Finiteness of Variance is Irrelevant in the Practice of Quantitative Finance

Second version, June 2008

Nassim Nicholas Taleb

Abstract: Outside the Platonic world of financial models, assuming the underlying distribution is a scalable "power law", we are unable to find a consequential difference between finite and infinite variance models –a central distinction emphasized in the econophysics literature and the financial economics tradition. While distributions with power law tail exponents $\alpha>2$ are held to be amenable to Gaussian tools, owing to their "finite variance", we fail to understand the difference **_in the application_** with other power laws ($1<\alpha<2$) held to belong to the Pareto-Lévy-Mandelbrot stable regime. The problem invalidates derivatives theory (dynamic hedging arguments) and portfolio construction based on mean-variance. This paper discusses methods to deal with the implications of the point in a real world setting.

1. The Four Problems of Practice

This note outlines problems viewed solely from the vantage point of practitioners of quantitative finance and derivatives hedging, and the uneasy intersection of theories and practice; it aims at asking questions and finding robust and practical methods around the theoretical difficulties. Indeed, practitioners face theoretical problems and distinctions that are not visibly relevant in the course of their activities; furthermore some central practical problems appear to have been neglected by theory. Models are Platonic: going from theory to practice appears to be a direction that is arduous to travel. In fact, the problem may be even worse: seen from a derivatives practitioner's vantage point, theory may be just fitting (albeit with considerable delay), rather than influence, practice. This article is organized around a class of such problems, those related to the effect of power laws and scalable distributions on practice. We start from the basis that we have *no evidence against* Mandelbrot's theory that financial and commodity markets returns obey power law distributions [Mandelbrot, 1963, 1997], (though of unknown parameters). We do not even have an *argument* to reject it. We therefore need to find ways to effectively *deal with the consequences*.

We find ourselves at the intersection of two lines of research from which to find guidance: orthodox financial theory and econophysics. Financial theory has been rather silent on power laws (while accepting some mild forms of "fat tails" though not integrating them or

taking them to their logical consequences) –we will see that power laws (even with finite variance) are totally incompatible with the foundations of financial economics, both derivatives pricing and portfolio theory. As to the econophysics literature: by adopting power laws, but with artificial separation parameters, using $\alpha=2$, it has remedied *some* of the deficits of financial economics but has not yet offered us help for our problems in practice[1].

We will first identify four problems confronting practitioners related to the fat tailed distributions under treatment of common finance models, 1) problems arising from the use by financial theory, as a proxy for "fat tails", of milder forms of randomness too dependent on the Gaussian (non-power laws), 2) problems arising from the abstraction of the models and properties that only hold asymptotically, 3) problems related to the temporal independence of processes that lead to assume rapid convergence to the Gaussian basin, and, 4) problems related to the

---

[1] There has been a family of econophysics papers that derive their principal differentiation from Mandelbrot (1963) on the distinction between Levy-Stable basins and other power laws (infinite v/s finite variance): Plerou et al (2001), Stanley et al (2001), Gabaix et al (2003a, 2003b), Gopikrishnan et al (1998, 1999, 2000) and others. Furthermore, Plerou et al (2002) derive the "cubic" tail exponent $\alpha=3$ from order flow (a combination of $\alpha=3/2$ for company and order size and nonlinear square root impact for order impact) –a causal argument that seems unconvincing to any professional, particularly when the paper ignore the share of equity variations that comes from jumps around corporate announcements –such as earnings or mergers. These discontinuities occur *_before_* the increase in volume.

calibration of scalable models and to the fitting of parameters.

Note here that we provide the heuristic attribute of a scalable distribution as one where for some "large" value of x, $P_{>x} \sim K x^{-\alpha}$, where $P_{>x}$ is the "exceedant probability", the probability of exceeding x, and K is a scaling constant. (Note that the same applies to the negative domain). The main property under concern here, which illustrates its scalability, is that, in the tails, $P_{>nx}/P_{>x}$ depends on n, rather than x.

The Criterion of Unboundedness: One critical point for deciding the controversial question "is the distribution a power law?" Unlike the econophysics literature, we do not necessarily believe that the scalability holds for *x* reaching infinity; but, in practice, so long as we do not know *where* the distribution is eventually truncated, or what the upper bound for *x* is, we are forced operationally to use a power law. Simply as we said, we cannot safely reject Mandelbrot [1963, 1997]. In other words, it is the uncertainty concerning such truncation that is behind our statement of scalability. it is easy to state that the distribution might be lognormal, which mimics a power law for a certain range of values of *x*. But the uncertainty coming from <u>where</u> the real distribution starts becoming vertical on a Log-Log plot (i.e. α rising towards infinity) is central – statistical analysis is marred with too high sample errors in the tails to help us. This is a common problem of practice v/s theory that we discuss later with the invisibility of the probability distribution[2] [For a typical misunderstanding of the point, see Perline 2005].

Another problem: the unknowability of the upper bound invites faulty stress testing. Stress testing (say, in finance) is based on a probability-free approach to simulate a single, fixed, large deviation – as if it were the known payoff from a lottery ticket. However the choice of a "maximum" jump or a "maximum likely" jump is itself problematic, as it assumes knowledge of the structure of the distribution in the tails[3]. By assuming that tails are power-law distributed, though of unknown exact parameter, one can project richer sets of possible scenarios.

---

[2] The inverse problem can be quite severe – leading to the mistake of assuming stochastic volatility (with the convenience of all moments) in place of a scale-free distribution (or, equivalently, one of an unknown scale). Cont and Tankov (2003) show how a Student T with 3 degrees of freedom (infinite kurtosis will mimic a conventional stochastic volatility model.

[3] On illustration of how stress testing can be deemed dangerous –as we do not have a typical deviation – is provided by the management of the 2007-2008 subprime crisis. Many firms, such as Morgan Stanley, lost large sums of their capital in the 2007 subprime crisis because their stress test underestimated the outcome –yet was compatible with historical deviations (see "The Risk Maverick", Bloomberg, May 2008).

