Available online at www.sciencedirect.com
ScienceDirect
International Journal of Forecasting 25 (2009) 744–759
international journal of forecasting
www.elsevier.com/locate/ijforecast

Errors, robustness, and the fourth quadrant

Nassim Nicholas Taleb

New York University–Polytechnic Institute and Universa Investments, United States

Abstract

The paper presents evidence that econometric techniques based on variance – L^2 norm – are flawed and do not replicate. The result is un-computability of the role of tail events. The paper proposes a methodology to calibrate decisions to the degree (and computability) of forecast error. It classifies decision payoffs in two types: simple (true/false or binary) and complex (higher moments); and randomness into type-1 (thin tails) and type-2 (true fat tails), and shows the errors for the estimation of small probability payoffs for type 2 randomness. The fourth quadrant is where payoffs are complex with type-2 randomness. We propose solutions to mitigate the effect of the fourth quadrant, based on the nature of complex systems.

Keywords: Complexity; Decision theory; Fat tails; Risk management

1. Background and purpose

It appears scandalous that, of the hundreds of thousands of professionals involved, including prime public institutions such as the World Bank, the International Monetary Fund, different governmental agencies and central banks, private institutions such as banks, insurance companies, and large corporations, and, finally, academic departments, only a few individuals considered the possibility of the total collapse of the banking system that started in 2007 (and is still worsening at the time of writing), let alone the economic consequences of such breakdown. Not a single official forecast turned out to be close to the outcome experienced—even those issuing "warnings" did not come close to the true gravity of the situation. A few warnings about the risks, such as Taleb (2007a) or the works of the economist Nouriel Roubini,[1] went unheeded, often ridiculed.[2] Where did such sophistication go? In the face of miscalculations of such proportion, it would seem fitting to start an examination of the conventional forecasting methods for risky outcomes and assess their fragility—indeed, the size of the damage comes from confidence in forecasting and the mis-estimation of potential forecast errors for a certain classes of variables and a certain type of exposures. However, this was not

E-mail address: nnt@fooledbyrandomness.com.

[1] "Dr. Doom", *New York Times*, August 15, 2008.

[2] Note the irony that the ridicule of the warnings in Taleb (2007a) and other ideas came from the academic establishment, not from the popular press.

0169-2070/$ - see front matter
doi:10.1016/j.ijforecast.2009.05.027

the first time such events have happened—nor was it a "Black Swan" (when capitalized, an unpredictable outcome of high impact) to the observer who took a close look at the robustness and empirical validity of the methods used in economic forecasting and risk measurement.

This examination, while grounded in economic data, generalizes to all decision-making under uncertainty in which there is a potential miscalculation of the risk of a consequential rare event. The problem of concern is the rare event, and the exposure to it, of the kind that can fool a decision maker into taking a certain course of action based on a misunderstanding of the risks involved.

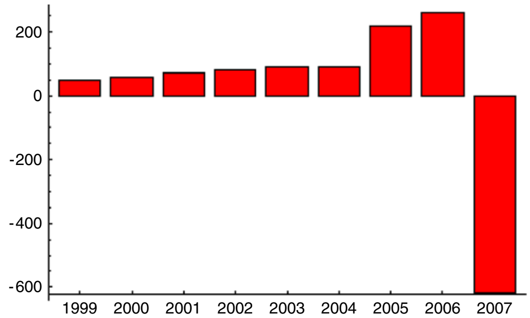

Fig. 1. Indy Mac's annual income (in millions) between 1998 and 2007. We can see fat tails at work. Tragic errors come from underestimating potential losses, with the best known cases being FNMA, Freddie Mac, Bear Stearns, Northern Rock, and Lehman Brothers, in addition to numerous hedge funds.

2. Introduction

Forecasting is a serious professional and scientific endeavor with a certain purpose, namely to provide predictions to be used in formulating decisions, and taking actions. The forecast translates into a decision, and, accordingly, the uncertainty attached to the forecast, i.e., the error, needs to be endogenous to the decision itself. This holds particularly true of risk decisions. In other words, the use of the forecast needs to be determined — or modified — based on the estimated accuracy of the forecast. This in turn creates an interdependency about what we should or should not forecast—as some forecasts can be harmful to decision makers.

Fig. 1 gives an example of harm coming from building risk management on the basis of extrapolative (usually highly technical) econometric methods, providing decision-makers with false confidence about the risks, and leaving society exposed to several trillions in losses that put capitalism on the verge of collapse.

A key word here, *fat tails*, implies the outsized role in the total statistical properties played by one single observation—such as one massive loss coming after years of stable profits or one massive variation unseen in past data.

- "Thin-tails" lead to ease in forecasting and tractability of the errors;
- "Thick-tails" imply more difficulties in getting a handle on the forecast errors and the fragility of the forecast.

Close to 1000 financial institutions have shut down in 2007 and 2008 from the underestimation of outsized market moves, with losses up to 3.6 trillion.[3] Had their managers been aware of the unreliability of the forecasting methods (which were already apparent in the data), they would have requested a different risk profile, with more robustness in risk management and smaller dependence on complex derivatives.

2.1. The smoking gun

We conducted a simple scientific examination of economic data, using a near-exhaustive set that includes 38 "tradable" variables[4] that allow for daily prices: major equity indices across the globe (US, Europe, Asia, Latin America), most metals (gold, silver), major interest rate securities, and main currencies — what we believe represents around 98% of tradable volume.

[3] Bloomberg, Feb 5, 2009.

[4] We selected a near-exhaustive set of economic data that includes "tradable" securities that allow for a future or a forward market: most equity indices across the globe, most metals, most interest rate securities, and most currencies. We collected all available traded futures data—what we believe represents around 98% of tradable volume. The reason we selected tradable data is because of the certainty of the practical aspect of a price on which one can transact: a nontradable currency price can lend itself to all manner of manipulation. More precisely we selected "continuously rolled" futures in which the returns from holding a security are built-in. For instance, analyses of Dow Jones that fail to account for dividend payments or analyses of currencies that do not include interest rates provide a bias in the measurement of the mean and higher moments.

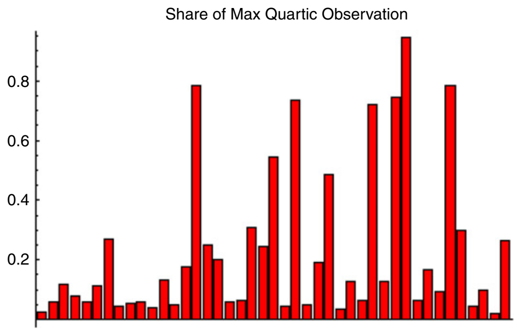

Fig. 2. The smoking gun: Maximum contribution to the fourth moment kurtosis coming from the largest observation in ∼10,000 (29–40 years of daily observations) for 43 economic variables. For the Gaussian the number is expected to be ∼0.006 for $n = 10{,}000$.

We analyzed the properties of the logarithmic returns $r_{t,\Delta t} = \text{Log}\left(\frac{X_t}{X_{t-\Delta t}}\right)$, where Δt can be 1 day, 10 days, or 66 days (non-overlapping intervals).[5]

A conventional test of nonnormality used in the literature is the excess kurtosis over the normal distribution. Thus, we measured the fourth noncentral moment $k(\Delta t) = \frac{\sum r_{t,\Delta t}^4}{n}$ of the distributions and focused on the stability of the measurements.

