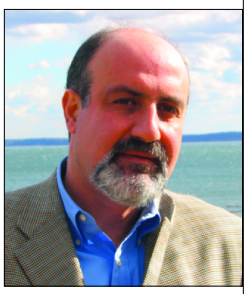
Interview: Nassim Nicholas Taleb

FIGURATIVE BLACK SWANS

Aaron Brown (AB): You seem to want to have it both ways with respect to Black Swans. You write, "I would recommend to someone to pick a profession that is not scalable," but also, "Remember that positive Black Swans have a necessary first step: you need to be exposed to them. … Seize any opportunity, or anything that looks like an opportunity." You also say to avoid situations with negative Black Swan potential.

That seems as if you want one pocket stuffed with lottery tickets and another stuffed with insurance policies, and you expect someone to pay you to stuff your pockets. Or, in Greek, you want that magical position that is long gamma and long theta. Are you telling people to risk blow-up when there are moderate-sized moves? Or are you just as irrational about risk as the rest of us?



Nassim Taleb

Nassim Nicholas Taleb (NNT): I am telling someone to seize opportunities because the asymmetry is rarely appreciated. Also, I am recommending the avoidance to big downside exposures, because these are not priceable. Many people are avoiding these risks not because they are courageous, but rather because they are both chicken and self-deluded.

Indeed I want to have it both ways: guaranteed floor and maximal upside — the barbell strategies I propose; the right mixture of greed and paranoia. I want a safety of a minimal income with the out-of-the-money dream. This is why the book advocates the stochasticity of trial-and-error as a device for research. In other words, you are not paying for that gamma — just look at the world today compared to what it was during the Pleistocene.

When it comes to professional choices, scalable activities are not utility efficient because it is hit or miss, misery or glory, and we might not enjoy glory as much as we hate misery. Even outside of that, it is not great economically because actors, entrepreneurs and writers make collectively less than publishers and venture capitalists — who besides being more successful are also more diversified. But people like to buy a dream: the pursuit of glory is a strong driver, as illustrated in my discussion of Giovanni Drogo's fate. It makes you avoid the complicated, hedonic satisficing of daily life.

So there is something existential about living in hope that I do not discount. It makes some people happy to be overconfident; there is something inherently human about shooting for glory. And there is something else: Black Swan hunting has an ethical connotation to me. It is honorable to avoid some classes of risks in spite of peer pressure.

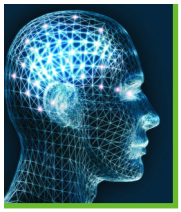
LITERAL BLACK SWANS

AB: For centuries "all swans are white" was an expression in Europe for something obviously true. In 1697, black swans were discovered in Australia. European philosophers switched to "all ravens are black," for which no clear-cut counterexamples have been discovered. Following the philosopher of science Karl Popper, you use this as an illustration of how extensive confirmation cannot prove a statement true.

But is this fair? Perhaps a typical European 400 years ago would say "all swans are white," but I'm pretty confident he wouldn't give you infinite odds if you offered to bet on it. I doubt you could have gotten even odds against the proposition that you could produce a black swan. Casual observation and word-of-mouth would convince people correctly that the swans they were likely to see would be overwhelmingly white, and perhaps exclusively so. But anyone with any money left understands the difference between this kind of reasoning and logical certainty.

There were some people for whom a black swan would be a Black Swan (an unexpected, life-changing event). It could make the reputation of a naturalist or the fortune of an exotic animal dealer. These people understood the limits of the evidence; they flocked to voyages of discovery and before that to remote places in Europe. They weren't looking specifically for black swans; the chances of finding one was too small. But they knew that if you looked in new places, you would find something new.

Granted, there were philosophers who claimed proofs that all swans were white. They did not immediately lose their reputations on the finding of a counterexample. As you point out in the book, experts can be consistently wrong without



much negative consequence. But in a larger view, medieval Aristotelian thought was overthrown by empirically based science, in part due to unexpected discoveries like black swans. Aren't people a bit smarter than you imply about black swans and Black Swans?