The finance literature uses variance as a measure of dispersion for the probability distributions, *even when dealing with fat tails*. This creates a severe problem outside the pure Gaussian nonscalable environment.

Financial economics is grounded in general Gaussian tools, or distributions that have all finite moments and correspondingly a characteristic scale, a category that includes the Log-Gaussian as well as subordinated processes with non-scalable jumps such as diffusion-Poisson, regime switching models, or stochastic volatility methods [see Hull, 1985;Heston,1993; Duffie, Pan, and Singleton, 2000, Gatheral, 2006], outside what Mandelbrot, 1997, bundles under the designation scale-invariant or fractal randomness.

All of these distributions can be called "fat-tailed", but not scalable in the above definition, as the finiteness of all moments makes them collapse into thin tails:

      1) for *some* extreme deviation (in excess of some known level) ,
or

      2) rapidly under convolution, or temporal aggregation. Weekly or monthly properties are supposed to be closer, in distribution, to the Gaussian than daily ones. Likewise, fat tailed securities are supposed to add up to thin-tailed portfolios, as portfolio properties cause the loss of fat-tailed character rather rapidly, thanks to the increase in the number of securities involved.

The dependence on these "pseudo-fat tails", or finite moment distributions, led to the building of tools based on the Euclidian norm, like variance, correlation, beta, and other matters in L^2 . It makes *finite* variance necessary for the modeling, and *not because the products and financial markets naturally require such variance*. We will see that the scaling of the distribution that affect the pricing of derivatives is the mean expected deviation, in L^1 , which does not justify such dependence on the Eucledian metric.

The natural question here is: why do we use variance? While it may offer some advantages, as a "summary measure" of the dispersion of the random variable, it is often meaningless outside of an environment in which higher moments do not lose significance.

But the practitioner use of variance can lead to additional pathologies. Taleb and Goldstein [2007] show that most professional operators and fund managers use a mental measure of mean deviation as a substitute for variance, without realizing it: since the literature focuses exclusively on L^2 metrics, such as

"Sharpe ratio", "portfolio deviations", or sigmas". Unfortunately the mental representation of these measures is elusive, causing a substitution. There seems to be a serious disconnect between decision making and projected probabilities. Standard deviation is exceedingly unstable compared to mean deviation in a world of fat tails (see an illustration in Figure 1).
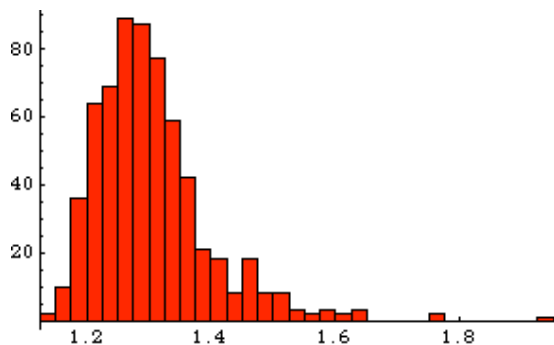

Figure 1 Distribution of the monthly STD/MAD ratio for the SP500 between 1955 and 2007

1.2 SECOND PROBLEM, "LIFE OUTSIDE THE ASYMPTOTE": QUESTIONS STEMMING FROM IDEALIZATION V/S PRACTICE

The second, associated problem comes from the idealization of the models, often inexactly the wrong places for practitioners, leading to the reliance on results that work in the asymptotes, *and only* in the asymptotes. Furthermore the properties outside the asymptotes are markedly different from those *at* the asymptote. Unfortunately, operators live far away from the asymptote, with nontrivial consequences for pricing, hedging, and risk management.

a) Time aggregation

Take the example of a distribution for daily returns that has a finite second moment, but infinite kurtosis, say a power-law with exponent <4, of the kind we observe routinely in the markets. It will eventually, under time aggregation, say if we lengthen the period to weekly, monthly, or yearly returns, converge to a Gaussian. But this will only happen *at infinity*. The distribution will become increasingly Gaussian in the center, **but not in the tails**. Bouchaud and Potters,2002, show how such convergence will be extremely slow, at the rate of $\sqrt{n\log(n)}$ standard deviations, where n is the number of observations. A distribution with a power law exponent α >2, even with a million convolutions, will eventually behave like a Gaussian up until about 3 standard deviations but conserve the power-law attributes outside of such regime. So, at best we are getting a mixed distribution, with the same fat tails as a nonGaussian --and the tails are where the problems reside.

More generally, the time-aggregation of probability distributions with some infinite moment will not obey the Central Limit Theorem in applicable time, thus leaving us with non-asymptotic properties to deal with in an effective manner Indeed it may not be even a matter of time-window being too short, but for distributions with finite second moment, but with an infinite higher moment, for CLT to apply we need an infinity of convolutions.

b) Discreteness

We operate in discrete time while much of the theory concerns mainly continuous time processes [Merton, 1973,1992], or finite time operational or computational approximations to true continuous time processes [Cox and Ross,1976; review in Baz and Chacko, 2004]. Accordingly, the results coming from taking the limits of continuous time models, all Gaussian-based (non-scalable) pose difficulties in their applications to reality.

A scalable, unlike the Gaussian, does not easily allow for continuous time properties, because the continuous time limit allowing for the application of Ito's lemma is not reached, as we will see in section 2.2.

1.3 THIRD PROBLEM- STABILITY AND TIME DEPENDENCE.

Most of the mathematical treatment of financial processes reposes on the assumption of time-independence of the returns. Whether it is for mathematical convenience (or necessity) it is hard to ascertain; but it remains that most of the distinctions between processes with finite second moment and others become thus artificial as they reposes heavily on such independence.

The consequence of such time dependence is the notion of distributional "stability", in the sense that a distribution loses its properties with the summation of random variables drawn from it. Much of the work discriminating between Levy-fat tails and non Levy fat tails reposes on the notion that a distribution with tail exponent $\alpha<2$ is held to converge to a Levy stable basin, while those with $\alpha\geq2$ are supposed to become Gaussian.

