No, Small Probabilities Are Not "Attractive to Sell": A Comment

N. N. Taleb

January 2013

ABSTRACT Owing to the convexity of the payoff of out -of-the money options, an extremely small probability of a large deviation unseen in past data justifies rationally buying them, or at least justifies excessive caution in not being exposed to them, particularly those options that are extremely nonlinear in response to market movement. On needs, for example, a minimum of 2000 years of stock market data to assert that some tail options are "expensive". The paper presents errors in Ilmanen (2012), which provides an exhaustive list of all arguments in favor of selling insurance on small probability events. The paper goes beyond Ilmanen (2012) and suggests an approach to analyze the payoff and risks of options based on the nonlinearities in the tails.

In answering the question posed by his recent article[1] (September/October 2012), Antti Ilmanen concluded—seemingly backed by a great deal of "empirical" examination and citing a large number of studies—that investors should not merely be uninsured but should also consider selling such insurance. Selling volatility on the left tail "adds value in the long term." He also included carry trades because they imply tail-selling risk insurance.

Perhaps Ilmanen cited too many papers and arguments for comfort. Just as in a complicated detective novel in which the character with the most alibis often turns out to be the murderer, the enumeration of "backup" arguments fails to mask a central methodological error: a combination of (1) cherry picking and (2) missing nonlinear effects and asymmetries in errors (deviations from the model result in considerably more harm when one is wrong than when one is right). Merely adding these nonlinear responses to tail events does more than reverse the result. Further, because Ilmanen included a review of all supporting arguments against the purchase of small-probability events, refuting his article allows the refutation of the prevailing arguments that posit the overpricing of small odds in finance. So, it turns out that there is not a single study that convincingly demonstrates the overpricing of small probabilities in finance or economics.

There are two elephants in the room in the form of exclusion of central (i.e., nonlinear) evidence:

First elephant. Ilmanen excluded the stock market crash of 1987 from his analysis. But because of the convexity of option payoffs, the return from such crashes is convex to distance from moneyness. So, to use a very extreme (but illustrative) case, an option located 20 standard deviations from the money would return 230,000 times its daily premium erosion in the event of a 20-standard-deviation move (i.e., standard deviations from the implied volatility at which the option was purchased). Hence, one would need more than 2,000 years of data showing an absence of 1987-style crashes—generously assuming that the environment is stable—to pronounce the sale of these options "safe." Even those options that are closer to the money (and commonly traded) deliver large enough a payoff to forbid us from making claims from few decades worth of data; for instance an option 12 standard deviations away from the money return 5,000 the daily erosion.

Another misunderstanding concerns the path dependence of these payoffs, which compounds the payoff asymmetry. When the implied volatility quadruples, a 15-standard-deviation out-of-the-money option becomes a 4-standard-deviation option and its value is multiplied by 3,300. Implied volatility (as represented by various volatility indices, such as the Chicago Board Options Exchange Volatility Index, or VIX) quadrupled at least six times over the past quarter century. Table 1 shows the convexity of options to changes in implied volatility. These changes in implied volatility induce a second layer of optionality that is missing from Ilmanen's analysis—with opportunities for the option owner and an squeeze for the seller. (In a well-publicized debacle, the speculator Victor Niederhoffer went bust because of explosive changes in implied volatility in his option portfolio, not because of market movement; moreover, the options that bankrupted his fund ended up expiring worthless weeks

[1] Ilmanen, Antti, 2012, "Do Financial Markets Reward Buying or Selling Insurance and Lottery Tickets?" *Financial Analysts Journal*, September/October, Vol. 68, No. 5 : 26 - 36. Abstract of Ilmanen's article: "Selling financial investments with insurance or lottery characteristics should earn positive long-run premiums if investors like positive skewness enough to overpay for these characteristics. The empirical evidence is unambiguous: Selling insurance and selling lottery tickets have delivered positive long-run rewards in a wide range of investment contexts. Conversely, buying financial catastrophe insurance and holding speculative lottery-like investments have delivered poor long-run rewards. Thus, bearing small risks is often well rewarded, bearing large risks not".

later. The same thing happened with Long-Term Capital Management in 1998.)

