific American,* September 26, 2006.

Neurobiology of decision making: For a general understanding of our knowledge about the brain's architecture: Gazzaniga et al. (2002). Gazzaniga (2005) provides literary summaries of some of the topics. More popular: Carter (1999). Also recommended: Ratey (2001), Ramachandran (2003), Ramachandran and Blakeslee (1998), Carter (1999, 2002), Conlan (1999), the very readable Lewis, Amini, and Lannon (2000), and Goleman (1995). See Glimcher (2002) for probability and the brain. For the emotional brain, the three books by Damasio (1994, 2000, 2003), in addition to LeDoux (1998) and the more detailed LeDoux (2002), are the classics. See also the shorter Evans (2002). For the role of vision in aesthetics, but also in interpretation, Zeki (1999).

General works on memory: In psychology, Schacter (2001) is a review work of the memory biases with links to the hindsight effects. In neurobiology, see Rose (2003) and Squire and Kandel (2000). A general textbook on memory (in empirical psychology) is Baddeley (1997).

Intellectual colonies and social life: See the account in Collins (1998) of the "lineages" of philosophers (although I don't think he was aware enough of the Casanova problem to take into account the bias making the works of solo philosophers less likely to survive). For an illustration of the aggressiveness of groups, see Uglow (2003).

Hyman Minsky's work: Minsky (1982).

Asymmetry: Prospect theory (Kahneman and Tversky [1979] and Tversky and Kahneman [1992]) accounts for the asymmetry between bad and good random events, but it also shows that the negative domain is convex while the positive domain is concave, meaning that a loss of 100 is less painful than 100 losses of 1 but that a gain of 100 is also far less pleasurable than 100 times a gain of 1.

Neural correlates of the asymmetry: See Davidson's work in Goleman (2003), Lane et al. (1997), and Gehring and Willoughby (2002). Csikszentmihalyi (1993, 1998) further explains the attractiveness of steady payoffs with his theory of "flow."

Deferred rewards and its neural correlates: McLure et al. (2004) show the brain activation in the cortex upon making a decision to defer, providing insight on the limbic impulse behind immediacy and the cortical activity in delaying. See also Loewenstein et al. (1992), Elster (1998), Berridge (2005). For the neurology of preferences in Capuchin monkeys, Chen et al. (2005).

Bleed or blowup: Gladwell (2002) and Taleb (2004c). Why bleed is painful can be explained by dull stress; Sapolsky et al. (2003) and Sapolsky (1998). For how companies like steady returns, Degeorge and Zeckhauser (1999). Poetics of hope: Mihailescu (2006).

Discontinuities and jumps: Classified by René Thom as constituting seven classes; Thom (1980).

Evolution and small probabilities: Consider also the naïve evolutionary thinking positing the "optimality" of selection. The founder of sociobiology, the great E. O. Wilson, does not agree with such optimality when it comes to rare events. In Wilson (2002), he writes:

> The human brain evidently evolved to commit itself emotionally only to a small piece of geography, a limited band of kinsmen, and two or three generations into the future. To look neither far ahead nor far afield is elemental in a Darwinian sense. *We are innately inclined to ignore any distant possibility not yet requiring examination. It is, people say, just good common sense.* Why do they think in this shortsighted way?
>
> The reason is simple: it is a hardwired part of our Paleolithic heritage. For hundreds of millennia, those who worked for short-term gain within a small circle of relatives and friends lived longer and left more offspring—even when their collective striving caused their chiefdoms and empires to crumble around them. The long view that might have saved their distant descendants required a vision and extended altruism instinctively difficult to marshal.

See also Miller (2000): *"Evolution has no foresight. It lacks the long-term vision of drug company management. A species can't raise venture capital to pay its bills while its research team . . . This makes it hard to explain innovations."*

Note that neither author considered my age argument.

CHAPTER 8

Silent evidence bears the name *wrong reference class* in the nasty field of philosophy of probability, *anthropic bias* in physics, and *survivorship bias* in statistics (economists present the interesting attribute of having rediscovered it a few times while being severely fooled by it).

Confirmation: Bacon says in *On Truth*, "No pleasure is comparable to the standing upon the vantage ground of truth (a hill not to be commanded and where the air is always clear and serene), and to see the errors, and wanderings, and mists, and tempests, in the vale below." This easily shows how great intentions can lead to the confirmation fallacy.

Bacon did not understand the empiricists: He was looking for the golden mean. Again, from *On Truth:*

> There are three sources of error and three species of false philosophy; the sophistic, the empiric and the superstitious. . . . Aristotle affords the most eminent instance of the first; for he corrupted natural philosophy by logic—thus he formed the world of categories. . . . Nor is much stress to be laid on his frequent recourse to experiment in his books on animals, his problems and other treatises, for he had already decided, without having properly consulted experience as the basis of his decisions and axioms. . . . The empiric school produces dogmas of a more deformed and monstrous nature than the sophistic or theoretic school; not being founded in the light of common notions (which however poor and superstitious, is yet in a manner universal and of general tendency), but in the confined obscurity of a few experiments.

Bacon's misconception may be the reason it took us a while to understand that they treated history (and experiments) as mere and vague "guidance," i.e., epilogy.

Publishing: Allen (2005), Klebanoff (2002), Epstein (2001), de Bellaigue (2004), and Blake (1999). For a funny list of rejections, see Bernard (2002) and White (1982). Michael Korda's memoir, Korda (2000), adds some color to the business. These books are anecdotal, but we will see later that books follow steep scale-invariant structures with the implication of a severe role for randomness.

