By the Numbers: 10 things my hobbies[1] have taught me about investing

Vineer Bhansali

JOIM Conference, September 2014

Data Science, Institutional & Personal Investing

(Version September 29, 2014)

Abstract

I discuss ten common themes between non-investment related activities that shed practical and useful light on investing. While readers might not be familiar with these particular activities, I believe that combining analogies from any accumulated skill in intrinsically rewarding activities (also known as hobbies), with a disciplined analytical approach yields significant benefits.

When presenting a paper at an esteemed venue like JOIM (this is my third time), I have generally played cautious. My presentations have been either about (1) rigorous mathematical models applicable to finance, or (2) the application of financial structuring technology to solving problems we as investors face (For example, at the JOIM in 2009, right in the middle of the financial crisis, I spoke about Tail Risk Hedging). When I have acted as a discussant, it also generally has been about relatively technical issues (such as the one earlier this year discussing the brilliant work by Marty Leibowitz on yield convergence strategies).

But today's topic is "Data Science, Institutional and Personal Investing". My initial inclination was to stick to the quant framework again, and report some very interesting findings on the application of high frequency data (at 100 millisecond frequency) to low frequency phenomena. In particular a discussion of how high frequency data can be used to expand rare event data sets many fold. This analysis allows us to say useful things about correlations between assets in real-time; and anticipate identification of stress points and techniques to protect against them. I found some interesting statistical results and mathematical tricks to correct for things like the Epps effect (which result in correlations going to zero at high sampling frequency). But the more I thought about the context of today's conference, the more I started to fixate towards the "personal" investing part. So I decided to put the mathematics and quant modeling aside and talk about insights that I have gained about investing over the last 25 years from my personal hobbies. To go "beyond the quant model", while sticking to "the numbers".

[1] According to Wikipedia a hobby " is a regular activity that is done for pleasure, typically during one's leisure time. By continually participating in a particular hobby, one can acquire substantial skill and knowledge in that area. Generally speaking, a person who engages in an activity solely for fun is called a hobbyist, whereas 'professional' generally engages in an activity for reward…"

The reason this is relevant is the fact that most of us do not get a chance to invest using all the power of mathematics we have at our disposal – due to both constraints and behavioral responses we end up relying on rules of thumb and "gut feelings" for our investing. While they are not completely non-numeric, having a few operational rules of thumb that are sound makes the decision making process easier.[2] There are also practical constraints such as limited data and information, limited capital, limited access to markets, and relatively high transactions costs. So the question I asked myself was this – are the things that I do for fun, and intuitively, and which have no direct financial reward good labs for teaching robust analogies that can be used in investing; in other words do the things that I do because of the intrinsic value of the activity itself, tell me something about how to approach active investing? In other words can the things that make us better at other intrinsic activities make us better investors?

Rather than being descriptive, the discussion today will be more of an "how to", i.e. you should think of it more as a plumber, carpenter or electrician describing what works in an activity rather than a professor describing how to optimally to approach a problem when all the facts are present.

Of course, at this stage you are wondering "should I stay or should I go"! So to induce you to stay, I will ask the collective you to write down whether you have ever relied on analogies from your own hobby like activities for your own investing, and if you believe that they add any value.

I will use examples from five different activities that are not directly related to investing to distill some principles. I do not have any of the traditional hobbies (golf, coin, stamp, antique book, map, wine collection). To establish my credibility in these activities (which I suspect I have to do), let me list them and my highest qualification in each of them. My first hobby is ultra-running. An ultra-running race is anything more than the 26.2 mile marathon race. I have completed 28 of them, including the 100 mile Western States Endurance run five times. In June of this year I broke the sub-24 hour barrier to earn the coveted "silver" buckle. My second hobby is flying. I have been flying since I received my first pay check 27 years ago, and have single, multi-engine, floatplane, instrument and commercial pilot ratings. I have also been flying jets for the last five years at the level of an ATP (airline transport pilot), as a hobby. My third hobby, which was once almost a profession, is theoretical particle physics. I received my Ph.D. from Harvard in 1992 before accidentally entering finance. This hobby is now supported by stimulating conversations with my ex-collaborators who are professors (one of who takes joy in "curling" as a hobby). My fourth hobby is screenwriting (somewhat dormant right now). I forced myself to go to night school to learn screenplay writing so I could convert some video footage into a real documentary (still pending). Finally, my last hobby is to write programs, and this hobby has made my job as a quant enormously more productive and "fun". I find that I am most focused in my research when I can convert ideas into code. Over the last couple of

[2] See "Mental Math for Pilots", for an application to computing while flying.

decades, I have written most of my code in Mathematica, primarily because it's a functional programming language.[3] So with this preamble, here are my top 10 lessons.