Table 1 *The effect of an explosion of implied volatility on the pricing of options, expressed in a multiplier of original premium -The at-the money option is linear to volatility, while out-of-the-money options are increasingly convex.* [2]

	Volatility doubles	Volatility triples	Volatility quadruples
ATM	2	3	4
5 σ OTM	5	10	16
10 σ OTM	27	79	143
15 σ OTM	302	1,486	3,298
20 σ OTM	7,686	72,741	208,429

Second elephant. Ilmanen discussed "carry trades" but ignored the disastrous effect of bank loans (small-probability selling) in the 2008 debacle (as well as those bank loans during the 1982 and 1991 credit problems); he even cited a 2004 paper of mine that includes bank loans as a domain of tail selling.[3] The losses of 2008, estimated by the International Monetary Fund to be more than $5 trillion (before the government transfers and bailouts), would offset every single gain from tail selling in the history of economics.

Excluding the crash of 1987 and bank loans would be like claiming that the 20th century was extremely peaceful by excluding World Wars I and II. These two fallacies alone would be devastating for the entire idea. But let us examine additional errors related to a misunderstanding of nonlinearities.

Convexity bias. Ilmanen made the severe error of ignoring the effect of Jensen's inequality on the nonlinearity of the difference between the VIX and delivered volatility. The VIX, by design, delivers a payoff that is closer to the variance swap. Let's say that the VIX is "bought" at 10%—that is, the component options are purchased at a combination of volatilities that corresponds to a VIX at that level. Because of nonlinearity, it could benefit from an episode of 4% volatility followed by an episode of 15%, for an average of 9.5%; Ilmanen seemed to treat this 0.5 percentage point gap as a loss.

Misuse of the VIX. Using the VIX to gauge small probabilities is inappropriate. The VIX is not quite representative of the "tails"; its value is dominated by at-the-money options, and the fact that at-the-money options can be expensive has no bearing on the argument because we are concerned with the tails. When betting on fat-tailedness, I used to sell at-the-money options because we can safely say—in agreement with Ilmanen—that, owing to their linearity, they are patently expensive, and such a statement is robust to the first elephant.

Ludic fallacy. Real life has little to do with lottery tickets where the probabilities and maximum payoff are generally known. Ilmanen noted the phenomenon called "long-shot bias" while citing papers on bounded payoffs and binary payoffs in unrelated domains (what I call the "ludic fallacy"). These packages, discussed in several papers[4] cited in the Ilmanen article, are not sensitive to fat tails (there are no true exposures to explosive tail payoffs); I have written a brief note on the problem.[5] Ilmanen also conflated long volatility trading (a more or less convex strategy) with investment in high- or low-volatility stocks.

Finally, linking all these errors is a misunderstanding of the effect of the severe nonlinearity of the payoff of out-of-the-money options on inference and decisions. We check people getting on airplanes without "evidence" that they are terrorists simply because the consequence of letting terrorists board planes would be monstrous; likewise, there are some inferential mistakes that people are unwilling to make. Ilmanen failed to understand that in the tails, the difference between absence of evidence and evidence of absence is compounded. Alas, such arguments—based on supernaive inference from the past, not on assessment of fragility—led banks to blow up in 2008: They had "empirical evidence" that their payoffs were "safe."

Nassim N. Taleb

Polytechnic Institute of New York University and Universa Investments

[2] N.N. Taleb and R. Douady " Mathematical Definition and Mapping of (Anti)Fragility" forthcoming *Quantitative Finance,* shows how one can detect the probability of extreme loss from the convexity of the payoff to changes in standard deviation.

[3] N.N. Taleb, "Bleed or Blowup? Why Do We Prefer Asymmetric Payoffs?" *Journal of Behavioral Finance*, vol. 5, no. 1 (2004):2–7.

[4] For example, Joseph Golec and Maurry Tamarkin, "Bettors Love Skewness, Not Risk, at the Horse Track," *Journal of Political Economy*, vol. 106, no. 1 (February 1998):205–225; Erik Snowberg and Justin Wolfers, "Explaining the Favorite–Long Shot Bias: Is It Risk-Love or Misperceptions?" *Journal of Political Economy*, vol. 118, no. 4 (August 2010):723–746.

[5] N.N. Taleb, "A Short Note to Explain Why 'Prediction Markets' & Game Setups Have Little to Do with Real-World Exposures,", *in Metaprobability, Convexity, and Heuristics, the Technical Companion for the Incerto*, electronic book, www.fooledbyrandomness.com .