The problem is that such notion of independence is a bit too strong for us to take it at face value. There may be serial independence in returns, but coexisting with some form of serial dependence in absolute returns -- and the consequences on the tools of analysis are momentous[4].

---

[4] One candidate process is Mandelbrot's multifractal model [Mandelbrot,2004] in which the tail exponent conserves under addition. Daily returns can have a $\alpha=3$, so will monthly returns.

In other words, even in the asymptote, a process with finite variance that is not independently distributed is not guaranteed to become Gaussian.

This point further adds to the artificiality of the distinction between $\alpha<2$ and $\alpha\geq2$. Results of derivatives theory in the financial literature exclude path dependence and memory, which causes the aggregation of the process to hold less tractable properties than expected--making the convergence to a Gaussian basin of attraction less granted. While the returns may be independent, absolute values of these returns may not be, which causes extreme deviations to cumulate in a manner to fatten the tails at longer frequencies [See Sornette,2004, for the attributes of the drawdowns and excursions as these are more extreme than regular movements; deviations in the week of the stock market crash of 1987 were more extreme, statistically, than the day of the greatest move].Naively, if you measure the mean average deviation of returns over a period, then lag them you will find that the measure of deviation is sensitive to the lagging period.[Also see Lo and McKinley,1995,for a Gaussian test].

1.4 FOURTH PROBLEM- THE VISIBILITY OF STATISTICAL PROPERTIES IN THE DATA

The final problem is that operators do not observe probability distributions, only realizations of a stochastic process, with a spate of resulting mistakes and systematic biases in the measurement process [Taleb and Pilpel, 2004; Taleb, 2007]. Some complicated processes with infinite variance will tend exhibit finite variance under the conventional calibration methods, such as the Hill estimator or the Log-Linear regressions [Weron,2001]. In other words, for some processes, the typical error can be tilted towards the underestimation of the thickness of the tails. A process with an $\alpha=1.8$ can easily yield $\alpha >2$ in observations.

An argument in favor of "thin tails", or truncated power laws, is usually made with representations of the exceedant frequencies in log-log space that show the plot line getting vertical at some point, indicating an a pulling towards infinity. The problem is that it is hard to know whether this cutoff is genuine --and not the result of sample insufficiency. Such perceived cutoff can easily be the result of sampling error, given that we should find fewer data points in the far tails [Taleb, 2005]. But in fact, assuming truncation is acceptable, we do not know *where* the distribution is to be truncated. Relying on the past yields in-sample obvious answers [Mandelbrot, 2004] --but it does not reveal the true nature of the generator of the series. In the same vein, many researchers suggest the lognormal [Perline, 2005, see Mandelbrot 1997 for the review], or stretched exponential [Sornette, 2004]. On that score, the financial economics literature presents circular

arguments, favoring Poisson jumps, and using the same assumed distribution to gauge the sufficiency of the sample, without considering the limitations of the sample in revealing tail events [Coval and Shumway,2000; Bodarenko, 2004 [5]]. (Simply, rare events are less likely to show in a finite sample; assuming homogeneous past data, 20-year history will not reveal one-in-50-year events.) The best answer, for a practitioner, is to plead ignorance: so long as we do not know where the truncation starts, it is safer to stick to the assumption of power laws.

1.5 THE MAJOR CONSEQUENCE OF THESE FOUR PROBLEMS

The cumulation of these four problems results in the following consequence: in "real life", the problems incurred when the tail exponent $\alpha<2$ effectively prevail just as well when $\alpha>2$. The literature [Gopikrishnan et al, 1998, Gabaix et al, 2003, 2004] reports "evidence" in the equities markets, of a cubic α, i.e., a tail exponent around 3. Their dataset of around 18 million observations is available to most practitioners (including this author); it is extremely easy to confirm the result -- in sample. The econonophysics literature thus makes the distinction between Levy-regime and other whereas for us practitioners, because of the time aggregation problem, there is no such distinction. *A scalable is a scalable*: the tails never become thin enough to allow the use of Gaussian methods.

We will need to consider the consequences of the following two considerations:

> 1) The infinite moments never allow for derivations based on expansions and Ito's lemma
>
> 2) Processes do not necessarily converge to the Gaussian basin, making conventional tools like standard deviations inapplicable.
>
> 3) Parameter discovery is not as obvious as in the Gaussian world.

The main option pricing and hedging consequences of scalability do not arise from the finiteness of the variance but rather from the lack of convergence of higher moments. Infinite kurtosis, which is what empirical data seems to point to in almost every market examined, has the same effect. There are no tangible, or qualitative differences in practice between such earlier models such as Mandelbrot [1963], on one hand and later expositions showing finite variance models with a "cubic" tail exponent. Fitting these known

---

[5] In addition, these tests are quite inadequate outside of L^2 since they repose of measurement and forecast of variance.

processes induces the cost of severe mistracking of empirical reality.

The rest of this article will focus on the application of the above four problems and its consequence to derivatives pricing. We will first present dynamic hedging that is at the center of modern finance's version of derivatives pricing, and its difficulty outside of the Gaussian case owing to incompressible tracking errors. We then show how variance does not appear relevant for an option operator and that distributions with infinite variance are not particularly bothersome outside of dynamic hedging. Then we examine methods of pricing followed with the common difficulties in working with NonGaussian distributions with financial products. We examine how these results can be extended to portfolio theory.

2- The applications

2.1 FINITE VARIANCE IS INSUFFICIENT FOR PORTFOLIO THEORY

First, let us consider portfolio theory. There appears to be an accepted truism [after Markowitz, 1952], that mean-variance portfolio allocation requires, but can be satisfied with, only the first two moments of the distribution –and that fatness of tails do not invalidate the arguments presented. I leave aside the requirement for a certain utility structure (a quadratic function) to make the theory work, and assume it to be acceptable in practice[6].