Anthropic bias: See the wonderful and comprehensive discussion in Bostrom (2002). In physics, see Barrow and Tipler (1986) and Rees (2004). Sornette (2004) has Gott's derivation of survival as a power law. In finance, Sullivan et al. (1999) discuss survivorship bias. See also Taleb (2004a). Studies that ignore the bias and state inappropriate conclusions: Stanley and Danko (1996) and the more foolish Stanley (2000).

Manuscripts and the Phoenicians: For survival and science, see Cisne (2005). Note that the article takes into account physical survival (like fossil), not cultural, which implies a selection bias. Courtesy Peter Bevelin.

Stigler's law of eponymy: Stigler (2002).

French book statistics: *Lire,* April 2005.

Why dispersion matters: More technically, the distribution of the extremum (i.e., the maximum or minimum) of a random variable depends more on the variance of the process than on its mean. Someone whose weight tends to fluctuate a lot is more likely to show you a picture of himself very thin than someone else whose weight is on average lower but remains constant. The mean (read skills) sometimes plays a very, very small role.

Fossil record: I thank the reader Frederick Colbourne for his comments on this subject. The literature calls it the "pull of the recent," but has difficulty estimating the effects, owing to disagreements. See Jablonski et al. (2003).

Undiscovered public knowledge: Here is another manifestation of silent evidence: you can actually do lab work sitting in an armchair, just by linking bits and pieces of research by people who labor apart from one another and miss on connections. Using bibliographic analysis, it is possible to find links between published information that had not been known previously by researchers. I "discovered" the vindication of the armchair in Fuller (2005). For other interesting discoveries, see Spasser (1997) and Swanson (1986a, 1986b, 1987).

Crime: The definition of economic "crime" is something that comes in hindsight. Regulations, once enacted, do not run retrospectively, so many activities causing excess are never sanctioned (e.g., bribery).

Bastiat: See Bastiat (1862–1864).

Casanova: I thank the reader Milo Jones for pointing out to me the exact number of volumes. See Masters (1969).

Reference point problem: Taking into account background information requires a form of thinking in *conditional* terms that, oddly, many scientists (especially the better ones) are incapable of handling. The difference between the two odds is called, simply, conditional probability. We are computing the probability of surviving *conditional* on our being in the sample itself. Simply put, you cannot compute probabilities if your survival is part of the condition of the realization of the process.

Plagues: See McNeill (1976).

CHAPTER 9

Intelligence and Nobel: Simonton (1999). If IQ scores correlate, they do so very weakly with subsequent success.

"Uncertainty": Knight (1923). My definition of such risk (Taleb, 2007c) is that it is a normative situation, where we can be certain about probabilities, i.e., no metaprobabilities. Whereas, if randomness and risk result from epistemic opacity, the difficulty in seeing causes, then necessarily the distinction is bunk. Any reader of Cicero would recognize it as his probability; see epistemic opacity in his *De Divinatione*, Liber primus, LVI, 127:

> Qui enim teneat causas rerum futurarum, idem necesse est omnia teneat quae futura sint. Quod cum nemo facere nisi deus possit, relinquendum est homini, ut signis quibusdam consequentia declarantibus futura praesentiat.

"He who knows the causes will understand the future, except that, given that nobody outside God possesses such faculty . . ."

Philosophy and epistemology of probability: Laplace. *Treatise,* Keynes (1920), de Finetti (1931), Kyburg (1983), Levi (1970), Ayer, Hacking (1990, 2001), Gillies (2000), von Mises (1928), von Plato (1994), Carnap (1950), Cohen (1989), Popper (1971), Eatwell et al. (1987), and Gigerenzer et al. (1989).

History of statistical knowledge and methods: I found no intelligent work in the history of statistics, i.e., one that does not fall prey to the ludic fallacy or Gaussianism. For a conventional account, see Bernstein (1996) and David (1962).

General books on probability and information theory: Cover and Thomas (1991); less technical but excellent, Bayer (2003). For a probabilistic view of information theory: the posthumous Jaynes (2003) is the only mathematical book other than de Finetti's work that I can recommend to the general reader, owing to his Bayesian approach and his allergy for the formalism of the idiot savant.

Poker: It escapes the ludic fallacy; see Taleb (2006a).

Plato's normative approach to left and right hands: See McManus (2002).

Nietzsche's *bildungsphilister:* See van Tongeren (2002) and Hicks and Rosenberg (2003). Note that because of the confirmation bias academics will tell you that intellectuals "lack rigor," and will bring examples of those who do, not those who don't.

Economics books that deal with uncertainty: Carter et al. (1962), Shackle (1961, 1973), Hayek (1994). Hirshleifer and Riley (1992) fits uncertainty into neoclassical economics.

Incomputability: For earthquakes, see Freedman and Stark (2003) (courtesy of Gur Huberman).

Academia and philistinism: There is a round-trip fallacy; if academia means rigor (which I doubt, since what I saw called "peer reviewing" is too often a masquerade), nonacademic does not imply nonrigorous. Why do I doubt the "rigor"? By the confirmation bias they show you their contributions yet in spite of the high number of laboring academics, a relatively minute fraction of our results come from them. A disproportionately high number of contributions come from freelance researchers and those dissingly called amateurs: Darwin, Freud, Marx, Mandelbrot, even the early Einstein. Influence on the part of an academic is usually accidental. This even held in the Mid-

dle Ages and the Renaissance, see Le Goff (1985). Also, the Enlightenment figures (Voltaire, Rousseau, d'Holbach, Diderot, Montesquieu) were all nonacademics at a time when academia was large.