No, Small Probabilities Are Not “Attractive to Sell”: A Comment

ADDITIONAL NOTES & DERIVATIONS

A common error is made in the interpretation of time series by confusing an estimated value for a “true” parameter, without establishing the link between estimator from the realization of a statistical process and the estimated value. Take M the true mean of a distribution in the statistical sense, with M_N^* the observed arithmetic mean of the realizations *N*.

Note that scientific claims can only be made off M, not M*. Limiting the discussion to M* is mere journalistic reporting. Mistaking M* for M is what the author calls the “Pinker mistake” of calling journalistic claims “empirical evidence”, as science and scientific explanations are about theories and interpretation of claims that hold outside a sample, not discussions that limit to the sample.

Assume we satisfy standard statistical convergence where our estimator approaches the true mean (convergence in probability, that is, $\text{limit}_{N\to\infty}$ P[$|M_N^*-M|>\epsilon$]=0). The problem is that for a finite M, when all realizations of the process when the distribution is skewed-left, say bounded by 0 on one side, and is monomodal, E[M_N^*] > M. Simply, returns in the far left tail are less likely to show up in the distribution, while these have large contributions to the first moment.

A simple way to see the point: the study assumes that one can derive strong conclusions from a single historical path not taking into account sensitivity to counterfactuals and completeness of sampling. It assumes that what one sees from a time series is the entire story.

Figure 1: The Small Sample Effect and Naive Empiricism: When one looks at historical returns that are skewed to the left, most missing observations are in the left tails, causing an overestimation of the mean. The more skewed the payoff, and the thicker the left tail, the worst the gap between observed and true mean.

Now of concern for us is assessing the stub, or tail bias, that is, the difference between M and M*, or the potential contribution of tail events not seen in the window used for the analysis. When the payoff in the tails is powerful from convex responses, the stub becomes extremely large. So the rest of this note will go beyond the Ilmanen (2012) to explain the convexities of the payoffs in the tails and generalize to classical mistakes of testing strategies with explosive tail exposures on a finite simple historical sample. It will be based on the idea of metaprobability (or metamodel): by looking at effects of errors in models and representations. All one needs is an argument for a *very* small probability of a large payoff in the tail (devastating for the option seller) to reverse long shot arguments and make it uneconomic to sell a tail option. All it takes is a small model error to reverse the argument.

The Nonlineatities of Option Packages

There is a compounding effect of rarety of tail events and highly convex payoff when they happen, a convexity that is generally missed in the literature. To illustrate the point, we construct a "return on theta" (or return on time-decay) metric for a delta-neutral package of an option, seen at t_0 o given a deviation of magnitude $N\,\sigma_K$.

$$\Pi(N, K) \equiv \frac{1}{\theta_{S_0,t_0},\ \delta}\left(O\left(S_0\, e^{N\,\sigma_K\sqrt{\delta}},\ K,\ T-t_0,\ \sigma_K\right) - O\left(S_0,\ K,\ T-t_0-\delta,\ \sigma_K\right) - \Delta_{S_0,t_0}\left(1-S_0\right) e^{N\,\sigma_K\sqrt{\delta}}\right) \qquad (1)$$

Where $O(S_0, K, T-t_0-\delta, \sigma_K)$ is the European option price valued at time t_0 off an initial asset value S_0, with a strike price K, a final expiration at time T, and priced using an “implied” standard deviation σ_K. The payoff of Π is the same whether O is a put or a call, owing to the delta-neutrality by hegding using a hedge ratio Δ_{S_0,t_0} (thanks to put-call parity, Δ_{S_0,t_0} is negative if O is a call and positive otherwise).