By examining Table 1 and Figs. 2 and 3, it appears that:

(1) Economic variables (currency rates, financial assets, interest rates, commodities) are patently fat tailed—with no known exception. The literature (Bundt & Murphy, 2006) shows that this also applies to data not considered here, owing to a lack of daily changes, such as GDP, or inflation.

(2) Conventional methods, not just those relying on a Gaussian distribution, but those based on least-square methods, or using variance as a measure of dispersion, are, according to the data, incapable of tracking the kind of "fat-tails" we see (more technically, in the L^2 norm, as will be discussed in Section 5). The reason is that most of the kurtosis is concentrated in a few observations, making it practically unknowable using conventional methods—see Fig. 2. Other tests in Section 5 (the conditional expectation above a threshold) show further instability. This incapacitates least-square methods, linear regression, and similar tools, including risk management methods such as "Gaussian Copulas" that rely on correlations or any form of the product of random variables.[6, 7, 8]

[5] By convention we use $t = 1$ as one business day.

[6] This should predict, for instance, the total failure in practice of the ARCH/GARCH methods (Engle, 1982), in spite of their successes in-sample, and in academic citations, as they are based on the behavior of squares.

[7] One counterintuitive result is that sophisticated operators do not seem to be aware of the norm they are using, thus mis-estimating volatility, see Goldstein and Taleb (2007).

[8] Practitioners have blamed the naive L^2 reliance on the risk management of credit risk for the blowup of banks in the crisis that started in 2007. See Felix Salmon's "Recipe For Disaster: The Formula That Killed Wall Street" in Wired. 02/23/2009.

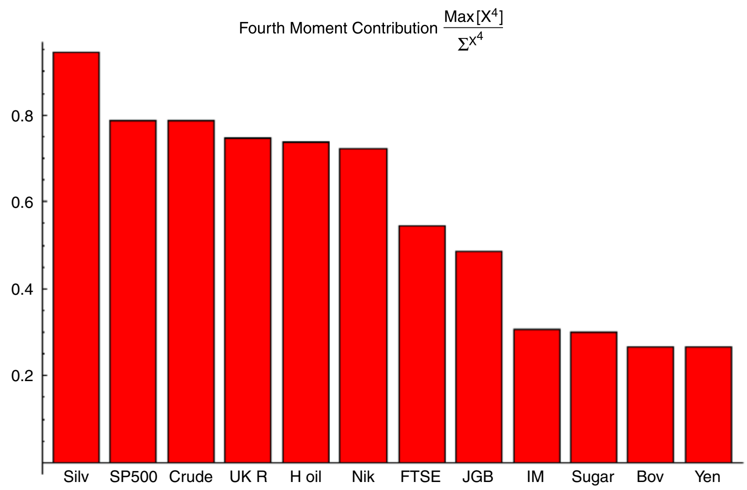

Fig. 3. A selection of the 12 most acute cases among the 43 economic variables.

Table 1
Fourth Noncentral Moment at daily, 10 day, and 66 day windows for the random variables.

	K (1)	K (10)	K (66)	Max quartic	Years
Australian Dollar/USD	6.3	3.8	2.9	0.12	22
Australia TB 10y	7.5	6.2	3.5	0.08	25
Australia TB 3y	7.5	5.4	4.2	0.06	21
BeanOil	5.5	7.0	4.9	0.11	47
Bonds 30Y	5.6	4.7	3.9	0.02	32
Bovespa	24.9	5.0	2.3	0.27	16
British Pound/USD	6.9	7.4	5.3	0.05	38
CAC40	6.5	4.7	3.6	0.05	20
Canadian Dollar	7.4	4.1	3.9	0.06	38
Cocoa NY	4.9	4.0	5.2	0.04	47
Coffee NY	10.7	5.2	5.3	0.13	37
Copper	6.4	5.5	4.5	0.05	48
Corn	9.4	8.0	5.0	0.18	49
Crude Oil	29.0	4.7	5.1	0.79	26
CT	7.8	4.8	3.7	0.25	48
DAX	8.0	6.5	3.7	0.2	18
Euro Bund	4.9	3.2	3.3	0.06	18
Euro Currency/DEM previously	5.5	3.8	2.8	0.06	38
Eurodollar Depo 1M	41.5	28.0	6.0	0.31	19
Eurodollar Depo 3M	21.1	8.1	7.0	0.25	28
FTSE	15.2	27.4	6.5	0.54	25
Gold	11.9	14.5	16.6	0.04	35
Heating Oil	20.0	4.1	4.4	0.74	31
Hogs	4.5	4.6	4.8	0.05	43
Jakarta Stock Index	40.5	6.2	4.2	0.19	16
Japanese Gov Bonds	17.2	16.9	4.3	0.48	24
Live Cattle	4.2	4.9	5.6	0.04	44
Nasdaq Index	11.4	9.3	5.0	0.13	21
Natural Gas	6.0	3.9	3.8	0.06	19
Nikkei	52.6	4.0	2.9	0.72	23
Notes 5Y	5.1	3.2	2.5	0.06	21
Russia RTSI	13.3	6.0	7.3	0.13	17
Short Sterling	851.8	93.0	3.0	0.75	17
Silver	160.3	22.6	10.2	0.94	46
Smallcap	6.1	5.7	6.8	0.06	17
SoyBeans	7.1	8.8	6.7	0.17	47
SoyMeal	8.9	9.8	8.5	0.09	48
Sp500	38.2	7.7	5.1	0.79	56
Sugar # 11	9.4	6.4	3.8	0.3	48
SwissFranc	5.1	3.8	2.6	0.05	38
TY10Y Notes	5.9	5.5	4.9	0.1	27
Wheat	5.6	6.0	6.9	0.02	49
Yen/USD	9.7	6.1	2.5	0.27	38

(3) There is no evidence of "convergence to normality" by aggregation, i.e., looking at the kurtosis of weekly or monthly changes. The "fatness" of the tails seems to be conserved under aggregation.

Clearly, had decision-makers been aware of such facts, and such unreliability of conventional methods in tracking large deviations, fewer losses would have been incurred, as they would have reduced exposures in some areas rather than rely on more "sophisticated" methods. The financial system has been fragile, as this simple test shows, with the evidence staring at us all along.

2.2. The problem of large deviations

2.2.1. The empirical problem of small probabilities

The central problem addressed in this paper is that small probabilities are difficult to estimate empirically (since the sample set for these is small), with a greater error rate than that for more frequent events. But since, in some domains, their effects can be consequential, the error concerning the contribution of small probabilities to the total moments of the distribution becomes disproportionately large. The problem has been dealt with by assuming a probability distribution and extrapolating into the tails—which brings model error into play. Yet, as we will discuss, model error plays a larger role with large deviations.

2.2.2. Links to decision theory

It is not necessary here to argue that a decision maker needs to use a full tableau of payoffs (rather than the simple one-dimensional average forecast) and that payoffs from decisions vary in their sensitivity to forecast errors. For instance, while it is acceptable to take a medicine that might be effective with a 5% error rate, but offers no side effects otherwise, it is foolish to play Russian roulette with the knowledge that one should win with a 5% error rate—indeed, standard theory of choice under uncertainty requires the use of full probability distributions, or at least a probability associated with every payoff. But so far this simple truism has not been integrated into the forecasting activity itself—as no classification has been made concerning the tractability and consequences of the errors. To put it simply, the mere separation between forecasting and decisions is lacking in both rigor and practicality, as it ruptures the link between forecast error and the quality of the decision.