CHAPTER 10

Overconfidence: Albert and Raiffa (1982) (though apparently the paper languished for a decade before formal publication). Lichtenstein and Fischhoff (1977) showed that overconfidence can be influenced by item difficulty; it typically diminishes and turns into underconfidence in easy items (compare with Armelius [1979]). Plenty of papers since have tried to pin down the conditions of calibration failures or robustness (be they task training, ecological aspects of the domain, level of education, or nationality): Dawes (1980), Koriat, Lichtenstein, and Fischhoff (1980), Mayseless and Kruglanski (1987), Dunning et al. (1990), Ayton and McClelland (1997), Gervais and Odean (1999), Griffin and Varey (1996), Juslin (1991, 1993, 1994), Juslin and Olsson (1997), Kadane and Lichtenstein (1982), May (1986), McClelland and Bolger (1994), Pfeifer (1994), Russo and Schoernaker (1992), Klayman et al. (1999). Note the decrease (unexpectedly) in overconfidence under group decisions: see Sniezek and Henry (1989)—and solutions in Plous (1995). I am suspicious here of the Mediocristan/Extremistan distinction and the unevenness of the variables. Alas, I found no paper making this distinction. There are also solutions in Stoll (1996), Arkes et al. (1987). For overconfidence in finance, see Thorley (1999) and Barber and Odean (1999). For cross-boundaries effects, Yates et al. (1996, 1998), Angele et al. (1982). For simultaneous overconfidence and underconfidence, see Erev, Wallsten, and Budescu (1994).

Frequency vs. probability—the ecological problem: Hoffrage and Gigerenzer (1998) think that overconfidence is less significant when the problem is expressed in frequencies as opposed to probabilities. In fact, there has been a debate about the difference between "ecology" and laboratory; see Gigerenzer et al. (2000), Gigerenzer and Richter (1990), and Gigerenzer (1991). We are "fast and frugal" (Gigerenzer and Goldstein [1996]). As far as the Black Swan is concerned, these problems of ecology do not arise: we do not live in an environment in which we are supplied with frequencies or, more generally, for which we are fit. Also in ecology, Spariosu (2004) for the ludic aspect, Cosmides and Tooby (1990). Leary (1987) for Brunswikian ideas, as well as Brunswik (1952).

Lack of awareness of ignorance: "In short, the same knowledge that underlies the ability to produce correct judgment is also the knowledge that underlies the ability to recognize correct judgment. To lack the former is to be deficient in the latter." From Kruger and Dunning (1999).

Expert problem in isolation: I see the expert problem as indistinguishable from Matthew effects and Extremism fat tails (more later), yet I found no such link in the literatures of sociology and psychology.

Clinical knowledge and its problems: See Meehl (1954) and Dawes, Faust, and Meehl (1989). Most entertaining is the essay "Why I Do Not Attend Case Conferences" in Meehl (1973). See also Wagenaar and Keren (1985, 1986).

Financial analysts, herding, and forecasting: See Guedj and Bouchaud (2006), Abarbanell and Bernard (1992), Chen et al. (2002), De Bondt and Thaler (1990), Easterwood and Nutt (1999), Friesen and Weller (2002), Foster (1977), Hong and Kubik (2003), Jacob et al. (1999), Lim (2001), Liu (1998), Maines and Hand (1996), Mendenhall (1991), Mikhail et al. (1997, 1999), Zitzewitz (2001), and El-Galfy and Forbes (2005). For a comparison with weather forecasters (unfavorable): Tyszka and Zielonka (2002).

Economists and forecasting: Tetlock (2005), Makridakis and Hibon (2000), Makridakis et al. (1982), Makridakis et al. (1993), Gripaios (1994), Armstrong (1978, 1981);

and rebuttals by McNees (1978), Tashman (2000), Blake et al. (1986), Onkal et al. (2003), Gillespie (1979), Baron (2004), Batchelor (1990, 2001), Dominitz and Grether (1999). Lamont (2002) looks for reputational factors: established forecasters get worse as they produce more radical forecasts to get attention—consistent with Tetlock's hedgehog effect. Ahiya and Doi (2001) look for herd behavior in Japan. See McNees (1995), Remus et al. (1997), O'Neill and Desai (2005), Bewley and Fiebig (2002), Angner (2006), Bénassy-Quéré (2002); Brender and Pisani (2001) look at the Bloomberg consensus; De Bondt and Kappler (2004) claim evidence of weak persistence in fifty-two years of data, but I saw the slides in a presentation, never the paper, which after two years might never materialize. Overconfidence, Braun and Yaniv (1992). See Hahn (1993) for a general intellectual discussion. More general, Clemen (1986, 1989). For Game theory, Green (2005).

Many operators, such as James Montier, and many newspapers and magazines (such as *The Economist*), run casual tests of prediction. Cumulatively, they must be taken seriously since they cover more variables.

Popular culture: In 1931, Edward Angly exposed forecasts made by President Hoover in a book titled *Oh Yeah?* Another hilarious book is Cerf and Navasky (1998), where, incidentally, I got the pre-1973 oil-estimation story.

Effects of information: The major paper is Bruner and Potter (1964). I thank Danny Kahneman for discussions and pointing out this paper to me. See also Montier (2007), Oskamp (1965), and Benartzi (2001). These biases become ambiguous information (Griffin and Tversky [1992]). For how they fail to disappear with expertise and training, see Kahneman and Tversky (1982) and Tversky and Kahneman (1982). See Kunda (1990) for how preference-consistent information is taken at face value, while preference-inconsistent information is processed critically.