Where x is the payoff (or wealth), and U the utility function:

$$U(x) = a\,x - b\,x^2\,,\ a,\ b>0,$$

By taking expectations, the utility of x

$$E[U(x)] = a\,E[x] - b\,E[x^2]$$

So seemingly higher moments do not matter. Such reasoning may work in the Platonic world of models, but, when turned into an application, even without relaxing any of the assumptions, it reveals a severe defect: where do we get the parameters from? $E[x^2]$, even if finite, *is not observable*. A distribution with infinite higher moments $E[x^7]$ (with n>2) will not reveal its properties in finite sample. Simply, if $E[x^4]$ is infinite, $E[x^2]$ will not show itself easily. The expected utility will remain stochastic, i.e., unknown. Much of the problems in financial theory come from the dissipation upon application of one of the central hypotheses: that the operator knows the parameters of the distribution (an

application of what Taleb, 2007, calls the "ludic fallacy").

The idea of mean-variance portfolio theory then has no possible practical justification.

2.2 DIFFICULTIES WITH FINANCIAL THEORY'S APPROACH TO OPTION PRICING

Technë-Epistemë: The idea that operators need theory, rather than the other way around, has been contradicted by historical evidence [Taleb and Haug, 2008]. They showed how option traders managed in a quite sophisticated manner to deal with option pricing and hedging –there is a long body of literature, from 1902, ignored by the economics literature presenting trading techniques and heuristics. The literature had been shy in considering the hypothesis that option price formation stems from supply and demand, and that traders manage to develop tricks and methods to eliminate obvious mispricing called "free lunches". So the result is that option theory seems to explain (or simplify) *what is being done* rather than drive price formation. Gatheral (2006) defines the profession of modelers as someone who finds equations that fit prices in the market prices with minimal errors, rather than the reverse. Accordingly the equations about the stochastic process do not have much beyond an instrumental use (to eliminate inconsistencies) –they do not correspond to a representation of future states of the world[7].

Standard Financial Theory: Let us now examine how financial theory "values" financial products (using an engineering/practitioner exposition). The principal difference in paradigm between the one presented by Bachelier, 1900, and the modern finance one known as Black-Scholes-Merton [Black-Scholes,1973, and Merton, 1973] lies in the following main point. Bachelier's model is based on an actuarial expectation of final payoffs. The same method was later used by a series of researchers, such as Sprenkle (1964), Boness (1964), Kassouf and Thorp (1967), Thorp (1970). They all encountered the following problem: how to produce a risk parameter –a risky asset discount rate -- to make it compatible with portfolio theory? The Capital Asset Pricing Model requires that securities command an expected rate of return in proportion to their "riskiness". In the Black-Scholes-Merton approach, an option price is derived from continuous-time dynamic hedging, and only in properties obtained from continuous time dynamic hedging --we will describe dynamic hedging in some details further down. Thanks to such a method, an option collapses into a

---

[6] This is one of the arguments against the results of Mandelbrot (1963) using the "evidence" provided by Officer (1972).

[7] In the same vein, we repeat that the use of power-laws does not necessarily correspond to the belief that the distribution is truly parametrized as a power law, rather selected owing to the absence of knowledge of the properties in the tails.

deterministic payoff and provides returns independent of the market; hence it does not require any risk premium.

The problem we have with the Black-Scholes-Merton approach is that the requirements for dynamic hedging are extremely idealized, requiring the following strict conditions –idealization might have gone too far, and dangerously so, of the style "assume the earth was square". The operator is assumed to be able to buy and sell in a frictionless market, incurring no transaction costs. The procedure does not allow for the price impact of the order flow--if an operator sells a quantity of shares, it should not have consequences on the subsequent price. The operator knows the probability distribution, which is the Gaussian, with fixed and constant parameters through time (all parameters do not change). Finally, the most significant restriction: no scalable jumps. In a subsequent revision [Merton, 1976] allows for jumps but these are deemed to be Poisson arrival time, and fixed or, at the worst, Gaussian. The framework does not allow the use of power laws both in practice and mathematically. Let us examine the mathematics behind the stream of dynamic hedges in the Black-Scholes-Merton equation.

Assume the risk-free interest rate $r=0$ with no loss of generality. The canonical Black-Scholes-Merton model consists in building a dynamic portfolio by selling a call and purchasing shares of stock that provide a hedge against instantaneous moves in the security. Thus the value of the portfolio π locally "hedged" against exposure to the first moment of the distribution is the following:

$$\pi = -C + \frac{\partial C}{\partial S} S$$

where C is the call price, and S the underlying security.

By expanding around the initial values of the underlying security S, we get the changes in the portfolio in discrete time. Conventional option theory applies to the Gaussian in which all orders higher than ΔS^2 and Δt (including the cross product $\Delta S \, \Delta t$) are neglected.

$$\Delta \pi = \frac{\partial C}{\partial t} \Delta t - \frac{1}{2} \frac{\partial^2 C}{\partial S^2} \Delta S^2 + O(\Delta S^3)$$

Taking expectations on both sides, we can see very strict requirements on moment finiteness: *all* moments need to converge for us to be comfortable with this framework. If we include another term, ΔS^3, it may be of significance in a probability distribution with significant cubic or quartic terms. Indeed, although the nth derivative with respect to S can decline very sharply, for options that have a strike K away from the initial price S, it remains that the moments are rising

disproportionately fast, enough to cause potential trouble.

So here we mean all moments need to be finite and losing in impact --no approximation would do. Note here that the jump diffusion model (Merton,1976) does not cause much trouble for researchers since it has all the moments --which explains its adoption in spite of the inability to fit jumps in a way that tracks them out-of-sample. And the annoyance is that a power law will have every moment higher than a infinite, causing the equation of the Black-Scholes-Merton portfolio to fail.

As we said, the logic of the Black-Scholes-Merton so-called "solution" is that the portfolio collapses into a deterministic payoff. But let us see how quickly or effectively this works in practice.