The extensive literature on decision theory and choices under uncertainty so far has limited itself to (1) assuming *known* probability distributions (except for a few exceptions in which this type of uncertainty has been called “ambiguity”[9]), and (2) ignoring fat tails. This paper introduces a new structure of fat tails and classification of classes of randomness into the analysis, and focuses on the interrelation between errors and decisions. To establish a link between decision and quality of forecast, this analysis operates along two qualitative lines: qualitative differences between decisions along their vulnerability to error rates on one hand, and qualitative differences between two types of distributions of error rates. So there are two *distinct* types of decisions, and two *distinct* classes of randomness.

This classification allows us to isolate situations in which forecasting needs to be suspended—or a revision of the decision or exposure may be necessary. What we call the “fourth quadrant” is the area in which both the magnitude of forecast errors is large and the sensitivityt to these errors is consequential. What we recommend is either changes in the payoff itself (clipping exposure) or the shifting of exposures away from that part. For that we will provide precise rules.

The paper is organized as follows. First, we classify decisions according to targeted payoffs. Second, we discuss the problem of rare events, as these are the ones that are both consequential and hard to predict. Third, we present the classification of the two categories of probability distributions. Finally, we present the “fourth quadrant” and what we need to do to escape it, thus answering the call for how to handle “decision making under low predictability”.

3. The different types of decisions

The first type of decisions is simple, it aims at “binary” payoffs, i.e. you just care whether something is true or false. Very true or very false does not matter. Someone is either pregnant or not pregnant. A biological experiment in the laboratory or a bet about the outcome of an election belong to this category. A scientific statement is traditionally considered “true” or “false” with some confidence interval. More technically, they depend on the zeroth moment, namely just on the probability of events, and not their magnitude —for these one just cares about “raw” probability.[10]

[9] Ellsberg’s paradox, Ellsberg (1961); see also Gardenfors and Sahlin (1982) and Levi (1986).

[10] The difference can be best illustrated as follows. One of the most erroneous comparisons encountered in economics is the one between the “wine rating” and “credit rating” of complex securities. Errors in wine rating are hardly consequential for the buyer (the “payoff” is binary); errors in credit ratings have bankrupted banks, as these carry massive payoffs.

Clearly these are not very prevalent in life—they mostly exist in laboratory experiments and in research papers.

The second type of decisions depends on more complex payoffs. The decision maker does not just care about the frequency, but about the impact as well, or, even more complex, some function of the impact. So there is another layer of uncertainty of impact. These depend on higher moments of the distribution. When one invests one does not care about the frequency, how many times he makes or loses, he cares about the expectation: how many times money is made or lost *times* the amount made or lost. We will see that there are even more complex decisions.

More formally, where $p[x]$ is the probability distribution of the random variable x, and D the domain on which the distribution is defined, the payoff $\lambda(x)$ is defined by integrating on D as:

$$\lambda(x) = \int f(x)p(x)\mathrm{d}x.$$

Note that we can incorporate utility or nonlinearities of the payoff in the function $f(x)$. But let us ignore utility for the sake of simplification.

For a simple payoff, $f(x) = 1$. So $L(x)$ becomes the simple probability of exceeding x, since the final outcome is either 1 or 0 (or 1 and -1).

For more complicated payoffs, $f(x)$ can be complex. If the payoff depends on a simple expectation, i.e., $\lambda(x) = E[x]$, the corresponding function $f(x) = x$, and we need to ignore frequencies since it is the payoff that matters. One can be right 99% of the time, but this does not matter at all, since with some skewed distributions, the consequences of the expectation of the 1% error can be too large. Forecasting typically has $f(x) = x$, a linear function of x, while measures such as least squares depend on the higher moments $f(x) = x^2$.

Note that some financial products can even depend on the fourth moment (see Table 2).[11]

Next we turn to a discussion of the problem of rare events.

[11] More formally, a linear function with respect to the variable x has no second derivative; a convex function is one with a positive second derivative. By expanding the expectation of $f(x)$ we end up with $E[f(x)] = f(x)e[\Delta x] + 1/2 f''(x)E[\Delta x2] + \cdots$, and hence higher orders matter to the extent of the importance of higher derivatives.

4. The problem of rare events

The passage from theory to the real world presents two distinct difficulties: "inverse problems" and "pre-asymptotics".

4.1. Inverse problems

It is the greatest difficulty one can encounter in deriving properties. In real life we do not observe probability distributions, we just observe events. So we do not know the statistical properties — until, of course, after the fact — as we can see in Fig. 1. Given a set of observations, plenty of statistical distributions can correspond to the exact same realizations—each would extrapolate differently outside the set of events on which it was derived. The inverse problem is more acute when more theories, more distributions can fit a set of data—particularly in the presence of nonlinearities or nonparsimonious distributions.[12]

So this inverse problem is compounded of two problems:

+ *The small sample properties of rare events*, as these will be naturally rare in a past sample. This is also acute in the presence of nonlinearities, as the families of possible models/parametrization explode in numbers.
+ *The survivorship bias effect of high impact rare events*. For negatively skewed distributions (with a thicker left tail), the problem is worse. Clearly, catastrophic events will be necessarily absent from the data, since the survivorship of the variable itself will depend on such effect. Thus, left tailed distributions will (1) overestimate the mean; (2) underestimate the variance and the risk.

Fig. 4 shows how we normally lack data in the tails; Fig. 5 shows the empirical effect (see Fig. 6).

4.2. Pre-asymptotics

Theories can be extremely dangerous when they were derived in idealized situations, the asymptote, but are used outside the asymptote (at its limit, say infinity

[12] A Gaussian distribution is parsimonious (with only two parameters to fit). But the problem of adding layers of possible jumps, each with a different probabilities, opens up endless possibilities of combinations of parameters.

Table 2
Tableau of decisions.

Mo	M1	M2+
“True/False”	Expectations	
	LINEAR PAYOFF	NONLINEAR PAYOFF
$f(x) = 0$	$f(x) = 1$	$f(x)$ nonlinear($= x^2, x^3$, etc.)
Medicine (health not epidemics)	Finance: nonleveraged investment	Derivative payoffs
Psychology experiments	Insurance, measures of expected shortfall	Dynamically hedged portfolios
Bets (prediction markets)	General risk management	Leveraged portfolios (around the loss point)
Binary/Digital derivatives	Climate	Cubic payoffs (strips of out of the money options)
Life/Death	Economics (Policy)	Errors in analyses of volatility
	Security: Terrorism, Natural catastrophes	Calibration of nonlinear models
	Epidemics	Expectation weighted by nonlinear utility
	Casinos	Kurtosis-based positioning (“volatility trading”)

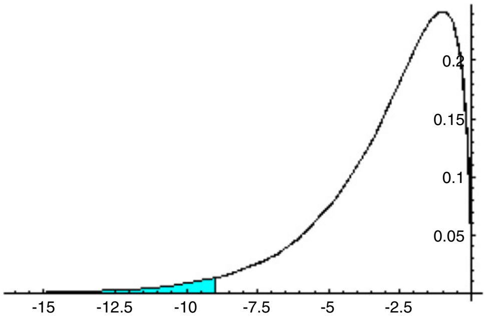

Fig. 4. The confirmation bias at work. The shaded area shows what tend to be missing from the observations. For negatively-skewed, fat-tailed distributions, we do not see much of negative outcomes for surviving entities AND we have a small sample in the left tail. This illustrates why we tend to see a better past for a certain class of time series than is warranted.