Θ_{S_0,t_0} is the discrete change in value of the option over a time increment δ (changes of value for an option in the absence of changes in any other variable). With the increment $\delta = \frac{1}{252}$, this would be a single business day. We assumed interest rate are 0, with no loss of generality (it would be equivalent of expressing the problem under a risk-neutral measure). What Eq (1) did is re-express the Fokker-Plank-Kolmogorov differential equation (Black Scholes), $\frac{\partial f}{\partial t} = -\frac{1}{2} S^2 \sigma^2 \frac{\partial^2 f}{\partial S^2}$ in discrete terms, away from the limit of $\delta \to 0$. In the standard Black-Scholes World, the expectation of $\Pi(N,K)$ should be zero, as N follows a Gaussian distribution with mean $-\frac{1}{2}\sigma^2$. But we are not about the Black Scholes world and we need to examine payoffs to potential distributions. The use of σ_K neutralizes the effect of "expensive" for the option as we will be using a multiple of σ_K as N standard deviations; if the option is priced at 15.87% volatility, then one standard deviation would correspond to a move of about 1%, Exp[$\sqrt{\frac{1}{252}}$. 1587].

Clearly, for all K, Π[0,K]=-1 , Π[$\sqrt{\frac{2}{\pi}}$,K]= 0 close to expiration (the break-even of the option without time premium, or when $T - t_0 = \delta$, takes place one mean deviation away), and Π[1,K]= 0.

Convexity and Explosive Payoffs

Of concern to us is the explosive nonlinearity in the tails. Let us examine the payoff of Π across many values of K= S_0 Exp[$\Lambda\ \sigma_K \sqrt{\delta}$], in other words how many "sigmas" away from the money the strike is positioned. A package about 20 σ out of the money , that is, Λ=20, the crash of 1987 would have returned 229,000 days of decay, compensating for > 900 years of wasting premium waiting for the result. An equivalent reasoning could be made for subprime loans. From this we can assert that we need a minimum of 900 years of data to start pronouncing these options 20 standard deviations out-of-the money "expensive", in order to match the frequency that would deliver a payoff, and, more than 2000 years of data to make conservative claims. Clearly as we can see with Λ=0, the payoff is so linear that there is no hidden tail effect.

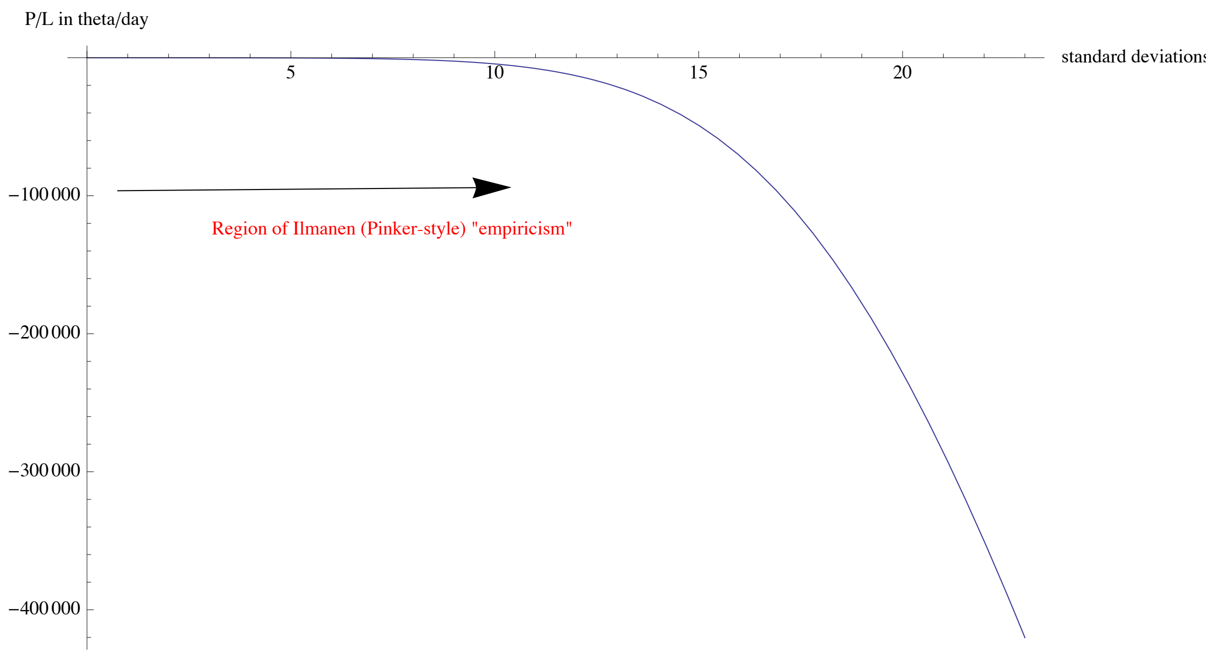

Figure 2: The extreme convexity of an extremely out of the money option, with Λ=20. We show how the effect is totally undetectable by looking at the "regular". The solution is to perturbate the standard deviation as per the Taleb-Douady transfer method (illustrated in Table 1)