NNT: Primo, Karl Popper dealt with a logical problem, the error of stating no black swans. My Black Swan has one additional component: impact. Being wrong can be extremely costly, so my Black Swan starts where conventional philosophers stop. My idea of domains where this matters changes the whole story; in fact, the logical component for me is minimal. You make a mistake about a bird: so what! But here I am discussing some domains in which you can be a sucker.

Secundo, more technically, in logic it is a problem of probability. For me it is one of probability times the payoff (i.e., the moments of the distribution). This makes it easier, much easier, than the black swan problem. Make sure you look for these probability-impact pairs where you can be in trouble.

TEN DAYS IN 50 YEARS

AB: You make the point that the largest 10 daily moves in the S&P 500 over the last 50 years account for half the total return over the period. But that's taking a total, not an average. The longer the period it takes for the 10 most extreme log returns to add to an absolute value above ln(2), the less volatile the market. Looking at things another way, if the S&P 500 doesn't change for the next 50 years, you'll be able to say "10 days in a century" instead of "10 days in 50 years" — even though there was no market volatility at all in the second half of the period.

Mediocristan statisticians say enough data washes out the effect of individual observations when computing averages, but not when computing totals. Those ten days will continue to represent half the total return of the stock market forever into the future, because you're compounding rather than computing average returns. In average terms, missing those 10 days would raise your average log return from 7.0% to 8.4%, a significant change, but not half the return (and even smaller-looking if we add dividend return to both).

Isn't Extremistan the place where people compute totals instead of averages, rather than a place where underlying phenomena have fat-tails and non-linear effects?

NNT: Here I agree with you, but let's not nitpick: I wrote on the graph that it was pedagogical, not as compelling scientifically as other violations of normality. It is the weakest point in the book from a mathematical standpoint, as the 20 sigmas and the plethora of 10 sigmas in single stocks and interest rates (where I've seen 50 sigmas!) are far more convincing — but these examples are not noticed. In reality, it does illustrate skewness far more than fat tails. It is also the effect of a multiplicative process, which, as you noticed, compounds variations.

But I disagree with the idea that the effect is benign for multiplicative returns: in fact, the multiplication compounds the errors, and we must worry a lot more about misspecification with such a process than an arithmetic one. Many people such as the economist Robert Gibrat argued that log returns can cause inequalities such as the ones we see today — the 5-, 10-, and 20-sigma occurrences that are in log returns. Another few huge days and your average would be totally shellacked.

The way I view it, for the broad market, the power law scale exponent is often calculated as $a = 3$ for downside deviations and close to 5 for upside ones. These are not situations of hugely fat tails, but they are sufficient to make portfolio theory completely nonsense mathematically — along with dynamic option replication. Where things turn scary is with the nonlinear payoffs like those of derivatives as the "a" of these payoffs becomes very, very low and their effects turn explosive.

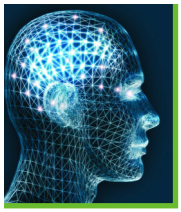
Finally, a little more technical point: my true problem is that the a can be as low as 1.8 for very, very large deviations, from something I call the Weron effect, which was rediscovered by my colleagues Mark Spitznagel and Pallop Angsupun at Universa (formerly of Empirica). If you generate by Monte Carlo an infinite variance process, a statistician is more likely to mistake it for a process with much thinner tails.

CONFIRMATION BIAS

AB: You discuss Peter Wason's famous 2-4-6 task. Subjects are given 2-4-6 as a triple of numbers that obeys some rule and asked to deduce that rule by suggesting other triples, to which the experimenter will say either "yes" or "no" according to whether they conform to the rule. Subjects seemed to guess initially the rule was consecutive even numbers and many neglected to test odd or nonsequential numbers. This 1960 experiment stimulated a large body of work.

You interpret this result as a confirmation bias; people give sequences that confirm their initial guess rather than sequences that test it. But why is this irrational? Suppose my first guess is 8/10/12, and the experimenter says, "yes." Now I can try to narrow the rule or broaden it. I could guess, for example, "-6/-4/-2," in case the rule is consecutive even positive numbers. If I get a "no," I have narrowed the rule. Or I could guess "1/3/5," in case the rule is ascending sequence with differences of 2. In that case, a "yes" broadens my rule.