1. Focus on Structure (and on the one idea that supports the structure):

We know that the objective of investing is to earn positive returns, and simultaneously reduce risk, if possible. For ultra-runners, the objective is to complete the course, in many cases 100 miles or more, over varied terrain (the Western States run has 18000 feet of gain and 22000 feet of loss in the Sierras), temperatures (vary from freezing to above 110 degrees often), and day and night, in the fastest possible time. For pilots, the objective is to get from point A to point B safely and in comfort. For screenplay writers the objective is to write a movie script that tells a story that resonates with the audience. For physicists, the objective is to make a theory that explains and predicts, while remaining quantifiable and falsifiable. For programming, the objective is to write a program that automates a task in the most efficient manner, i.e. achieve accurate results quickly and with the least amount of overhead.

We can probably see some commonality emerging by succinctly summarizing the objectives.

Structure helps primarily by creating a framework within which the most important and critical objectives can be stated and most efficiently realized; without getting too distracted by features, data and feedback loops that can distract from the objective. Structure also allows for simplicity in understanding the critical features, which allows for repeatability – a set of relatively simple patterns that can be recursively nested to build more complex outcomes out of simpler ones. Since most of the endeavors we discuss are naturally multi-period, having a structure allows us to apply the same systematic principles repeatedly. Finally, a robust structure allows for approaches that are resilient under stress.

For runners, the primary structural element that results in the fastest time over the distance is to focus on *pacing*. Exhibit 1 shows a table that compares the pace sustained over a 100 mile race compared to shorter races. Clearly the speed one can sustain over a longer distance falls monotonically with distance. The objective is to manage the pace such that the average pace over the distance is minimized. This is clearly a non-linear optimization problem. If you start out too fast, you will end up using up too much energy early, and in a finite energy system, this will result in a bigger fade later on. In a 100 mile race it does not matter at all who wins the race to the end of the first mile, or the 50th. The race is to finish first at the 100th mile. The reason for managing pace is really about managing energy; the simple fact is that as you run faster, you utilize more muscle glycogen, and the total amount of energy stored plus new energy that can be

[3] I also tried law school, and music, but did nor persevere in either.

processed (e.g. by taking in gels or sugared drinks) while on the run has an upper limit.[4] Exhibit 2 shows how increasing heart rates are correlated to energy burn. So the structure of pacing is to optimize the energy consumption over the course of the run.

Distance	Time(sec)	Distance in Mi	Speed	Pace (min/mi)
100	9.58	0.06	23.35	2.57
200	19.19	0.12	23.31	2.57
400	43.18	0.25	20.72	2.90
800	101.11	0.50	17.70	3.39
1000	131.96	0.62	16.95	3.54
1500	206	0.93	16.29	3.68
1600	223.13	0.99	16.04	3.74
2000	284.79	1.24	15.71	3.82
3000	440.67	1.86	15.23	3.94
5000	757.35	3.11	14.77	4.06
10000	1577.53	6.21	14.18	4.23
Half	3493.34	13.10	13.50	4.44
Marathon	7415.4	26.20	12.72	4.72
50K Track	10080	31.08	11.10	5.41
100K Track	22200	62.15	10.08	5.95
100 Mile Track	41280	100.00	8.72	6.88