Planning fallacy: Kahneman and Tversky (1979) and Buehler, Griffin, and Ross (2002). The planning fallacy shows a consistent bias in people's planning ability, even with matters of a repeatable nature—though it is more exaggerated with nonrepeatable events.

Wars: Trivers (2002).

Are there incentives to delay?: Flyvbjerg et al. (2002).

Oskamp: Oskamp (1965) and Montier (2007).

Task characteristics and effect on decision making: Shanteau (1992).

Epistēmē* vs. *Technē: This distinction harks back to Aristotle, but it recurs then dies? down—it most recently recurs in accounts such as tacit knowledge in "know how." See Ryle (1949), Polanyi (1958/1974), and Mokyr (2002).

Catherine the Great: The number of lovers comes from Rounding (2006).

Life expectancy: www.annuityadvantage.com/lifeexpectancy.htm. For projects, I have used a probability of exceeding with a power-law exponent of 3/2: $f = Kx^{3/2}$. Thus the conditional expectation of x, knowing that x exceeds a

$$\mathrm{E}[x|x>a] = \frac{\int_a^\infty x f(x)dx}{\int_a^\infty f(x)dx}.$$

CHAPTERS 11–13

Serendipity: See Koestler (1959) and Rees (2004). Rees also has powerful ideas on forecastability. See also Popper's comments in Popper (2002), and Waller (2002a), Cannon (1940), Mach (1896) (cited in Simonton [1999]), and Merton and Barber (2004). See Simonton (2004) for a synthesis. For serendipity in medicine and anesthesiology, see Vale et al. (2005).

"**Renaissance man**": See www.bell-labs.com/project/feature/archives/cosmology/.

Laser: As usual, there are controversies as to who "invented" the technology. After a successful discovery, precursors are rapidly found, owing to the retrospective distortion.

Charles Townsend won the Nobel prize, but was sued by his student Gordon Gould, who held that he did the actual work (see *The Economist,* June 9, 2005).

Darwin/Wallace: Quammen (2006).

Popper's attack on historicism: See Popper (2002). Note that I am reinterpreting Popper's idea in a modern manner here, using my own experiences and knowledge, not commenting on comments about Popper's work—with the consequent lack of fidelity to his message. In other words, these are not directly Popper's arguments, but largely mine phrased in a Popperian framework. The conditional expectation of an unconditional expectation is an unconditional expectation.

Forecast for the future a hundred years earlier: Bellamy (1891) illustrates our mental projections of the future. However, some stories might be exaggerated: "A Patently False Patent Myth still! Did a patent official really once resign because he thought nothing was left to invent? Once such myths start they take on a life of their own." *Skeptical Inquirer,* May–June, 2003.

Observation by Peirce: Olsson (2006), Peirce (1955).

Predicting and explaining: See Thom (1993).

Poincaré: The three body problem can be found in Barrow-Green (1996), Rollet (2005), and Galison (2003). On Einstein, Pais (1982). More recent revelations in Hladik (2004).

Billiard balls: Berry (1978) and Pisarenko and Sornette (2004).

Very general discussion on "complexity": Benkirane (2002), Scheps (1996), and Ruelle (1991). For limits, Barrow (1998).

Hayek: See www.nobel.se. See Hayek (1945, 1994). Is it that mechanisms do not correct themselves from railing by influential people, but either by mortality of the operators, or something even more severe, by being put out of business? Alas, because of contagion, there seems to be little logic to how matters improve; luck plays a part in how soft sciences evolve. See Ormerod (2006) for network effects in "intellectuals and socialism" and the power-law distribution in influence owing to the scale-free aspect of the connections—and the consequential arbitrariness. Hayek seems to have been a prisoner of Weber's old differentiation between *Natur-Wissenschaften* and *Geistes Wissenschaften*—but thankfully not Popper.

Insularity of economists: Pieters and Baumgartner (2002). One good aspect of the insularity of economists is that they can insult me all they want without any consequence: it appears that only economists read other economists (so they can write papers for other economists to read). For a more general case, see Wallerstein (1999). Note that Braudel fought "economic history." It was history.

Economics as religion: Nelson (2001) and Keen (2001). For methodology, see Blaug (1992). For high priests and lowly philosophers, see Boettke, Coyne, and Leeson (2006). Note that the works of Gary Becker and the Platonists of the Chicago School are all marred by the confirmation bias: Becker is quick to show you situations in which people are moved by economic incentives, but does not show you cases (vastly more numerous) in which people don't care about such materialistic incentives.

The smartest book I've seen in economics is Gave et al. (2005) since it transcends the constructed categories in academic economic discourse (one of the authors is the journalist Anatole Kaletsky).

General theory: This fact has not deterred "general theorists." One hotshot of the Platonifying variety explained to me during a long plane ride from Geneva to New York that the ideas of Kahneman and his colleagues must be rejected because they do not allow us to develop a general equilibrium theory, producing "time-inconsistent preferences." For a minute I thought he was joking: he blamed the psychologists' ideas and human incoherence for interfering with his ability to build his Platonic model.

Samuelson: For his optimization, see Samuelson (1983). Also Stiglitz (1994).