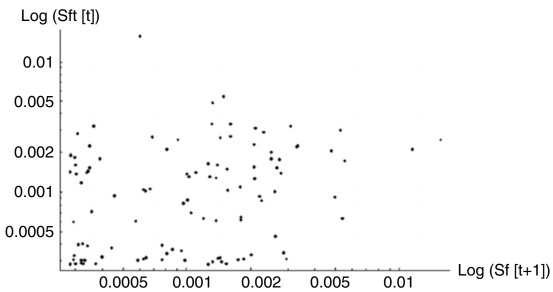

Fig. 5. Outliers don’t predict outliers. The plot shows (on a logarithmic scale) a shortfall in one given year against the shortfall the following one, repeated throughout for the 43 variables. A shortfall here is defined as the sum of deviations in excess of 7%. Past large deviations do not appear to predict future large deviations, at different lags.

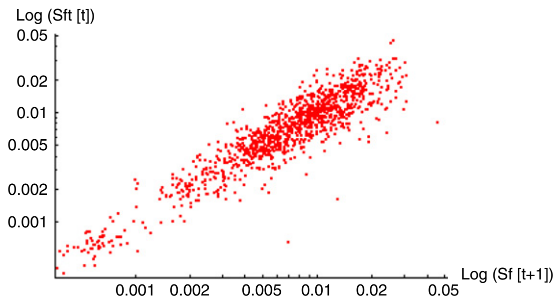

Fig. 6. Regular events predict regular events. This plot shows, by comparison with Fig. 5, how, for the same variables, the mean deviation in one period predicts the one in the subsequent period.

or the infinitesimal). Some asymptotic properties do work well pre-asymptotically (as we’ll see, with type-1 distributions), which is why casinos do well, but others do not, particularly when it comes to the class of fat-tailed distributions.

Most statistical education is based on these asymptotic, laboratory-style Platonic properties—yet we take economic decisions in the real world that very rarely resembles the asymptote. Most of what students of statistics do is assume a structure, typically with a known probability. Yet the problem we have is not so much making computations once you know the probabilities as finding the true distribution.

5. The two probabilistic structures

There are two classes of probability domains — very distinct qualitatively and quantitatively — according to precise mathematical properties. The first, Type-1, we call “benign” thin-tailed nonscalable, the second, Type 2, "wild" thick tailed scalable, or fractal (the attribution "wild" comes from the classification of Mandelbrot, 1963, 2001).

Taleb (2009) makes the distinction along the lines of convergence to the Central Limit Theorem. Type-1 converges in an acceptable form, while Type-2 either does not converge (infinite variance), or converges only in a remote asymptote and needs to be treated pre-asymptotically. Taleb (2009) also shows that one of the mistakes in the economics literature that "fattens the tails", with two main classes of nonparsimonious models and processes (the jump-diffusion processes of Merton, 1976,[13] or stochastic volatility models such as Engels' ARCH[14]) is to believe that the second type of distribution is amenable to analyses like the first—except with fatter tails. In reality, a fact commonly encountered by practitioners is that fat-tailed distributions are very unwieldy—as we can see in Fig. 2. Furthermore, we often face a problem of mistaking one for the other: a process that is extremely well behaved, but, on occasions, delivers a very large deviation, can easily be mistaken for a thin-tailed one—a problem known as the "problem of confirmation" (Taleb, 2007a,b). So we need to be suspicious of the mistake of taking Type-2 for Type-1, as it is more severe (and more readily made) than the one in the other direction.[15]

As we saw from the data presented, this classification of "fat tails" does not just mean having a fourth moment worse than the Gaussian. The Poisson distribution, or a mixed distribution with a known Poisson jump, would have tails thicker than the Gaussian; but this mild form of fat tails can be dealt with rather easily—the distribution has all its moments finite. The problem comes from the structure of the decline in probabilities for larger deviations and the ease with which the tools at our disposal can be tripped into producing erroneous results from observations of data in a finite sample and jumping to wrong decisions.

5.1. The scalable property of type-2 distributions

Take a random variable x. With scalable distributions, asymptotically, for x large enough (i.e. "in the tails"), $\frac{P[X>nx]}{P[X>x]}$ depends on n, not on x (the same property can hold for $P[X < nx]$ for negative values). This induces statistical self-similarities. Note that owing to the finiteness of the realizations of random variables, and the lack of samples in the tails, we might not be able to observe such a property, yet not be able to rule out.

For economic variables, there is no fundamental reason for the ratio of "exceedances" (i.e., the cumulative probability of exceeding a certain threshold) to decline, as both the numerator and the denominators are multiplied by 2.

This self-similarity at all scales generates power-law, or Paretian, tails, i.e., above a crossover point, $P[X > x] = Kx^{-\alpha}$.[16, 17]

Let us now draw the implications of type-2 distributions.

5.1.1. Finiteness of moments and higher order effects

For thick tailed distributions, moments higher than α are not "finite", i.e., they cannot be computed. They can certainly be measured in finite samples—thus giving the illusion of finiteness. But they typically show a great degree of instability. For instance, a distribution with an infinite variance will always provide, in a sample, the illusion of finiteness of variance.

In other words, while errors converge for type-1 distributions, the expectations of higher orders of x, say of order n, such as $1/n!E[x^n]$, where x is the error, do not decline; in fact, they become explosive (see Fig. 7).

13 See the general decomposition into diffusion and jump (non-scalable) in Duffie, Pan, and Singleton (2000) and Merton (1976); and the discussion in Baz and Chacko (2004) and Haug (2007).

14 Engle (1982).

15 Makridakis et al. (1993) and Makridakis and Hibon (2000) present evidence that more complicated methods of forecasting do not deliver superior results to simple ones (already bad). The obvious reason is that the errors in calibration swell with the complexity of the model.

16 Scalable discussions: introduced by Mandelbrot (1963), Mandelbrot (1997) and Mandelbrot and Taleb (in press).

17 Complexity and power laws: Amaral et al. (1997), Sornette (2004), and Stanley, Amaral, Gopikrishnan, and Plerou (2000); for scalability in different aspects of financial data, Gabaix, Gopikrishnan, Plerou, and Stanley (2003a,b), Gabaix, Ramalho, and Reuter (2003c), Gopikrishnan, Meyer, Amaral, and Stanley (1998), Gopikrishnan, Plerou, Amaral, Meyer, and Stanley (1999), and Gopikrishnan, Plerou, Gabaix, and Stanley (2000). For the statistical mechanics of scale-free networks see Albert, Jeong, and Barabási (2000), Albert and Barabasi (2002) and Barabási and Albert (1999). The "sandpile effect" (i.e., avalanches and cascades) is discussed by Bak (1996) and Bak, Tang, and Wiesenfeld (1987, 1988), as power laws arise from conditions of self-organized criticality.

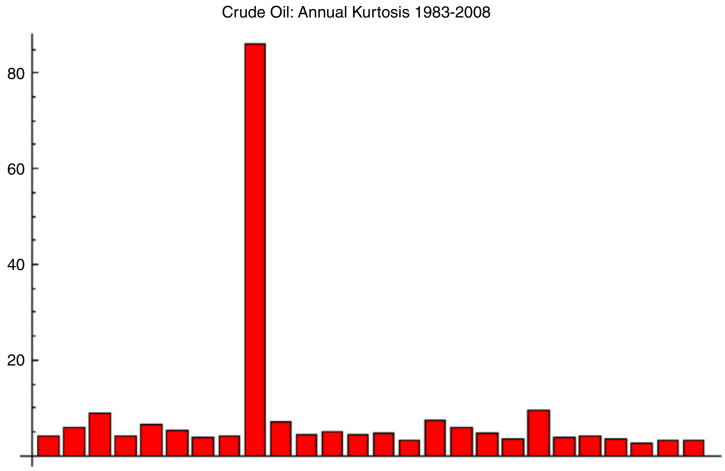

Fig. 7. Kurtosis over time: example of an "infinite moment". The graph shows the fourth moment for crude oil in annual nonoverlapping observations between 1982 and 2008. The instability shows in the dependence of the measurement on the observation window.