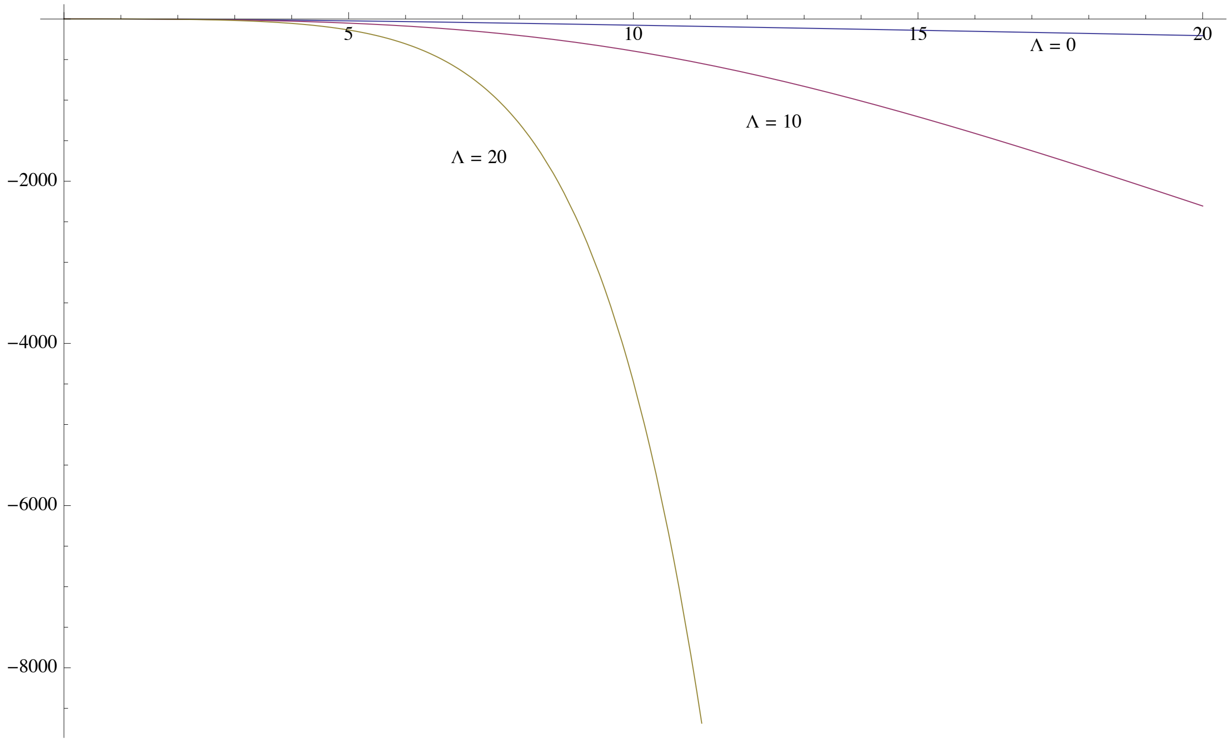

Figure 3: Different Levels of Convexity. Returns for package *Π(N,K=* S_0Exp[Λ σ_K]) at values of Λ= 0,10,20 and N, the conditional "sigma" deviations.

Visibly the convexity is compounded by the fat-tailedness of the process: intuitively a convex transformation of a fat-tailed process, say a powerlaw, produces a powerlaw of considerably fatter tails. The Variance swap for instance results in $\frac{1}{2}$the tail exponent of the distribution of the underlying security, so it would have infinite variance with tail $\frac{3}{2}$ off the "cubic" exonent discussed in the literature (Gabaix et al,2003; Stanley et al, 2000) -and some out-of-the money options are more convex than variance swaps, producing tail equivalent of up to $\frac{1}{5}$ over a broad range of fluctuations.

For specific options there may not be an exact convex transformation. But we can get a Monte Carlo simulation illustrating the shape of the distribution and visually showing how sjewed it is.