Plato's dogma on body symmetry: "Athenian Stranger to Cleinias: In that the right and left hand are supposed to be by nature differently suited for our various uses of them;

whereas no difference is found in the use of the feet and the lower limbs; but in the use of the hands we are, as it were, maimed by the folly of nurses and mothers; for although our several limbs are by nature balanced, we create a difference in them by bad habit," in Plato's *Laws*. See McManus (2002).

Drug companies: Other such firms, I was told, are run by commercial persons who tell researchers where they find a "market need" and ask them to "invent" drugs and cures accordingly—which accords with the methods of the dangerously misleading Wall Street security analysts. They formulate projections as if they know what they are going to find.

Models of the returns on innovations: Sornette and Zajdenweber (1999) and Silverberg and Verspagen (2005).

Evolution on a short leash: Dennet (2003) and Stanovich and West (2000).

Montaigne: We don't get much from the biographies of a personal essayist; some information in Frame (1965) and Zweig (1960).

Projectibility and the grue paradox: See Goodman (1955). See also an application (or perhaps misapplication) in King and Zheng (2005).

Constructionism: See Berger and Luckmann (1966) and Hacking (1999).

Certification vs, true skills or knowledge: See Donhardt (2004). There is also a franchise protection. Mathematics may not be so necessary a tool for economics, except to protect the franchise of those economists who know math. In my father's days, the selection process for the mandarins was made using their abilities in Latin (or Greek). So the class of students groomed for the top was grounded in the classics and knew some interesting subjects. They were also trained in Cicero's highly probabilistic view of things—and selected on erudition, which carries small side effects. If anything it allows you to handle fuzzy matters. My generation was selected according to mathematical skills. You made it based on an engineering mentality; this produced mandarins with mathematical, highly structured, logical minds, and, accordingly, they will select their peers based on such criteria. So the papers in economics and social science gravitated toward the highly mathematical and protected their franchise by putting high mathematical barriers to entry. You could also smoke the general public who is unable to put a check on you. Another effect of this franchise protection is that it might have encouraged putting "at the top" those idiot-savant-like researchers who lacked in erudition, hence were insular, parochial, and closed to other disciplines.

Freedom and determinism: a speculative idea in Penrose (1989) where only the quantum effects (with the perceived indeterminacy there) can justify consciousness.

Projectibility: uniqueness assuming least squares or MAD.

Chaos theory and the backward/forward confusion: Laurent Firode's *Happenstance,* a.k.a. *Le battement d'ailes du papillon / The Beating of a Butterfly's Wings* (2000).

Autism and perception of randomness: See Williams et al. (2002).

Forecasting and misforecasting errors in hedonic states: Wilson, Meyers, and Gilbert (2001), Wilson, Gilbert, and Centerbar (2003), and Wilson et al. (2005). They call it "emotional evanescence."

Forecasting and consciousness: See the idea of "aboutness" in Dennett (1995, 2003) and Humphrey (1992). However, Gilbert (2006) believes that we are not the only animal that forecasts—which is wrong as it turned out. Suddendorf (2006) and Dally, Emery, and Clayton (2006) show that animals too forecast!

Russell's comment on Pascal's wager: Ayer (1988) reports this as a private communication.

History: Carr (1961), Hexter (1979), and Gaddis (2002). But I have trouble with historians throughout, because they often mistake the forward and the backward processes. Mark Buchanan's *Ubiquity* and the quite confused discussion by Niall Ferguson in *Nature*. Neither of them seem to realize the problem of calibration with power laws. See also Ferguson, *Why Did the Great War?,* to gauge the extent of the forward-backward problems.

For the traditional nomological tendency, i.e., the attempt to go beyond cause into a general theory, see *Muqaddamah* by Ibn Khaldoun. See also Hegel's *Philosophy of History*.

Emotion and cognition: Zajonc (1980, 1984).

Catastrophe insurance: Froot (2001) claims that insurance for remote events is overpriced. How he determined this remains unclear (perhaps by backfitting or bootstraps), but reinsurance companies have not been making a penny selling "overpriced" insurance.

Postmodernists: Postmodernists do not seem to be aware of the differences between narrative and prediction.

Luck and serendipity in medicine: Vale et al. (2005). In history, see Cooper (2004). See also Ruffié (1977). More general, see Roberts (1989).

Affective forecasting: See Gilbert (1991), Gilbert et al. (1993), and Montier (2007).

CHAPTERS 14–17

This section will also serve another purpose. Whenever I talk about the Black Swan, people tend to supply me with anecdotes. But these anecdotes are just corroborative: you need to show that *in the aggregate* the world is dominated by Black Swan events. To me, the rejection of nonscalable randomness is sufficient to establish the role and significance of Black Swans.

Matthew effects: See Merton (1968, 1973a, 1988). Martial, in his *Epigrams*: "*Semper pauper eris, si pauper es, Aemiliane./Dantur opes nullis (nunc) nisi divitibus.*" (Epigr. V 81). See also Zuckerman (1997, 1998).

Cumulative advantage and its consequences on social fairness: review in DiPrete et al. (2006). See also Brookes-Gun and Duncan (1994), Broughton and Mills (1980), Dannefer (2003), Donhardt (2004), Hannon (2003), and Huber (1998). For how it may explain precocity, see Elman and O'Rand (2004).

Concentration and fairness in intellectual careers: Cole and Cole (1973), Cole (1970), Conley (1999), Faia (1975), Seglen (1992), Redner (1998), Lotka (1926), Fox and Kochanowski (2004), and Huber (2002).

Winner take all: Rosen (1981), Frank (1994), Frank and Cook (1995), and Attewell (2001).