5.1.2. "Atypicality" of moves

For thin tailed domains, the conditional expectation of a random variable X, conditional on its exceeding a number K, converges to K for larger values of K.

$$\lim_{K \to \infty} E[X|_{X>K}] = K.$$

For instance, the conditional expectation for a Gaussian variable (assuming a mean of 0) conditional on the variable exceeding 0 is approximately 0.8 standard deviations. But with K equals 6 standard deviations, the conditional expectation converges to 6 standard deviations. The same applies to all of the random variables that do not have a Paretian tail. This induces some "typicality" of large moves.

For fat tailed variables, such a limit does not seem to hold:

$$\lim_{K \to \infty} E[X|_{X>K}] = Kc,$$

where c is a constant. For instance, the conditional expectation of a market move, given that it is in excess of 3 mean deviations, will be around 5 mean deviations. The expectation of a move conditional on it being higher than 10 mean deviations will be around 18. This property is quite crucial.

The atypicality of moves has the following significance.

– One may correctly predict a given event, say, a war, a market crash, or a credit crisis. But the amplitude of the damage will be unpredicted. The open-endedness of the outcomes can cause a severe miscalculation of the expected payoff function. For instance, the investment bank Morgan Stanley predicted a credit crisis but was severely hurt (and needed to be rescued) because it did not anticipate the extent of the damage.
– Methods like Value-at-Risk[18] that may correctly compute, say, a 99% probability of not losing no more than a given sum, called "value-at-risk", will nevertheless miscompute the conditional expectation should such a threshold be exceeded. For instance, one has 99% probability of not exceeding a $1 million loss, but should such a loss occur, it can be $10 million or $100 million.

This lack of typicality is of some significance. Stress testing and scenario generation are based on assuming a "crisis" scenario and checking robustness to it. Unfortunately such luxury is not available for fat tails, as "crises" do not have a typical magnitude.

Tables 3 and 4 show the evidence of a lack of convergence to thin tails, and hence a lack of "typicality" of the moves. We stopped for segments for which the number of observations becomes small, since a lack of observations in the tails can provide the illusion of "thin" tails.

[18] For the definition of Value at Risk see, Jorion (2001); for a critique, see Joe Nocera, "Risk Mismanagement: What led to the Financial Meltdown", *New York Time Magazine*, Jan 2, 2009.

Table 3
Conditional expectation for moves $> K$, 43 economic variables.

K, Mean deviations	Mean move (in MAD) in excess of K	n
1	2.01443	65,958
2	3.0814	23,450
3	4.19842	8,355
4	5.33587	3,202
5	6.52524	1,360
6	7.74405	660
7	9.10917	340
8	10.3649	192
9	11.6737	120
10	13.8726	84
11	15.3832	65
12	19.3987	47
13	21.0189	36
14	21.7426	29
15	24.1414	21
16	25.1188	18
17	27.8408	13
18	31.2309	11
19	35.6161	7
20	35.9036	6

Table 4
Conditional expectation for moves $< K$, 43 economic variables.

K, Mean deviations	Average move (in MAD) below K	n
−1	−2.06689	62,803
−2	−3.13423	23,258
−3	−4.24303	8,676
−4	−5.40792	3,346
−5	−6.66288	1,415
−6	−7.95766	689
−7	−9.43672	392
−8	−11.0048	226
−9	−13.158	133
−10	−14.6851	95
−11	−17.02	66
−12	−19.5828	46
−13	−21.353	38
−14	−25.0956	27
−15	−25.7004	22
−16	−27.5269	20
−17	−33.6529	16
−18	−35.0807	14
−19	−35.5523	13
−20	−38.7657	11

5.1.3. Preasymptotics

Even if we eventually converge to a probability distribution of the kind well known and tractable, it is central that the time to convergence plays a large role.

For instance, much of the literature invokes the Central Limit Theorem to assume that fat-tailed distributions with a finite variance converge to a Gaussian under summation. If daily errors are fat-tailed, cumulative monthly errors will become Gaussian. In practice, this does not appear to hold. The data, as we saw earlier, show that economic variables do not remotely converge to the Gaussian under aggregation.

Furthermore, finiteness of variance is a necessary but highly insufficient condition. Bouchaud and Potters (2003) showed that the tails remain heavy while the body of the distribution becomes Gaussian (see Fig. 8).

5.1.4. Metrics

Much of time series work seems to be based on metrics which are in the square domain, and hence patently intractable. Define the norm L^p:

$$\left(\frac{1}{n}\sum |x|^p\right)^{\frac{1}{p}};$$

it will increase along with p. The numbers can become explosive, with rare events taking a disproportionately larger share of the metric at higher orders of p. Thus the variance/standard deviation ($p = 2$), as a measure of dispersion, will be far more unstable than mean deviation ($p = 1$). The ratio of mean-deviation to variance (Taleb, 2009) is highly unstable for economic variables. Thus, modelizations based on variance become incapacitated. More practically, this means that for distributions with a finite mean (tail exponent greater than 1), the mean deviation is more "robust".[19]

[19] A note on the weaknesses of nonparametric statistics: the mean deviation is often used as a robust, nonparametric or distribution-free statistic. It does work better than the variance, as we saw, but does not contain information on rare events, by the argument seen before. Likewise, nonparametric statistical methods (relying on the empirical frequency) will be extremely fragile to the "black swan problem", since the absence of large deviations in the past leave us in a near-total opacity about their occurrence in the future—as we saw in Fig. 4, these are confirmatory. In other words, nonparametric statistics that consist of fitting a kernel to empirical frequencies, assume, even more than other methods, that a large deviation will have a predecessor.

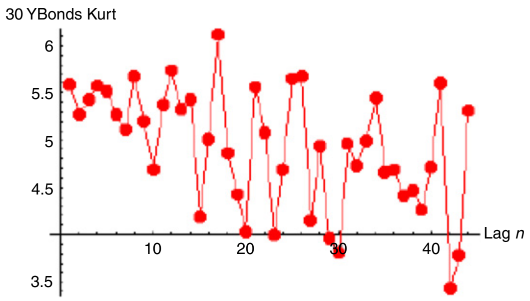

Fig. 8. Behavior of kurtosis under aggregation: we lengthen the window of changes from 1 day to 50 days. Even for variables with an infinite fourth moment, the kurtosis tends to drop under aggregation in small samples, then rise abruptly after a large observation.

5.1.5. Incidence of rare events

One common error is to believe that thickening the tails leads to an *increase* of the probability of rare events. In fact, it usually leads to a decrease of the incidence of such events, but the magnitude of the event, should it happen, will be much larger.

Take, for instance, a normally distributed random variable. The probability of exceeding 1 standard deviation is about 16%. Observed returns in the markets, with a higher kurtosis, present a lower probability of exceeding the same threshold, around 7%–10%, but the depth of the excursions is greater.

5.1.6. Calibration errors and fat tails

One does not need to accept power laws to use them. A convincing argument is that if we don't know what a "typical" event is, fractal power laws are the most effective way to *discuss* the extremes mathematically. It does not mean that the real world generator is actually a power law—it means that we don't understand the structure of the external events it delivers and need a tool of analysis. Also, fractals simplify the mathematical discussions because all you need to do is to perturbate one parameter, here the α, and it increases or decreases the role of the rare event in the total properties.