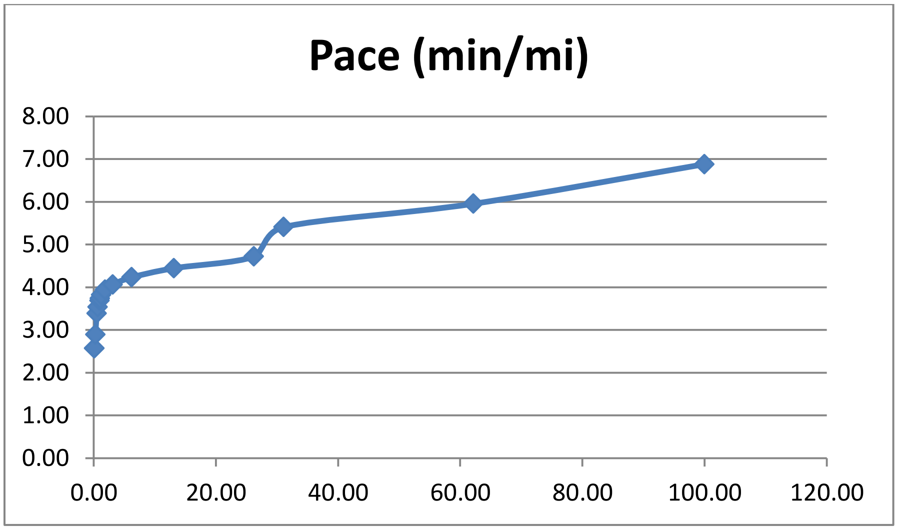


Exhibit 1: Approximate world record pace over different distances (Source: author)

[4] Indeed, in my view even the "Central Governor" model of Noakes is an advanced version of this theory of finite resources, actual or perceived.

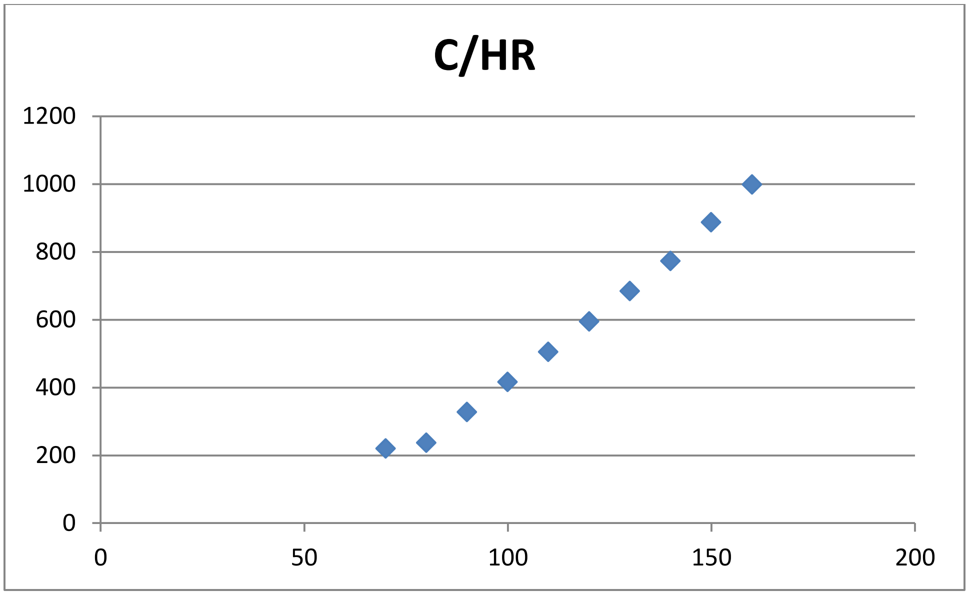


Exhibit 2: Calories/Hour as a function of heart-rate. (Source: author)

When flying an aircraft, the structure of the airfoil determines how lift and drag interact, resulting in a "performance envelope". The core idea is to fly at airspeeds where the drag from the airflow is compensated by the lift produced due to the Bernoulli effect. An immediate consequence of this is that if you slow down too much, to maintain the same lift you have to increase the angle of attack of the wing. There is a critical angle of attack, regardless of other environmental features, that will result in a sharp loss of laminar flow, and which will result in a stall. This leads to an "optimal" airspeed for each airfoil that minimizes total drag (Exhibit 3).