Arts: Bourdieu (1996), Taleb (2004e).

Wars: War is concentrated in an Extremistan manner: Lewis Fry Richardson noted last century the uneveness in the distribution of casualties (Richardson [1960]).

Modern wars: Arkush and Allen (2006). In the study of the Maori, the pattern of fighting with clubs was sustainable for many centuries—modern tools cause 20,000 to 50,000 deaths a year. We are simply not made for technical warfare. For an anecdotal and causative account of the history of a war, see Ferguson (2006).

S&P 500: See Rosenzweig (2006).

The long tail: Anderson (2006).

Cognitive diversity: See Page (2007). For the effect of the Internet on schools, see Han et al. (2006).

Cascades: See Schelling (1971, 1978) and Watts (2002). For information cascades in economics, see Bikhchandani, Hirshleifer, and Welch (1992) and Shiller (1995). See also Surowiecki (2004).

Fairness: Some researchers, like Frank (1999), see arbitrary and random success by others as no different from pollution, which necessitates the enactment of a tax. De Vany, Taleb, and Spitznagel (2004) propose a market-based solution to the problem of allocation through the process of voluntary self-insurance and derivative products. Shiller (2003) proposes cross-country insurance.

The mathematics of preferential attachment: This argument pitted Mandelbrot against the cognitive scientist Herbert Simon, who formalized Zipf's ideas in a 1955 paper (Simon [1955]), which then became known as the Zipf-Simon model. Hey, you need to allow for people to fall from favor!

Concentration: Price (1970). Simon's "Zipf derivation," Simon (1955). More general bibliometrics, see Price (1976) and Glänzel (2003).

Creative destruction revisited: See Schumpeter (1942).

Networks: Barabási and Albert (1999), Albert and Barabási (2000), Strogatz (2001, 2003), Callaway et al. (2000), Newman et al. (2000), Newman, Watts, and Strogatz (2000), Newman (2001), Watts and Strogatz (1998), Watts (2002, 2003), and Amaral et al. (2000). It supposedly started with Milgram (1967). See also Barbour and Reinert (2000), Barthélémy and Amaral (1999). See Boots and Sasaki (1999) for infections. For extensions, see Bhalla and Iyengar (1999). Resilence, Cohen et al. (2000), Barabási and Bonabeau (2003), Barabási (2002), and Banavar et al. (2000). Power laws and the Web, Adamic and Huberman (1999) and Adamic (1999). Statistics of the Internet: Huberman (2001), Willinger et al. (2004), and Faloutsos, Faloutsos, and Faloutsos (1999). For DNA, see Vogelstein et al. (2000).

Self-organized criticality: Bak (1996).

Pioneers of fat tails: For wealth, Pareto (1896), Yule (1925, 1944). Less of a pioneer Zipf (1932, 1949). For linguistics, see Mandelbrot (1952).

Pareto: See Bouvier (1999).

Endogenous vs. exogenous: Sornette et al. (2004).

Sperber's work: Sperber (1996a, 1996b, 1997).

Regression: If you hear the phrase *least square regression,* you should be suspicious about the claims being made. As it assumes that your errors wash out rather rapidly, it underestimates the total possible error, and thus overestimates what knowledge one can derive from the data.

The notion of central limit: very misunderstood: it takes a long time to reach the central limit—so as we do not live in the asymptote, we've got problems. All various random variables (as we started in the example of Chapter 16 with a +1 or –1, which is called a Bernouilli draw) under summation (we did sum up the wins of the 40 tosses) become Gaussian. Summation is key here, since we are considering the results of adding up the 40 steps, which is where the Gaussian, under the first and second central assumptions becomes what is called a "distribution." (A distribution tells you how you are likely to have your outcomes spread out, or distributed.) However, they may get there at different speeds. This is called the central limit theorem: if you add random variables coming from these individual tame jumps, it will lead to the Gaussian.

Where does the central limit not work? If you do not have these central assumptions, but have jumps of random size instead, then we would not get the Gaussian. Furthermore, we sometimes converge very slowly to the Gaussian. For preasymptotics and scalability, Mandelbrot and Taleb (2007a), Bouchaud and Potters (2003). For the problem of working outside asymptotes, Taleb (2007).

Aureas mediocritas: historical perspective, in Naya and Pouey-Mounou (2005) aptly called *Éloge de la médiocrité.*

Reification (hypostatization): Lukacz, in Bewes (2002).

Catastrophes: Posner (2004).

Concentration and modern economic life: Zajdenweber (2000).

Choices of society structure and compressed outcomes: The classical paper is Rawls (1971), though Frohlich, Oppenheimer, and Eavy (1987a, 1987b), as well as Lissowski, Tyszka, and Okrasa (1991), contradict the notion of the desirability of Rawl's veil (though by experiment). People prefer maximum average income subjected to a floor constraint on some form of equality for the poor, inequality for the rich type of environment.

Gaussian contagion: Quételet in Stigler (1986). Francis Galton (as quoted in Ian Hacking's *The Taming of Chance*): "I know of scarcely anything so apt to impress the imagination as the wonderful form of cosmic order expressed by 'the law of error.' "

"Finite variance" nonsense: Associated with CLT is an assumption called "finite variance" that is rather technical: none of these building-block steps can take an infinite value if you square them or multiply them by themselves. They need to be bounded

at some number. We simplified here by making them all one single step, or finite standard deviation. But the problem is that some fractal payoffs may have finite variance, but still not take us there rapidly. See Bouchaud and Potters (2003).