The actual replication process: According to standard financial economics [Merton, 1992], the payoff of a call is expected to be replicated *in practice* with the following stream of dynamic hedges. The procedure is as follows (again, assuming 0 interest rates): Take C the initial call price, S_t the underlying security at initial period t, and T the final expiration of the option. The performance will have three components: 1) C the initial call value as cash earned by the option seller, 2) Max $(S_T–K,0)$, the final call value (intrinsic value) that the option seller needs to disburse, and 3) the stream of dynamic hedges aiming at offsetting the pair C - Max $(S_T–K,0)$, in quantities of the underlying held in inventory, revised at different periods.

So we are concerned with the evolution between the two periods t and T –and the stream of dynamic hedges. Break up the period (T-t) into n increments Δt. The operator changes the hedge ratio, i.e. the quantities of the underlying security he is supposed to have in inventory, as of time $t +(i-1) \Delta t$, then gets the difference between the prices of S at periods $t +(i-1) \Delta t$ and $t +i \Delta t$ (called nonanticipating difference). Where P is the final profit/loss:

$$P = -C + (K - S_T) + \sum_{t=1}^{n=\frac{T-t}{\Delta t}} \frac{\partial C}{\partial S}\Bigg|_{S=S_{t+(t-1)\Delta t}, t=t+(i-1)\Delta t} (S_{t+i\Delta t} - S_{t+(t-1)\Delta t})$$

Standard option theory considers that the final package P of the three components will become deterministic at the limit of $\Delta t \to 0$, as the stream of dynamic hedges reduces the portfolio variations to match the option. This seems mathematically and operationally impossible[8].

Failure: How hedging errors can be prohibitive.

---

[8] See Bouchaud and Potters (2002) for a critique.

6

As a consequence of the mathematical property seen above, hedging errors in an cubic α appear to be indistinguishable from those from an infinite variance process. Furthermore such error has a disproportionately large effect on strikes, as we illustrate in Figures 2, 3 and 4. Figure 2 illustrate the portfolio variations under a finite variance power law distribution, subjected to the same regime of revision ($\Delta t = 1$ business day, 1/252), compared to the Gaussian in Figure 3. Finally Figure 4 shows the real market (including the crash of 1987).
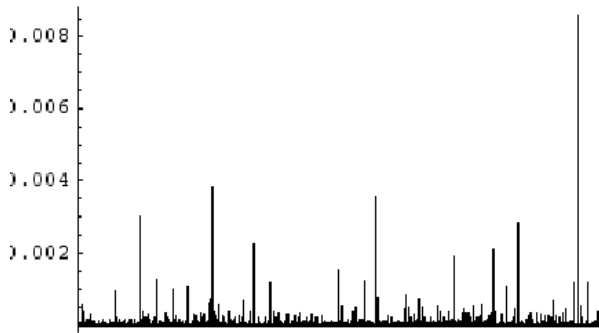

Figure 2 shows the hedging errors for an option portfolio (under a daily revision regime) over 3000 days, under a constant volatility Student T with tail exponent $\alpha=3$. Technically the errors should not converge in finite time as their distribution has infinite variance.
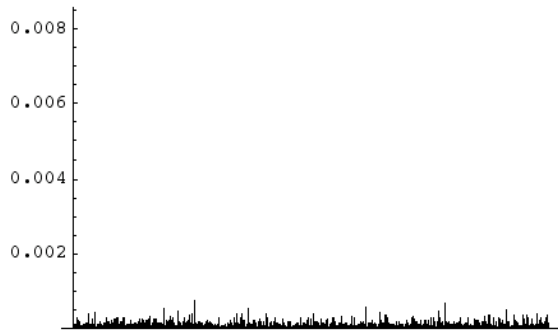

Figure 3 -Hedging errors for an option portfolio (equivalent daily revision) under an equivalent "Black-Scholes" world.
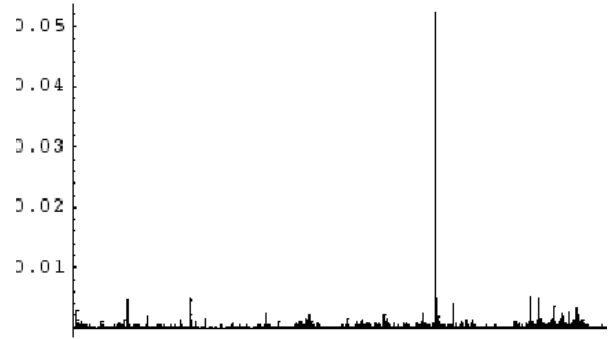

Figure 4 -Portfolio Hedging errors including the stock market crash of 1987.

In short: dynamic hedging in a power law world does not remove risk.

Options without variance

Based on this dynamic hedging problem, we look at which conditions we need to "price" the option. Most models base the option on variance. Clearly options do not depend on variance, but on mean average deviation, --but *it is expressed in terms of variance*.

Consider the Bachelier expectation framework --the actuarial method of discounting the probabilistic payoffs of the options. Where F is the forward, the price in the market for the delivery of S at period T, and regardless of the probability distribution φ, under the sole restriction that the first moment $\int F \varphi(F) \, dF$ exists, the puts and calls can be priced as follows (assuming, to simplify, 0 financing rates):

$$C(K) = \int_{K}^{\infty} (F - K)\phi(F)\,dK$$

$$P(K) = \int_{0}^{K} (K - F)\phi(F)\,dK$$

Where C(K) and P(K) are the call and put struck at K, respectively. Thus when the options is exactly at-the-money by the forward, i.e. *K=F*, each delivers half the discounted mean absolute deviation

$$C(K) = P(K) = \frac{1}{2}\int_{0}^{\infty} (\Delta F)\phi(F)\,dK$$

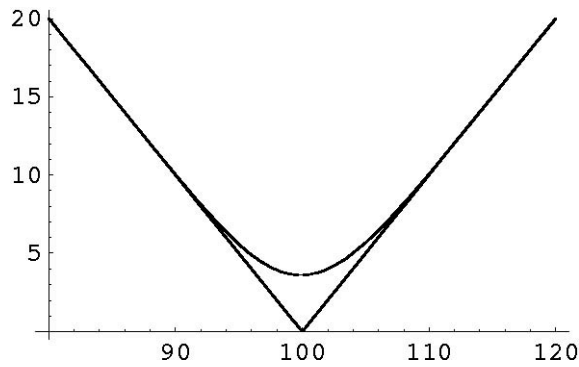An option's payoff is piecewise linear as can be shown in Figure 5.