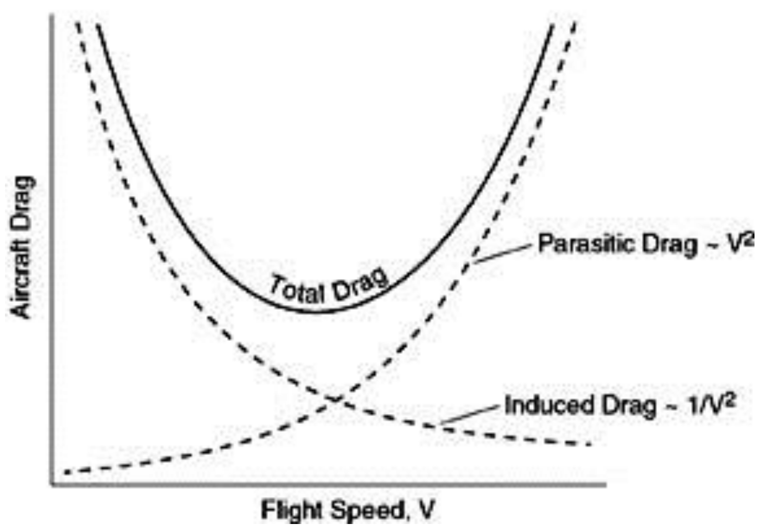


Exhibit 3: Aircraft drag vs. flight velocity. Source: FAA

You might not have realized it, but when you are watching a movie, you are also being gently manipulated by structure. Almost all movies (certainly the ones that appeal to us), are based on screenplays that are in a three-act structure Within the three act structure, there are 5 critical points, and amongst the five critical points there are a few more important moments. The most important structural moment is the 17 minute point. Next time you watch a movie stop it at the 17 minute point. With plus or minus a minute accuracy, that moment will identify what the movie is about. It sets up the core situation that the rest of the movie will work to resolve. Try it! Usually the 17 minute point also tells you what the movie's title is all about. There are also various types of structural symmetries used on you subconsciously by every screenwriter that create familiarity and an intense experience from repetition (plot symmetries, dialogue symmetries, triple repeats, etc.).[5]

The structure of physical theories relies upon the fundamental principles of symmetries and the conservation laws that arise from the symmetries (Noether's theorem). Special relativity arises from Lorentz invariance; General relativity from the conservation of the stress-energy tensor. Even energy conservation is a consequence of time invariance of the Lagrangian of a system. String theory relies on more fancy Lie group symmetries. Symmetries, invariants and conservations laws are the "holy trinity" of physical theories. Pretty much everything in physics flows from the concept of symmetries combined with the "principle of least action", which can be succinctly summarized in the "sum over histories" (or path integral) approach to all of modern physics.[6] Not surprisingly the connection between this approach in physics and the evaluation of path dependent options in finance is deep.[7] And this connection is not only conceptual, but

---

[5] "Why is plot symmetry so satisfying for the audience? Same reason as for triple repeats and dialogue symmetry: by giving the brain a generous dose of carefully designed narrative pattern, a feeling of intense understanding is induced, and as humans we can't get enough of that.Triple repeats, dialogue symmetry and plot symmetry are most definitely the pleasure-inducing drugs of screenplays, and as such they must be used with great care. Use them in a miscalibrated way and your screenplay will die of an overdose." Source: http://www.lavideofilmmaker.com/screenplays/the-red-hot-screenplay-blueprint-10-screenplay-elements-of-films-that-captured-hearts-and-made-serious-money.html

[6] In quantum mechanical terms this states that the probability of a particle to go from point a to point b equals the absolute square of the amplitude $K(b,a)$, i.e. $P(b,a) = |K(b,a)|^2$. $K(b,a) = \sum_{\text{all paths from a to b}} \phi[x(t)]$, and $\phi[x(t)] \propto e^{(i/\hbar)S[x(t)]}$ where $S[x(t)]$ is the "action", i.e. the total of the energy (Lagrangian) under the trajectory marked by $x(t)$. The actual trajectory followed is the one where the action is minimal (see Feynman and Hibbs (1965)). To evaluate the sum over histories we need to evaluate the action over all (infinitely many) paths weighted by their phase, so the sum is replaced by an integral over all paths $K(b,a) = \int_a^b e^{(1/\hbar)S[b,a]} \mathcal{D}x$.

[7] The price of a general, path dependent option is given by the Feynman-Kac formula $\mathcal{O}_F(S,t) = e^{-r\tau} E_{(t,S)}[F[S(t')]] = e^{-r\tau} \int_{-\infty}^{\infty} (\int_{x(t)=x}^{x(T)=X_T} F(e^{x(t')}) e^{-S_{BS}[x(t')]} \mathcal{D}x(t'))\, dx_T$, where $S_{BS}[x(t')] =$

actually leads to better tools (try computing the price of an Asian option, i.e. an option on a geometric average, with and without the use of the path integral technology).