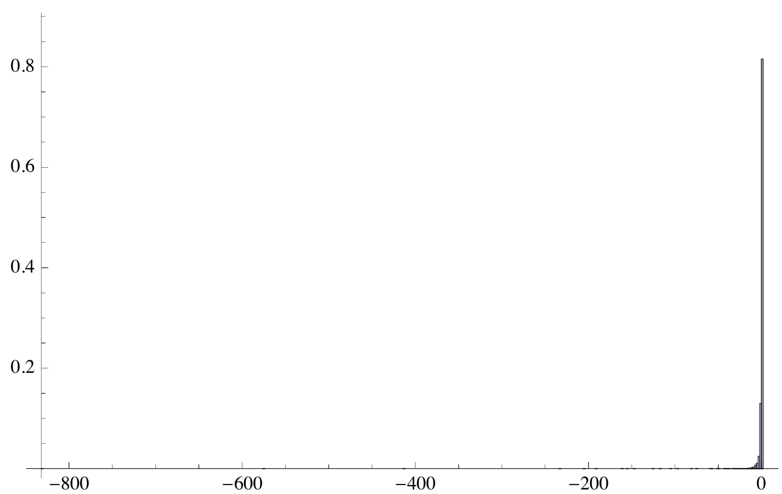

Figure 4: In probability space. Histogram of the distribution of the returns Λ=10 using powerlaw returns for underlying distribution with α tail exponent =3.

Footnote 1: This convexity effect can be mitigated by some dynamic hedges, assuming no gaps but, because of "local time" for stochastic processes, in fact, some smaller deviations can carry the cost of larger ones: for a move of -10 sigmas followed by an upmove of 5 sigmas revision can end up costing a lot more than a mere -5 sigmas. Tail events can come from a volatile sample path snapping back and forth.

Fragility Heuristic and Nonlinear Exposure to Implied Volatility

Most of the losses from option portfolios tend to take place from the explosion of implied volatility, therefore acting as if the market had already experienced a tail event (say in 2008). The same result as Figure 3 can be seen for changes in implied volatility: an explosion of volatility by 5 × results in a 10 σ option gaining 270 × (the VIX had at least five episodes of explosive rises in excess of x 5 since 1987). (In a well publicized debacle, the speculator Niederhoffer went bust because of explosive changes in implied volatility in his option portfolio, not from market movement; further, the options that bankrupted his fund ended up expiring worthless weeks later).

The Taleb and Douady (2012), Taleb Canetti et al (2012) fragility heuristic identifies convexity to significant parameters as a metric to assess fragility to model error or representation: by theorem, model error maps directly to nonlinearity of parameters. The heuristic corresponds to the perturbation of a parameter, say the scale of a probability distribution and looks at the effect of the expected shortfall; the same theorem asserts that the asymmetry between gain and losses (convexity) maps directly to the exposure to model error and to fragility. The exercise allows us to re-express the idea of convexity of payoff by ranking effects.

	× 2	× 3	× 4
ATM	2	3	4
Λ=5	5	10	16
Λ=10	27	79	143
Λ=15	302	1486	3298
Λ=20	7686	72 741	208 429

Table 1 It shows differents results (in terms of multiples of option premia over intrinsic value) by multiplying implied volatility by 2, 3,4. An option 5 conditional standard deviations out of the money gains 16 times its value when implied volatility is multiplied by 4. Further out of the money options gain exponentially. Note the linearity of at-the-money options

Conclusion: The Asymmetry in Decision Making

To assert overpricing (or refute underpricing) of tail events expressed by convex instruments requires an extraordinary amount of “evidence”, a much longer time series about the process and strong assumptions about temporal homogeneity. Out of the money options are so convex to events that a single crash (say every 50, 100, 200, even 900 years) could be sufficient to justify skepticism about selling *some* of them (or avoiding to sell them) --those whose convexity matches the frequency of the rare event. The further out in the tails, the less claims one can make about their "value", state of being "expensive', etc. One can make claims on "bounded" variables perhaps, not for the tails.

Acknowledgments

Bruno Dupire, Xavier Gabaix, Mark Spitznagel, Michael Mauboussin, Phil Maymin, Aaron Brown, Brandon Yarkin, Yechezkel Zilber