Say, for instance, that, in an analysis, you move α from 2.3 to 2 for data in the publishing business; the sales of books in excess of 1 million copies would triple! This method is akin to generating combinations of scenarios with series of probabilities and series of payoffs, fattening the tail at each time.

The following argument will help illustrate the general problem with forecasting under fat tails. Now the problem: *Parametrizing a power law lends itself to extremely large estimation errors* (since heavy tails have inverse problems). Small changes in the α main parameter used by power laws lead to extremely large effects in the tails.

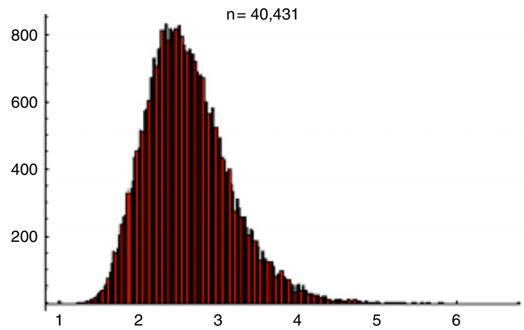

Fig. 9. Estimation error from 40 thousand economic variables.

And we don't observe the α—an uncertainty that comes from the measurement error. Fig. 9 shows more than 40 thousand computations of the tail exponent α from different samples of different economic variables (data for which it is impossible to refute fractal power laws). We clearly have problems figuring out what the α is: our results are marred by errors. The mean absolute error in the measurement of the tail exponent is in excess of 1 (i.e. between $\alpha = 2$ and $\alpha = 3$). Numerous papers in econophysics found an "average" alpha between 2 and 3—but if you process the >20 million pieces of data analyzed in the literature, you find that the variations between single variables are extremely significant.[20]

Now this mean error has massive consequences. Fig. 10 shows the effect: the expected value of your losses in excess of a certain amount (called the "shortfall") is multiplied by >10 from a small change in the α that is less than its mean error.[21]

[20] One aspect of this inverse problem is even pervasive in Monte Carlo experiments (much better behaved than the real world), see Weron (2001).

[21] Note that the literature on extreme value theory (Embrechts, Klüppelberg, & Mikosch, 1997) does not solve much of the problem, as the calibration errors stay the same. The argument about calibration we saw earlier makes the values depend on the unknowable tail exponent. This calibration problem explains how Extreme Value Theory works better on computers than in the real world (and has failed completely in the economic crisis of 2008–2009).

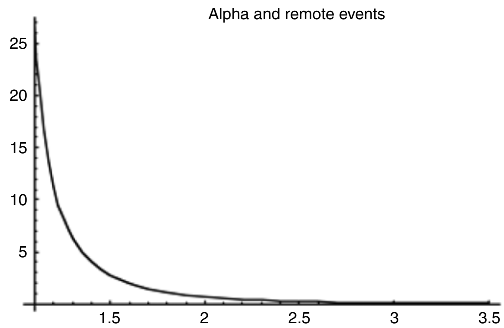

Fig. 10. The value of the expected shortfall (expected losses in excess of a certain threshold) in response to changes in the tail exponent α. We can see it explode by an order of magnitude.

6. The map

First quadrant: Simple binary decisions, under type-1 distributions: forecasting is safe. These situations are, unfortunately, more common in laboratories and games than in real life. We rarely observe these in payoffs in economic decision making. Examples: some medical decisions, casino bets, prediction markets.

Second quadrant: Complex decisions under type-1 distributions: Statistical methods may work satisfactorily, though there are some risks. True, thin-tails may not be a panacea, owing to preasymptotics, lack of independence, and model error. There are clearly problems there, but these have been addressed extensively in the literature (see Freedman, 2007).

Third quadrant: Simple decisions, under type-2 distributions: there is little harm in being wrong—the tails do not impact the payoffs.

Fourth quadrant: Complex decisions under type-2 distributions: this is where the problem resides. We need to avoid the prediction of remote payoffs—though not necessarily ordinary ones. Payoffs from remote parts of the distribution are more difficult to predict than closer parts.

A general principle is that, while in the first three quadrants you can use *the best* model you can find, this is dangerous in the fourth quadrant: no model should be better than just any model. So the idea is to exit the fourth quadrant.

The recommendation is to move into the third quadrant—it is not possible to change the distribution; but it is possible to change the payoff, as will be discussed in Section 7 (see Table 5).

The subtlety is that, while we have a poor idea about the expectation in the 4th quadrant, exposures to rare events are not symmetric.

7. Decision-making and forecasting in the fourth quadrant

7.1. Solutions by changing the payoff

Finally, the main idea proposed in this paper is to endogenize decisions, i.e., escape the 4th quadrant whenever possible by changing the payoff in reaction to the high degree of unpredictability and the harm it causes. How?

Just consider that the property of "atypicality" of the moves can be compensated by truncating the payoffs, thus creating an organic "worst case" scenario that is resistant to forecast errors. Recall that a binary payoff is insensitive to fat tails precisely because above a certain level, the domain of integration, changes in probabilities do not impact the payoff. So making the payoff no longer open-ended mitigates the problems, thus making it more tractable mathematically.

A way to express it using moments: all moments of the distribution become finite in the absence of open-ended payoffs, by putting a floor L below which $f(x) = 0$, as well a ceiling H. Just consider that if you are integrating payoffs in a finite, rather than an open-ended domain, i.e. between L and H, respectively, *the tails of the distributions outside that domain no longer matter*. Thus the domain of integration becomes the domain of payoff.

$$\lambda(x) = \int_L^H f(x)\, p(x)\mathrm{d}x.$$

With an investment portfolio, for instance, it is possible to "put a floor" on the payoff using insurance, or, even better, by changing the allocation. Insurance products are tailored with a maximum payoff; catastrophe insurance products are also set with a "cap", though the cap might be high enough to allow for a dependence on the error of the distribution.[22]

[22] Insurance companies might cap the payoff of a single claim, but a collection of capped claims might represent some problems, as the maximum loss becomes so large as to be almost undistinguishable from that with an uncapped payoff.

Table 5
The four quadrants.

	Simple payoffs	Complex payoffs
Distribution 1 ("thin tailed")	First quadrant: Extremely safe	Second quadrant: Safe
Distribution 2 (no or unknown characteristic scale)	Third quadrant: Safe	Fourth quadrant: Dangers[a]

[a] The dangers are limited to exposures in the negative domain (i.e., adverse payoffs). Some exposures, we will see, can only be "positive".

7.1.1. *The effect of skewness*

We omitted earlier to discuss asymmetry in either the payoff or the distribution. Clearly, the fourth quadrant can present either left or right skewness. If we suspect right-skewness, the true mean is more likely to be underestimated by the measurement of past realizations, and the total potential is likewise poorly gauged. A biotech company (usually) faces positive uncertainty, a bank faces almost exclusively negative shocks.

More significantly, by raising the L (the lower bound), one can easily produce positive skewness, with a set floor for potential adverse outcomes and open upside. For instance, what Taleb (2007a) calls a "barbell" investment strategy consists of allocating a high portion of a portfolio to T-Bills (or equivalent), say α, with $0 < \alpha < 1$, and a small portion $(1 - \alpha)$ to high-variance securities. While the total portfolio has medium variance, $L = (1 - \alpha)$ times the face value invested, another portfolio of the same variance might lose 100%.

7.1.2. *Convex and concave to error*

If a source of uncertainty can offer more benefits than a potential harm, then there may be gains from it—which we label "convex" or "concave".