And any programmer will tell you that the structure of a good program depends on the principle of modularity. In other words a program is made out of functions, and if we can structure a program in terms of a coherent set of functions we will get a good, modular program in most cases. There are languages like Mathematica (my favorite), where everything is a function. Exhibit 4 is a one line function in Mathematica that generates a random walk (and by a simple extension a Monte-Carlo simulation, the workhorse of "experimental theoretical finance"):

```
ListLinePlot[Table[Accumulate[Prepend[RandomVariate[NormalDistribution[0,1],100],0]],{5}]]
```

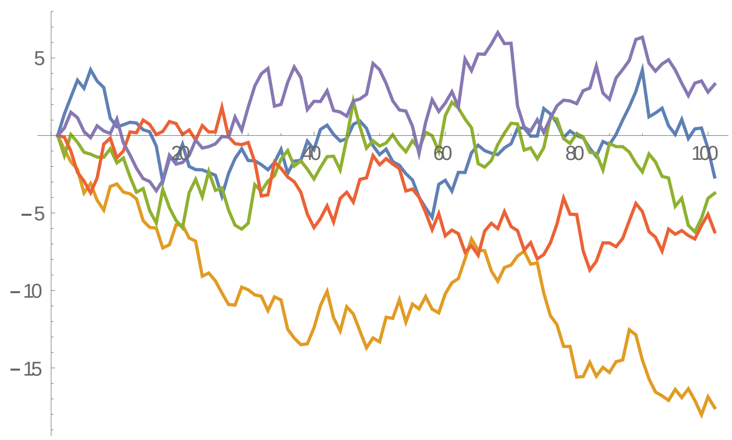


Exhibit 4: A one line functional program in Mathematica. (Source: author)

Or how about this from the latest version that even takes care of all the stochastic calculus for you and creates internally the relevant structures (generate random functions and operate on them analytically)?

$\int_t^T \mathcal{L}_{BS} dt'$is the "action" for Black-Scholes and is given by $\mathcal{L}_{BS} = \frac{1}{2\,\sigma^2}\left(\frac{dx}{dt'} - \mu\right)^2$. With this tool the Black-Scholes option pricing formula can be derived in two lines.

```
ListLinePlot[RandomFunction[GeometricBrownianMotionProcess[μ,σ,100]/.{μ->0.05,σ->0.20},{0,1,1/250},10]]
```

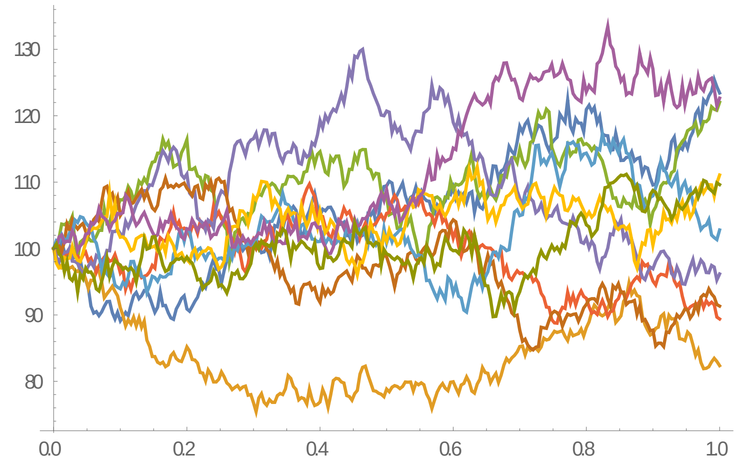


Exhibit 5: A fully functional stochastic simulation (Source: author)

So this brings us to the "structure" that lies behind investing. To me, the structure of investing is dominated primarily by the concepts of "carry" and "trend".[8] If, we as investors can maximize carry per unit of risk, over the long haul we will do well. If can stay out of the market's way (and our own way), by being on the right side of the trend, we will not succumb to generating losses that give away the hard-earned carry (see exhibit 6). We will talk more about trend later, so let us focus here on carry. Whether the carry comes from financing longer maturity bonds with shorter maturity securities, or harvesting dividends, or even selling volatility and options, the accumulation of carry returns over time are hard to beat. Of course, leverage and greed can lead markets into situations where everyone has too much exposure to carry, which can result in shocks shaking out the weak players.[9] But structurally carry is earned due to the sale of an implicit or explicit option to someone else who values it more than you do, and this "risk-premium" is the structural underpinning of most sustainable returns over time.[10]

[8] Please recall that I am discussing "investing" as opposed to financial theory. For the latter, there would be no doubt that the fundamental principle from which everything else emanates is the absence of arbitrage. The absence of arbitrage has the same role in financial theory as the principle of least action and energy conservation in physics.
[9] Which leads to the adage "the road to hell is paved with carry".
[10] See, for example, my book "Bond Portfolio Investing and Risk Management", McGraw-Hill, 2010.