Why is one preferable to the other? Suppose I find that



chewing on the bark of a certain tree reduces fever. As a medical investigator, I could try to broaden that rule by having people with fevers chew bark from other trees or people with other symptoms chew bark from the first tree. I learn from a positive result. Or I could try to narrow it by taking the bark from different places on the tree or at different times. Then, I learn from a negative result. Both investigations are valuable, yet you seem to label the first one "empirical medicine" and the second one "confirmation error."

Why is performing experiments in which you expect a positive result inherently inferior to performing experiments in which you expect a negative result?

NNT: I agree with you that there are situations in which it is efficient to confirm and generalize. Indeed, as always, in Mediocristan it is both efficient and a good approximation. My point again takes us to Mediocristan. In a fat-tailed environment, negative confirmation works much better than in thin-tailed ones. With high-impact events, also. You can confirm guilt more easily than innocence. In thin-tailed environments, you may have situations of equality.

Let me mention something else here: statistics is inherently a confirmatory exercise, and what I want to propose is a veto mechanism to lift confirmation in some domains. Again, the try-not-to-be-a-sucker theory.

SCALABILITY

AB: When giving examples of scalable phenomena, you seem to mix two different types. Things like wealth, earthquakes and stock market returns are clearly non-Gaussian, even in small samples. If you give me the wealth of 100 randomly selected individuals, I'm likely to notice that the data appear to have an exponential distribution, which will lead me to predict a significant probability of Bill-Gates wealth, even if no one in the sample comes close to that level. Even intraday stock market returns exhibit significant heteroskedasticity, which should lead me to adopt a GARCH or similar model that predicts significant probability of extreme movements.

Of course, in both cases, small data sets will have a lot of estimation error when extrapolating to extreme events, but the small sample data should alert me to the possibility of events many (small sample) standard deviations away from the (small sample) mean, which implies the possibility that the true population mean may be quite different from the small sample mean. People do get fooled by this all the time, but good statistical practice can avoid it.

Far more problematic are things that appear perfectly Gaussian in small samples, but have rare extreme jumps. For example, if you study estimated worldwide influenza mortality rates over the last 300 years leaving out 1732, 1781, 1802, 1830, 1847, 1857, 1918, 1957 and 1968, you could conclude you were looking at a Gaussian distribution. You could study 30 years of consecutive data without a clue that global influenza pandemics are possible and therefore understating the mean and wildly understating the probability of extreme events. No doubt there are a huge number of other things which are more than 10 times as rare, or for which we have less than a 10th as much data, that are entirely unsuspected and will appear as Black Swans.

Most of the examples in your book appear to be things of the first type, which only require improvements in statistical practice to address. But most of the discussion concerns things of the second type, for which statistics can offer no help. Are you mixing apples and oranges?

NNT: I disagree. Many of these are in fact scalable, but we don't know it because their sample is too small.

GARCH does not work out of sample. It is a good story, but I was unable to use it in predicting squared deviations or mean deviations. Heteroskedasticity — that is, stochastic volatility — is an epiphenomenon, not a phenomenon. It may come from the measurement of a process without tractable higher moments.

NARRATIVE FALLACY

AB: You accuse newspapers in particular of the narrative fallacy, assigning importance to events after the fact to make a coherent story, which is then "confirmed" by those same events. Conflicting evidence is deemed unimportant since, by definition, it did not contribute to the story.

Aren't you guilty of the same thing when you claim Black Swans dominate history and current events? Newspaper headlines are often surprising. As the saying goes, "Man Bites Dog" is a headline, "Dog Bites Man" isn't. Those surprising events were often unpredictable, sometimes even indefinable, before they happened. But the rest of the newspaper, the part that influences far more people far more of the time, seems to change slowly and predictably, either in cycles, in trends or not at all. The weather, fashion pages, comics, food section, want ads, graduation announcements, local celebrations and other matters contain few surprises and are not hard to predict. Headlines are caused by Black Swans, because headlines need Black Swans, and newspapers will create them if they don't occur naturally.