More generally, we can be concave to model error if the payoff from the error (obtained by changing the tails of the distribution) has a negative second derivative with respect to the change in the tails, or is negatively skewed (like the payoff of a short option). It will be convex if the payoff is positively skewed (like the payoff of a long option).

7.1.3. *The effect of leverage in operations and investment*

Leveraging in finance has the effect of increasing concavity to model error. As we will see, it is exactly the opposite of redundancy—it causes payoffs to increase, but at the costs of an absorbing barrier should there be an extreme event that exceeds the allowance made in the risk measurement. Redundancy, on the other hand, is the equivalent of de-leveraging, i.e. by having more idle "inefficient" capital on the side. But a a second look at such funds can reveal that there may be a direct expected value from being able to benefit from opportunities in the event of asset deflation, and hence "idle" capital needs to be analyzed as an option.

7.2. *Solutions by mitigating forecasting errors*

7.2.1. *Optimization vs. redundancy*

The optimization paradigm of the economics literature meets some problems in the fourth quadrant: what if we have a consequential forecasting error? Aside from the issue that the economic agent is optimizing on the future states of the world, with a given probability distribution, nowhere[23] have the equations taken into account the possibility of a large deviation that would allow *not* optimizing consumption and having idle capital. Also, the psychological literature on well-being (Kahneman, 1999) shows an extremely concave utility function of income—if one spends such income. But if one hides it under the mattress, one will be less vulnerable to an extreme event. So there is an enhanced survival probability for those who have additional margin.

While economics have been mired in conventional linear analysis, stochastic optimization with Bellman-style equations that fall into the category Type-1, a different point of view is provided by complex systems analysis. One of the central attributes of complex systems is redundancy (May, Levin, & Sugihara, 2008).

[23] See Merton (1992) for a discussion of the general consumption Capital Asset Pricing Market.

Biological systems — those that have survived millions of years — include a large share of redundancies.[24, 25] Just consider the number of double organs (lungs, kidneys, ears). This may suggest an option-theoretic analysis: redundancy is like an option. One certainly pays for it, but it may be necessary for survival. And while redundancy means similar functions used by identical organs or resources, biological systems have, in addition, recourse to "degeneracy", the possibility of one organ to perform more than one function, which is the analog of redundancy at a functional level (Edelman & Gally, 2001).

When institutions such as banks optimize, they often do not realize that a simple model error can blow through their capital (as it just did) (see Fig. 11).

Examples: In one day in August 2007, Goldman Sachs experienced 24 times the average daily transaction volume[26]—would 29 times have blown up the clearing system? Another severe instance of an extreme "spike" lies in an event of September 18, 2008, in the aftermath of the Lehman Bothers Bankruptcy. According to congress documents, only made public in February 2009:

> On Thursday (Sept 18), at 11 am the Federal Reserve noticed a tremendous draw-down of money market accounts in the US, to the tune of $550 billion[27] was being drawn out in the matter of an hour or two.
>
> If they had not done that [add liquidity], their estimation is that by 2 pm that afternoon, $5.5 trillion would have been drawn out of the money market system of the U.S., which would have collapsed the entire economy of the U.S., and within 24 h the world economy would have collapsed. It would have been the end of our economic system and our political system as we know it.[28]

For naive economics, the best way to effectively reduce costs is to minimize redundancy, and hence avoiding the option premium of insurance. Indeed, some systems tend to optimize and therefore become more fragile. Albert and Barabasi (2002) and Barabási and Albert (1999) warned (ahead of the North Eastern power outage of August 2003) how electricity grids, for example, optimize to the point of not coping with unexpected surges—which predicted the possibility of a blackout of the magnitude of the one that took place in the North Eastern U.S. in August 2003. We cannot discuss "flat earth" globalization without realizing that it is overoptimized to the point of maximal vulnerability.

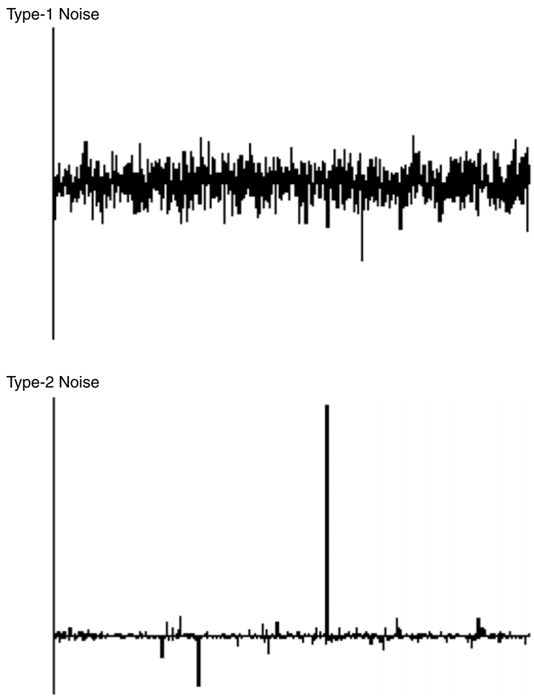

Fig. 11. Comparison between Gaussian-style noise and Type-2 noise with extreme spikes—which necessitates more redundancy (or insurance) than normally required. Policymakers and forecasters were not aware that complex systems tend to produce the second type of noise.

7.2.2. *Time and sample size*

It takes much, much longer for a fat-tailed time series to reveal its properties—in fact, many can, in short episodes, masquerade as thin-tailed. At the worst, we don't know how long it would take to know. But we can have a pretty clear idea whether *organically*, because of the nature of the payoff, the "Black Swan" can hit on the left (losses) or on the

[24] May et al. (2008).

[25] For the scalability of biological systems, see Burlando (1993), Enquist and Niklas (2001), Harte, Kinzig, and Green (1999), Ritchie and Olff (1999) and Solé, Manrubia, Benton, Kauffman, and Bak (1999).

[26] Personal communication, Pentagon Highland Forum, April meeting, 2008.

[27] Even if the number, as is possible, is off by one order of magnitude, the consequences remain extremely severe.

[28] http://www.liveleak.com/view?i=ca2_1234032281.

right (profits). This point can be used in climatic analysis. Things that have worked for a long time are preferable—they are more likely to have reached their ergodic states.

Likewise, portfolio diversification needs to be larger, much larger than anticipated. A mean variance Markowitz-style portfolio construction fails in the real world on several accounts. Taleb (2009) shows that, even if we assume finite variance, but fat tails and an unknown variance, the process of discovery of the variance itself makes portfolio theory totally unusable. DeMiguel, Garlappi, and Uppal (2007) show that a naive $1/n$ allocation outperforms out-of-sample any form of "optimal" portfolio—compatible with the notion that fat tails (and unknown future properties from past samples) require much broader diversification than is required by modern portfolio theory.

7.2.3. The problem of moral hazard

It is optimal (both economically and psychologically) to make a series of annual bonuses betting on hidden risks in the fourth quadrant, then "blow up" (Taleb, 2004). The problem is that bonus payments are made with a higher frequency (i.e. annually) than is warranted from the statistical properties (when it takes longer to capture the statistical properties).

7.2.4. Metrics

Conventional metrics based on type 1 randomness fail to produce reliable results—while the economics literature is grounded in them. Concepts like "standard deviation" are not stable and do not measure anything in the fourth quadrant. This is also true for "linear regression" (the errors are in the fourth quadrant), "Sharpe ratio", the Markowitz optimal portfolio,[29] ANOVA, Least squares, etc. "Variance" and "standard deviation" are terms invented years ago when we had no computers. Note that from the data shown and the instability of the kurtosis, no sample will ever deliver the true variance in a reasonable time. However, note that truncating payoffs blunts the effects of the inadequacy of the metrics.