Figure 5 Straddle price: options are piecewise linear, with a hump at the strike –the variance does not enter the natural calculation.

In a Gaussian world, we have the mean absolute deviation over standard deviation as folllows:

$$\frac{\overline{m}}{\sigma} = \sqrt{\frac{2}{\pi}}$$

But, with fat tails, the ratio of the dispersion measures drops, as σ reaches infinity when $1 < \alpha < 2$. This means that a simple sample of activity in the market will not reveal much since most of the movements become concentrated in a fewer and fewer number of observations. Intuitively 67% of observations take place in the "corridor" between +1 and -1 standard deviations in a Gaussian world. In the real world, we observe between 80% and 99% of observations in that range –so large deviations are rare, yet more consequential. Note that for the conventional results we get in finance ["cubic"] α , about 90.2% of the time is spent in the [-1,+1] standard deviations corridor. Furthermore, with α=3, the previous ratio becomes

$$\frac{\overline{m}}{\sigma} = \frac{2}{\pi}$$

Case of a variance swap

There is an exception to the earlier statement that derivatives do not depend on variance. The only common financial product that depends on L^2 is the "variance swap": a contract between two parties agreeing to exchange the difference between an initially predetermined price and the delivered variance of returns in a security. However the product is not replicable with single options.

The exact replicating portfolio is constructed [Gatheral, 2006] *in theory* with an infinity of options spanning all possible strikes, weighted by K^{-2} , where K is the strike price.

$$\int_0^\infty \frac{1}{K^2} C(K)\, dK - \Delta S$$

However, to turn this into practice in the real world requires buying an infinite amount of options of a strike K approaching 0, and an inifinitesimal amount of options with strike K approaching infinity. In the real world, strikes are discrete and there is a lower bound K_L and a highest possible one K_H. So what the replication leaves leaves out, $< K_L$ and $> K_H$ leaves us exposed to the large deviations; and would cost an infinite amount to purchase when options have an infinite variance.

The discrete replicating portfolio would be as follows: by separating the options into n strikes between K_L and K_H incrementing with ΔK.

$$\sum_{i=1}^{n} \frac{C(i\Delta K + K_L) + P(i\Delta K + K_L)}{\left(\frac{(K_H - K_L)i}{n} + K_L\right)^2}$$

Such portfolio will be extremely exposed to mistracking upon the occurrence of tail events.

We conclude this section with the remark that, effectively, and without the granularity of the market, the dynamic hedging idea seriously underestimates the effectiveness of hedging errors –from discontinuities and tail episodes. As a matter of fact it plainly does not seem to work both practically and mathematically.

The minimization of daily variance may be effectual in smoothing performance from small moves, but it fails during large variations. In a Gaussian basin very small probability errors do not contribute to too large a share of total variations; in a true fat tails environment, and with nonlinear portfolios, the extreme events dominate the properties. More specifically, an occasional sharp move, such a "22 sigma event" (expressed in Gaussian terms, by using the standard deviation to normalize the market variations), of the kind that took place during the stock market crash of 1987, would cause a severe loss that would cost years to recover. Define the "daily time decay" as the drop in the value of the option over 1/252 years assuming no movement in the underlying security, a crash similar to 1987 would cause a loss of close to hundreds of years of daily time decay for a far out of the money option, and more than a year for the average option[9].

---

[9] One can also fatten the tails of the Gaussian, and get a power law by changing the standard deviation of the Gaussian: Dupire(1994), Derman and Kani (1994), Borland (2002), See Gastheral [2006] for a review.

8

It is not a matter of "can". We "need" to do so once we lift the idealized conditions --and we need to focus on the properties of the errors.

We just saw that scalability precludes dynamic hedging as a means to reach a deterministic value for the portfolio. There are of course other impediments for us --merely the fact that in practice we cannot reach the level of comfort owing to transaction costs, lending and borrowing restrictions, price impact of actions, and a well known problem of granularity that prevent us from going to the limit. In fact continuous-time finance [Merton,1992] is an idea that got plenty of influence in spite of both its mathematical stretching and its practical impossibility.

If we accept that returns are power-law distributed, then finite or infinite variance matter little. We need to use expectations of terminal payoffs, and not dynamic hedging. The Bachelier framework, which is how option theory started, does not require dynamic hedging. Derman and Taleb (2005) argue how one simple financing assumption, the equality of the cost-of-carry of both puts and calls, leads to the recovery of the Black-Scholes equation in the Bachelier framework without any dynamic hedging and the use of the Gaussian. Simply a European put hedged with long underlying securities has the same payoff as the call hedged with short underlying securities and we can safely assume that cash flows must be discounted in an equal manner. This is common practice in the trading world; it might disagree with the theories of the Capital Asset Pricing Model, but this is simply because the tenets behind CAPM do not appear to draw much attention on the part of practitioners, or because the bulk of tradable derivatives are in fixed-income and currencies, products concerned by CAPM. Futhermore, practitioners, are not concerned by CAPM (see Taleb, 1997, Haug, 2006, 2007). We do not believe that we are modeling a true expectation, rather fitting an equation to work with prices. We do not "value". Finally, thanks to this method, we no longer need to assume continuous trading, absence of discontinuity, absence of price impact, and finite higher moments. In other words, we are using a more sophisticated version of the Bachelier equation; but it remains the Bachelier (1900) nevertheless. And, of concern here, we can use it with power laws with or without finite variance.