Lognormal: There is an intermediate variety that is called the lognormal, emphasized by one Gibrat (see Sutton [1997]) early in the twentieth century as an attempt to explain the distribution of wealth. In this framework, it is not quite that the wealthy get wealthier, in a pure preferential attachment situation, but that if your wealth is at 100 you will vary by 1, but when your wealth is at 1,000, you will vary by 10. The relative changes in your wealth are Gaussian. So the lognormal superficially resembles the fractal, in the sense that it may tolerate some large deviations, but it is dangerous because these rapidly taper off at the end. The introduction of the lognormal was a very bad compromise, but a way to conceal the flaws of the Gaussian.

Extinctions: Sterelny (2001). For extinctions from abrupt fractures, see Courtillot (1995) and Courtillot and Gaudemer (1996). Jumps: Eldredge and Gould.

FRACTALS, POWER LAWS, and SCALE-FREE DISTRIBUTIONS

Definition: Technically, $P_{>x} = K\, x^{-\alpha}$ where α is supposed to be the power-law exponent. It is said to be scale free, in the sense that it does not have a characteristic scale: relative deviation of $\frac{P_{>x}}{P_{>nx}}$ does not depend on x, but on n—for x "large enough." Now, in the other class of distribution, the one that I can intuitively describe as nonscalable, with the typical shape p(x) = Exp[-a x], the scale will be *a*.

Problem of "how large": Now the problem that is usually misunderstood. This scalability might stop somewhere, but I do not know where, so I might consider it infinite. The statements *very large and I don't know how large* and *infinitely large* are epistemologically substitutable. There might be a point at which the distributions flip. This will show once we look at them more graphically.

Log P>x = $-\alpha$ Log X $+C^t$ for a scalable. When we do a log-log plot (i.e., plot P>x and x on a logarithmic scale), as in Figures 15 and 16, we should see a straight line.

Fractals and power laws: Mandelbrot (1975, 1982). Schroeder (1991) is imperative. John Chipman's unpublished manuscript *The Paretian Heritage* (Chipman [2006]) is the best review piece I've seen. See also Mitzenmacher (2003).

"To come very near true theory and to grasp its precise application are two very different things as the history of science teaches us. Everything of importance has been said before by somebody who did not discover it." Whitehead (1925).

Fractals in poetry: For the quote on Dickinson, see Fulton (1998).

Lacunarity: Brockman (2005). In the arts, Mandelbrot (1982).

Fractals in medicine: "New Tool to Diagnose and Treat Breast Cancer," *Newswise,* July 18, 2006.

General reference books in statistical physics: The most complete (in relation to fat tails) is Sornette (2004). See also Voit (2001) or the far deeper Bouchaud and Potters (2002) for financial prices and econophysics. For "complexity" theory, technical books: Bocarra (2004), Strogatz (1994), the popular Ruelle (1991), and also Prigogine (1996).

Fitting processes: For the philosophy of the problem, Taleb and Pilpel (2004). See also Pisarenko and Sornette (2004), Sornette et al. (2004), and Sornette and Ide (2001).

Poisson jump: Sometimes people propose a Gaussian distribution with a small probability of a "Poisson" jump. This may be fine, but how do you know how large the jump is going to be? Past data might not tell you how large the jump is.

Small sample effect: Weron (2001). Officer (1972) is quite ignorant of the point.

Recursivity of statistics: Taleb and Pilpel (2004), Blyth et al. (2005).

Biology: Modern molecular biology pioneers Salvador Luria and Max Delbrück witnessed a clustering phenomenon with the occasional occurrence of extremely large mutants in a bacterial colony, larger than all other bacteria.

Thermodynamics: Entropy maximization without the constraints of a second moment

FIGURE 15: TYPICAL DISTRIBUTION WITH POWER-LAW TAILS (HERE A STUDENT T)

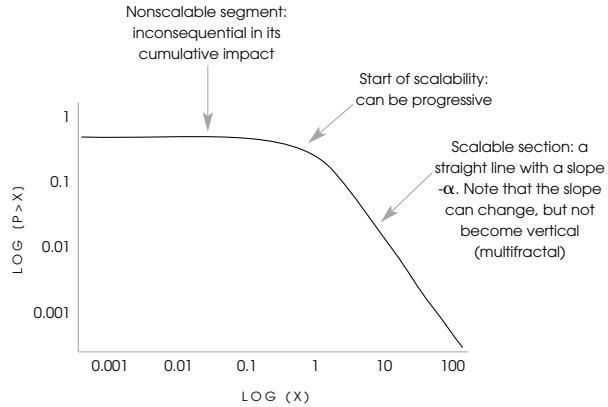
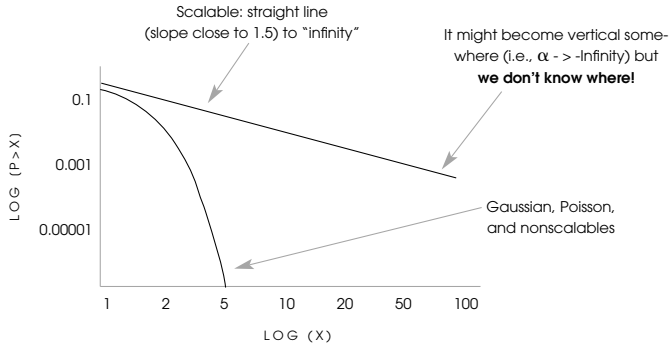


FIGURE 16



The two exhaustive domains of attraction: vertical or straight line with slopes either negative infinity or constant negative α. Note that since probabilities need to add up to 1 (even in France) there cannot be other alternatives to the two basins, which is why I narrow it down to these two exclusively.