[29] The framework of Markowitz (1952), as it is built on the L^2 norm, does not stand any form of empirical or even theoretical validity, owing to the dominance of higher moment effects, even in the presence of "finite" variance, see Taleb (2009).

8. Conclusion

To conclude, we offered a method of robustifying payoffs from large deviations and making forecasts possible to perform. The extensions can be generalized to a larger notion of society's safety—for instance how we should build systems (internet, banking structure, etc.) to be impervious to random effects.

Acknowledgements

A longer literary version of the ideas of this paper was posted on the web on the EDGE website at www.edge.org, eliciting close to 600 comments and letters, which helped in the elaboration of this version. The author thanks the commentators and various reviewers, and Yossi Vardi for the material on the events of Sept 18, 2008.

References

Albert, R., & Barabasi, A.-L. (2002). Statistical mechanics of complex network. *Review of Modern Physics*, *74*, 47–97.

Albert, R., Jeong, H., & Barabási, A.-L. (2000). Error and attack tolerance of complex networks. *Nature*, *406*, 378–382.

Amaral, L. A. N., Buldyrev, S. V., Havlin, S., Leschhorn, H., Maass, P., Salinger, M. A., et al. (1997). Scaling behavior in economics: I. Empirical results for company growth. *Journal de Physique I (France)*, *7*, 621.

Bak, P. (1996). *How nature works*. New York: Copernicus.

Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self-organized criticality: An explanation of 1/f noise. *Physical Review Letters*, *59*, 381–384.

Bak, P., Tang, C., & Wiesenfeld, K. (1988). Self-organized criticality. *Physical Review A*, *38*, 364–374.

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*, 509–511.

Baz, J., & Chacko, G. (2004). *Financial derivatives: Pricing, applications, and mathematics*. Cambridge University Press.

Bouchaud, J.-P., & Potters, M. (2003). *Theory of financial risks and derivatives pricing, from statistical physics to risk management* (2nd ed.). Cambridge University Press.

Bundt, T., & Murphy, R. P. (2006). *Are changes in macroeconomic variables normally distributed? Testing an assumption of neoclassical economics*. Preprint, NYU Economics Department.

Burlando, B. (1993). The fractal geometry of evolution. *Journal of Theoretical Biology*, *163*(2), 161–172.

DeMiguel, Garlappi, & Uppal, (2007). *Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy?* Working Paper.

Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and asset pricing for affine jump diffusions. *Econometrica*, *68*, 1343–1376.

Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the Nationall Academy of Sciences of the United States of America*, *98*(24), 13763–13768.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics*, *75*, 643–669.

Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling extremal events for insurance and finance*. Berlin: Spring Verlag.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom. *Econometrica*, *50*(4), 987–1007.

Enquist, B. J., & Niklas, K. J. (2001). Invariant scaling relations across tree-dominated communities. *Nature*, *410*(6829), 655–660.

Freedman, D. (2007). *Statistics* (4th ed.). W.W. Norton & Company.

Gabaix, X., Gopikrishnan, P., Plerou, V., & Stanley, H. E. (2003a). A theory of power law distributions in financial market fluctuations. *Nature*, *423*, 267–230.

Gabaix, X., Gopikrishnan, P., Plerou, V., & Stanley, H. E. (2003b). *Are stock market crashes outliers?* Mimeo 43.

Gabaix, X., Ramalho, R., & Reuter, J. (2003c). *Power laws and mutual fund dynamics*. MIT mimeo.

Gardenfors, P., & Sahlin, N. E. (1982). Unreliable probabilities, risk taking, and decision making. *Syn-these*, *53*, 361–386.

Goldstein, D. G., & Taleb, N. N. (2007). We don't quite know what we are talking about when we talk about volatility. *Journal of Portfolio Management*, *33*(4), 84–86.

Gopikrishnan, P., Meyer, M., Amaral, L., & Stanley, H. E. (1998). Inverse cubic law for the distribution of stock price variations. *European Physical Journal B*, *3*, 139–140.

Gopikrishnan, P., Plerou, V., Amaral, L., Meyer, M., & Stanley, H. E. (1999). Scaling of the distribution of fluctuations of financial market indices. *Physical Review E*, *60*, 5305–5316.

Gopikrishnan, P., Plerou, V., Gabaix, X., & Stanley, H. E. (2000). Statistical properties of share volume traded in financial markets. *Physical Review E*, *62*, R4493–R4496.

Harte, J., Kinzig, A., & Green, J. (1999). Self-similarity in the distribution and abundance of species. *Science*, *284*(5412), 334–336.

Haug, E. G. (2007). *Derivatives models on models*. New York: John Wiley & Sons.

Jorion, P. (2001). *Value-at-risk: The new benchmark for managing financial risk*. McGraw Hill.

Kahneman, D. (1999). Objective happiness. In D. Kahneman, E. Diener, & N. Schwartz (Eds.), *Well being: Foundations of hedonic psychology*. New York: Russell Sage Foundation.

Levi, I. (1986). The paradoxes of Allais and Ellsberg. *Economics and Philosophy*, *2*, 23–53.

Makridakis, S., & Hibon, M. (2000). The M3-competition: Results, conclusions and implications. *International Journal of Forecasting*, *16*, 451–476.

Makridakis, S., Chatfield, C., Hibon, M., Lawrence, M., Mills, T., Ord, K., et al. (1993). The M2-competition: A real-time judgmentally based forecasting study (with commentary). *International Journal of Forecasting*, *9*, 5–29.

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

Mandelbrot, B. (1997). *Fractals and scaling in finance*. Springer-Verlag.

Mandelbrot, B. (2001). *Quantitative Finance*, *1*, 113–123.

Mandelbrot, B., & Taleb, N. N. (2009). Mild vs. wild randomness: Focusing on risks that matter. In F. Diebold, N. Doherty, R. Herring (Eds.), *The known, the unknown and the unknowable in financial institutions*. Princeton, NJ: Princeton University Press (in press).

Markowitz, H. (1952). Portfolio Selection. *Journal of Finance*, *7*, 77–91.

May, R. M., Levin, S. A., & Sugihara, G. (2008). Complex systems: Ecology for bankers. *Nature*, *451*, 893–895.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, *3*, 125–144.

Merton, R. C. (1992). *Continuous-time finance* (revised edition). Blackwell.

Ritchie, M. E., & Olff, H. (1999). Spatial scaling laws yield a synthetic theory of biodiversity. *Nature*, *400*(6744), 557–560.

Solé, R. V., Manrubia, S. C., Benton, M., Kauffman, S., & Bak, P. (1999). Criticality and scaling in evolutionary ecology. *Trends in Ecological Evolution*, *14*(4), 156–160.

Sornette, D. (2004). *Critical phenomena in natural sciences: Chaos, fractals, self-organization and disorder: Concepts and tools* (2nd ed.). Berlin, Heidelberg: Springer.

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

Taleb, N. N. (2004). The practitioner's perspective: Bleed or blowup? Why do we prefer asymmetric payoffs? *Journal of Behavioral Finance*, *5*(1), 2–7.

Taleb, N. (2007a). *The black swan: The impact of the highly improbable*. US: Random House, Penguin (UK).

Taleb, N. N. (2007b). Black swans and the domains of statistics. *The American Statistician*, *61*(3), 198–200.

Taleb, N. N. (2009). Finiteness of variance is irrelevant in the practice of quantitative finance. *Complexity*, *14*(3), 66–76.

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