This said, while a Gaussian process provides a great measure of analytical convenience, we have difficulties building an elegant, closed-form stochastic process with scalables.

Working with conventional models present the following difficulties:

First difficulty: building a stochastic process

For pricing financial instruments, we can work with terminal payoff, except for those options that are path-dependent and need to take account of full-sample path.

Conventional theory prefers to concern itself with the stochastic process $dS/S = m\ dt + \sigma\ dZ$ (S is the asset price, m the drift, t, is time, and σ the standard deviation) owing to its elegance, as the relative changes result in exponential limit, leading to summation of instantaneous logarithmic returns, and allow the building of models for the distribution of price with the exponentiation of the random variable Z, over a discrete period Δt, $S_{t+\Delta t} = S_t\ e^{a+bz}$. This is convenient: for the expectation of S, we need to integrate an exponentiated variable

$$E[S_{t+\Delta t}] = S \int e^{a+bz}\ \phi(z)\ dz$$

which means that we need the density of z, in order to avoid finite expectations for S, to present a compensating exponential decline and allow bounding the integral. This explains the prevalence of the Lognormal distribution in asset price models –it is convenient, if unrealistic. Further, it allows working with Gaussian returns for which we have abundant mathematical results and well known properties.

Unfortunately, we cannot consider the real world as Gaussian, not even Log-Gaussian, as we've seen earlier.

How do we circumvent the problem? This is typically done at the cost of some elegance. Many operators [Wilmott, 2006] use, even with a Gaussian, the arithmetic process first used in Bachelier ,1900,

$$S_{t+\Delta t} = a + bz + S_t$$

which was criticized for delivering negative asset values (though with a minutely small probability). This process is used for interest rate changes, as monetary policy seems to be done by fixed cuts of 25 or 50 basis points regardless of the level of rates (whether they are 1% or 8%).

Alternatively we can have recourse to the geometric process for a , large enough, Δt (say one day)

$$S_{t+\Delta t} = S_t (1 + a + bz)$$

Such geometric process can still deliver negative prices (a negative, extremely large value for z) but is more in line with the testing done on financial assets, such as the SP500 index, as we track the daily returns $r_t=(P_t-P_{t-1})/P_{t-1}$ in place of $\log(P_t /P_{t-1})$. The distribution can be easily truncated to prevent negative prices (in practice the probabilities are so small that it does not have to be done as it would drown in the precision of the computation).

Second difficulty: dealing with time dependence

We said earlier that a multifractal process conserves its power law exponent across timescales (the tail exponent a remains the same for the returns between periods t and $t+\Delta t$, independently of Δt). We are not aware of an elegant way to express the process mathematically, even computationally --nor can we do so with any process that does not converge to the Gaussian basin. But for practitioners, theory is not necessary tricks. However this can be remedied in pricing of securities by, simply, avoiding to work with processes, and limiting ourselves to working with distributions between two discrete periods. Traders call that "slicing" [Taleb, 1997], in which we work with different periods, each with its own sets of parameters. We avoid studying the process between these discrete periods.

3- Concluding Remarks

This paper outlined the following difficulties: working in quantitative finance, portfolio allocation, and derivatives trading while being suspicious of the idealizations and assumptions of financial economics, but avoiding some of the pitfalls of the econophysics literature that separate models across tail exponent $\alpha=2$, truncate data on the occasion, and produce results that depend on the assumption of time independence in their treatment of processes. We need to find bottom-up patches that keep us going, in place of top-down, consistent but nonrealistic tools and ones that risk getting us in trouble when confronted with large deviations. We do not have many theoretical answers, nor should we expect to have them soon. Meanwhile option trading and quantitative financial practice will continue under the regular tricks that allow practice to survive (and theory to follow).

REFERENCES

Andersen, Leif, and Jesper Andreasen, 2000, Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing, Review of Derivatives Research 4, 231–262.

Bachelier, L. (1900): Theory of speculation in: P. Cootner, ed., 1964, The random character of stock market prices, MIT Press, Cambridge, Mass.

Baz, J., and G. Chacko, 2004, Financial Derivatives: Pricing, Applications, and Mathematics, Cambridge University Press

Black, F., and M. Scholes, 1973, The pricing of options and corporate liabilities, Journal of Political Economy 81, 631–659.

Boness, A. (1964): "Elements of a Theory of Stock-Option Value," Journal of Political Economy, 72, 163–175.

Borland, L.,2002, "Option pricing formulas based on a non-Gaussian stock price model, *Physical Review Letters*, 89,9.

Bouchaud J.-P. and M. Potters (2001): "Welcome to a non-Black-Scholes world", *Quantitative Finance*, Volume 1, Number 5, May 01, 2001 , pp. 482–483(2)

Bouchaud J.-P. and M. Potters (2003): Theory of Financial Risks and Derivatives Pricing, From Statistical Physics to Risk Management, 2nd Ed., Cambridge University Press.

Breeden, D., and R. Litzenberger, 1978, Prices of state-contingent claims implicit in option prices, Journal of Business 51, 621–651.

Clark, P. K., 1973, A subordinated stochastic process model with finite variance for speculative prices, Econometrica 41, 135–155.

Cont, Rama & Peter Tankov, 2003, Financial modelling with Jump Processes, Chapman & Hall / CRC Press, 2003.

Cox, J. C. and Ross, S. A. (1976), The valuation of options for alternative stochastic processes, Journal of Financial Economics 3: 145-166.

Demeterfi, Kresimir, Emanuel Derman, Michael Kamal, and Joseph Zou, 1999, A guide to volatility and variance swaps, Journal of Derivatives 6, 9–32.

Derman, Emanuel, and Iraj Kani, 1994, Riding on a smile, Risk 7, 32–39.

Derman, Emanuel, and Iraj Kani, 1998, Stochastic implied trees: Arbitrage pricing with stochastic term and strike structure of volatility, International Journal of Theoretical and Applied Finance 1, 61–110.