Similarly, if history is the story of kings and assassinations and revolutions and market crashes and inventions, then it is a flock of Black Swans. They fill the history books because they're exciting. But the day-to-day life of most people changes slowly and predictably. Fads come and go, and it

may be hard to predict the next one, but it's easy to predict there will be a next one. Hemlines go up and down, but people still wear clothes, and with remarkably conservative design, materials and functions. Technology changes, but people integrate it into their lives as convenient; only over long periods of time does it force change on individuals. Are Black Swans just the tip of the narrative fallacy?

NNT: Yes and no. My entire idea is that we overestimate some Black Swans, but not the idea of Black Swans. People make the mistake all the time of thinking that I am saying that all Black Swans are underestimated. I keep writing that we are shallowly suckers for the salient and people change my story to make me a predictor of Black Swans.

Myron Scholes, whom I enjoy putting in a state of rage, changed my story twice in order, attacking me. The first time he accused me of saying that fat tails come from psychology. The last time he used the counterargument that some people look at some extreme events, I guess to defend himself. That is not my point. My point, simply, is the abstract notion of the incomputability of extreme events. The larger the event, the less we can compute it.

As to the timescale, I wrote in Chapter 4 that some slowly-building-up historical events are Black Swans for a longer term predictor. It is a matter of timescales.

Also, I used a narrative to displace a narrative. *The Black Swan* is a good narrative — but it is an honest, non-distorting one.

BANKS

AB: Most people believe we need banks and that market processes alone are not enough to ensure they hold adequate capital. If you set the capital requirement too high, there aren't enough banks and they don't take enough risk, so the economy suffers. If you set the capital requirement too low, there are too many bank failures with the attendant disruption and expense. If you were forced to set the capital requirements for the world's banks, what would you say? The possibility of Black Swans makes holding capital irrelevant? Hold infinite amounts of capital? No capital? Hold 8% of notional amount of assets because more complicated analysis is worthless? Or simply, I don't know?

NNT: We do not need banks to gamble using public money and pay bonuses to employees when they are right. Hedge funds can screw their investors, but they are both mature individuals.

We should have let the banks go bankrupt. I am against worrying about the "health of the financial system" by putting Novacaine on it and increasing its long-term risks. Banks need to go bust once in a while so that we learn to rely on disintermediation with diversified ecology.

ICE CUBES AND PUDDLES

AB: You use the example of a melting ice cube to illustrate the problem with statistical models. If you see an ice cube, it's easy to predict the shape of the puddle that will result when it melts. But if you see a puddle, it's impossible to determine the shape of the ice that created it. Similarly, a statistical model will give precise predictions of possible outcomes; so if you have a model, you can put confidence intervals around events.

However, observing events does not help you build a model. There are an infinite number of models that could plausibly underlie historical events. Unless you have strong theoretical reasons for knowing the type of model, and there are a reasonably small number of parameters relative to the amount of data, modelling is as hopeless as trying to guess the shape of an ice cube from observing a puddle.

But isn't this the reason the field of non-parametric statistics was developed? It's possible to make some inferences and predictions from data with only weak, general model assumptions. Also, we have visualization methods, data analysis and robust estimators to give conclusions that are not sensitive to the underlying model. The heat equation that governs ice cube melting is the most famous example in physics of a problem that is easy to solve forward, but ill-posed backward. Most of the time, you can gain some knowledge about the past by examining the present. Isn't an ice cube an extreme example that overstates your case?

NNT: My point is far deeper, at the core of the knowledge in the tails, and cannot be remedied by non-parametric statistics. We do not observe probabilities; we estimate them from observations and samples.

The principal problem of probabilistic knowledge and discovery is as follows: without a strong *a priori* specification of the underlying reality of a process, epistemic confidence is inversely proportional to the consequences of the knowledge at hand. Large deviations, the one that can have the most meaningful effect on the total properties, are far more difficult to evaluate empirically than regular ones.

Non-parametric statistical methods will not do anything in the tails. Where it helps is in replacing Variance and Gaussian terms with robust ones like mean absolute deviation (MAD). I like working with higher moments like volatility of volatility using nested mean deviations. MAD works better than standard deviation, MAD of MAD works better than volatility of volatility.