My ideas are made very simple with this clean cut polarization—added to the problem of not knowing which basin we are in owing to the scarcity of data on the far right.

leads to a Levy-stable distribution—Mandelbrot's thesis of 1952 (see Mandelbrot [1997a]). Tsallis's more sophisticated view of entropy leads to a Student T.

Imitation chains and pathologies: An informational cascade is a process where a purely rational agent elects a particular choice ignoring his own private information (or judgment) to follow that of others. You run, I follow you, because you may be aware of a danger I may be missing. It is efficient to do what others do instead of having to reinvent the wheel every time. But this copying the behavior of others can lead to imitation chains. Soon everyone is running in the same direction, and it can be for spurious reasons. This behavior causes stock market bubbles and the formation of massive cultural fads. Bikhchandani et al. (1992). In psychology, see Hansen and Donoghue (1977). In biology/selection, Dugatkin (2001), Kirpatrick and Dugatkin (1994).

Self-organized criticality: Bak and Chen (1991), Bak (1996).

Economic variables: Bundt and Murphy (2006). Most economic variables seem to follow a "stable" distribution. They include foreign exchange, the GDP, the money supply, interest rates (long and short term), and industrial production.

Statisticians not accepting scalability: Flawed reasoning mistaking for sampling error in the tails for a boundedness: Perline (2005), for instance, does not understand the difference between absence of evidence and evidence of absence.

Time series and memory: You can have "fractal memory," i.e., the effect of past events on the present has an impact that has a "tail." It decays as power-law, not exponentially.

Marmott's work: Marmott (2004).

CHAPTER 18

Economists: Weintraub (2002), Szenberg (1992).

Portfolio theory and modern finance: Markowitz (1952, 1959), Huang and Litzenberger (1988) and Sharpe (1994, 1996). What is called the Sharpe ratio is meaningless outside of Mediocristan. The contents of Steve Ross's book (Ross [2004]) on "neoclassical finance" are completely canceled if you consider Extremistan in spite of the "elegant" mathematics and the beautiful top-down theories. "Anecdote" of Merton minor in Merton (1992).

Obsession with measurement: Crosby (1997) is often shown to me as convincing evidence that measuring was a great accomplishment not knowing that it applied to Mediocristan and Mediocristan only. Bernstein (1996) makes the same error.

Power laws in finance: Mandelbrot (1963), Gabaix et al. (2003), and Stanley et al. (2000). Kaizoji and Kaizoji (2004), Véhel and Walter (2002). Land prices: Kaizoji (2003). Magisterial: Bouchaud and Potters (2003).

Equity premium puzzle: If you accept fat tails, there is no equity premium puzzle. Benartzi and Thaler (1995) offer a psychological explanation, not realizing that variance is not the measure. So do many others.

Covered writes: a sucker's game as you cut your upside—conditional on the upside being breached, the stock should rally a lot more than intuitively accepted. For a representative mistake, see Board et al. (2000).

Nobel family: "Nobel Descendant Slams Economics Prize," *The Local,* September 28, 2005, Stockholm.

Double bubble: The problem of derivatives is that if the underlying security has mild fat tails and follows a mild power law (i.e., a tail exponent of three or higher), the derivative will produce far fatter tails (if the payoff is in squares, then the tail exponent of the derivatives portfolio will be half that of the primitive). This makes the Black-Scholes-Merton equation twice as unfit!

Poisson busting: The best way to figure out the problems of the Poisson as a substitute for a scalable is to calibrate a Poisson and compute the errors out of sample. The same applies to methods such as GARCH—they fare well in sample, but horribly, horribly

outside (even a trailing three-month past historical volatility or mean deviation will outperform a GARCH of higher orders).

Why the Nobel: Derman and Taleb (2005), Haug (2007).

Claude Bernard and experimental medicine: "*Empiricism pour le présent, avec direction a aspiration scientifique pour l'avenir.*" From Claude Bernard, *Principe de la médecine expérimentale*. See also Fagot-Largeault (2002) and Ruffié (1977). Modern evidence-based medicine: Ierodiakonou and Vandenbroucke (1993) and Vandenbroucke (1996) discuss a stochastic approach to medicine.

CHAPTER 19

Popper quote: From *Conjectures and Refutations*, pages 95–97.

The lottery paradox: This is one example of scholars not understanding the high-impact rare event. There is a well-known philosophical conundrum called the "lottery paradox," originally posed by the logician Henry Kyburg (see Rescher [2001] and Clark [2002]), which goes as follows: "I do not believe that any ticket will win the lottery, but I do believe that all tickets will win the lottery." To me (and a regular person) this statement does not seem to have anything strange in it. Yet for an academic philosopher trained in classical logic, this is a paradox. But it is only so if one tries to squeeze probability statements into commonly used logic that dates from Aristotle and is *all or nothing*. An *all or nothing* acceptance and rejection ("I believe" or "I do not believe") is inadequate with the highly improbable. We need shades of belief, degrees of faith you have in a statement other than 100% or 0%.

One final philosophical consideration. For my friend the options trader and Talmudic scholar Rabbi Tony Glickman: life is convex and to be seen as a series of derivatives. Simply put, when you cut the negative exposure, you limit your vulnerability to unknowledge, Taleb (2005).