Derman, E., and N. N. Taleb (2005): "The Illusion of Dynamic Delta Replication," *Quantitative Finance*, 5(4), 323–326. [49]

Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform analysis and asset pricing for affine jump diffusions, Econometrica 68, 1343– 1376.

Dupire, Bruno, 1994, Pricing with a smile, Risk 7, 18–20. , 1998, A new approach for understanding the. impact of volatility on option prices, Discussion paper Nikko Financial Products.

Friz, Peter, and Jim Gatheral, 2005, Valuation of volatility derivatives as an inverse problem, Quantitative Finance 5, 531—-542.

Gatheral, James, 2006, Stochastic Volatility Modeling, Wiley. Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou and H. Eugene Stanley, "A theory of power law distributions in financial market fluctuations," Nature, 423 (2003a), 267—230.

Gabaix, Xavier, Parameswaran Gopikrishnan and Vasiliki Plerou, H. Eugene Stanley, "Are stock market crashes outliers?", mimeo (2003b) 43

Gabaix, Xavier, Rita Ramalho and Jonathan Reuter (2003) "Power laws and mutual fund dynamics", MIT mimeo (2003c).

Gopikrishnan, Parameswaran, Martin Meyer, Luis Amaral and H. Eugene Stanley "Inverse Cubic Law for the Distribution of Stock Price Variations," European Physical Journal B, 3 (1998), 139-140.

Gopikrishnan, Parameswaran, Vasiliki Plerou, Luis Amaral, Martin Meyer and H. Eugene Stanley "Scaling of the Distribution of Fluctuations of Financial Market Indices," Physical Review E, 60 (1999), 5305-5316.

Gopikrishnan, Parameswaran, Vasiliki Plerou, Xavier Gabaix and H. Eugene Stanley "Statistical Properties of Share Volume Traded in Financial Markets," Physical Review E, 62 (2000), R4493-R4496.

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*

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

Haug and Taleb (2008) Why We Never Used the Black Scholes Mertn Option Pricing Formula, *Wilmott*, in press.

Heston, Steven L., 1993, A closed-form solution for options with stochastic volatility, with application to bond and currency options, Review of Financial Studies 6, 327–343. Hull, 1985

Kahl, Christian, and Peter Jackel, 2005, Not-so-complex logarithms in the Heston model, Wilmott Magazine pp. 94–103.

Lee, Roger W., 2001, Implied and local volatilities under stochastic volatility, International Journal of Theoretical and Applied Finance 4, 45–89.

Lee, Roger W. , 2004, The moment formula for implied volatility at extreme strikes, Mathematical Finance 14, 469–480.

Lee, Roger W., 2005, Implied volatility: Statics, dynamics, and probabilistic interpretation, in R. Baeza-Yates, J. Glaz, Henryk Gzyl, Jurgen Husler, and Jose Luis Palacios, ed.: Recent Advances in Applied Probability (Springer Verlag).

Lewis, Alan L., 2000, Option Valuation under Stochastic Volatility with Mathematica Code (Finance Press: Newport Beach, CA).

Mandelbrot, B. and N. N. Taleb (2008): "Mild vs. Wild Randomness: Focusing on Risks that Matter." Forthcoming in Frank Diebold, Neil Doherty, and Richard Herring, eds., The Known, the Unknown and the Unknowable in Financial Institutions. Princeton, N.J.: Princeton University Press.

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. (2001a): Quantitative Finance, 1, 113–123

Mandelbrot, B. (2001b): Quantitative Finance, 1, 124–130

Mandelbrot, 1997 Mandelbot, 2004a, 2004b, 2004c

Markowitz, Harry, 1952, Portfolio Selection, Journal of Finance 7: 77-91

Merton R. C. (1973): "Theory of Rational Option Pricing," Bell Journal of Economics and Management Science, 4, 141–183.

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

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

O'Connell, Martin, P. (2001): The Business of Options , New York: John Wiley & Sons.

Perline, R. "Strong, 2005, Wealk, and False Inverse Power Laws", Statistical Science, 20-1,February.

Plerou, Vasiliki, Parameswaran Gopikrishnan, Xavier Gabaix, Luis Amaral and H. Eugene Stanley, "Price Fluctuations, Market Activity, and Trading Volume," Quantitative Finance, 1 (2001), 262-269.

Rubinstein M. (2006): A History of The Theory of Investments. New York: John Wiley & Sons.

Sprenkle, C. (1961): "Warrant Prices as Indicators of Expectations and Preferences," Yale Economics Essays, 1(2), 178–231.

Stanley, H.E. , L.A.N. Amaral, P. Gopikrishnan, and V. Plerou, 2000, "Scale Invariance and Universality of Economic Fluctuations", Physica A, 283,31–41

Taleb, N. N, , 1996, Dynamic Hedging: Managing Vanilla and Exotic Options (John Wiley & Sons, Inc.: New York).

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

Taleb, N. N. (2007): Statistics and rare events, *The American Statistician*, August 2007, Vol. 61, No. 3

Thorp, E. O. (1969): "Optimal Gambling Systems for Favorable Games", Review of the International Statistics Institute, 37(3).

Thorp, E. O., and S. T. Kassouf (1967): Beat the Market. New York: Random House

Thorp, E. O. (2002): "What I Knew and When I Knew It—Part 1, Part 2, Part 3," Wilmott Magazine, Sep-02, Dec-02, Jan-03

Thorp, E. O. (2007): "Edward Thorp on Gambling and Trading," in Haug (2007)

Thorp, E. O., and S. T. Kassouf (1967): *Beat the Market*. New York: Random House

Weron, R., 2001, "Levy-Stable Distributions Revisited:Tail Index > 2 Does Not Exclude the Levy-Stable Regime." International Journal of Modern Physics12(2): 209–223

Wilmott, Paul, 2000, *Paul Wilmott on Quantitative Finance* (John Wiley & Sons: Chichester).