All Asset Classes, **1960-2014**, Annual Excess Returns

	With the Trend	Against the Trend	Totals By Carry Direction
Total (All Positions)	**5.3%**		
Positive Carry	5.1%	0.2%	**5.3%**
Negative Carry	1.9%	-1.9%	**0.0%**
Total by Trend Direction:	**7.0%**	**-1.7%**	
Trend Following (Always With the Trend)			**8.7%**
Carry Trade (Always in the direction of Carry)			**5.3%**

Exhibit 6: Returns to Carry and Trend across asset classes since 1960. Source: PIMCO

In fixed income markets there is a structural compensation for extension of duration. As researchers have found out over and over again, the shape of the yield curve is less a predictor of future yields, and more a predictor of future returns (the bond risk premium is positive). Other examples of such structural advantages persist across other fixed income markets (compensation for prepayment risk, credit risk, tax risk etc.) and even beyond fixed income markets. Thus, to position a portfolio to take optimal advantage of duration positioning in the curve makes structural sense, and it improves the long term expected returns of portfolios.[11] There is also a persistent excess risk-premium in the pricing of options, where the buyer of insurance (who is risk-averse), pays a seller of insurance (who is either less risk-averse or even risk-neutral) premium that can be a structural source of returns.

The wonderful thing about structural advantages across any of the aforementioned activities is that they do not dissipate quickly and can be repeated. Just like a nice stretch of downhill running, a tail-wind for a plane in flight, or a theorem for theoretical physicists, or an easily expandable modular program built on useful functions, structural advantages make returns and risks more predictable and are hence more repeatable.

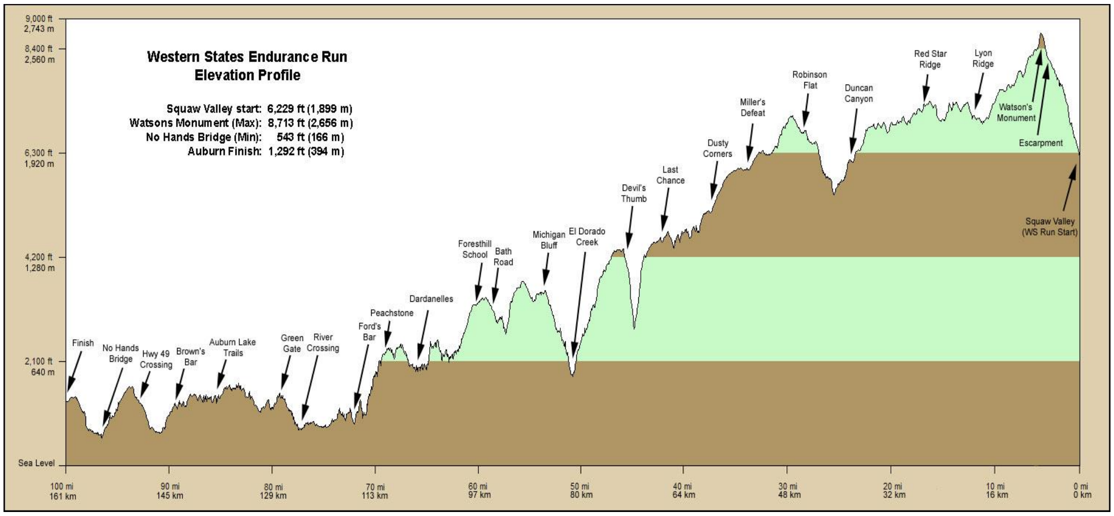
2. Let the data speak (and be Bayesian when you listen to it)

The second key element that is common among all these disciplines is the realization that even despite the best inputs and intentions, despite the best tools and machinery, results in each of them are realized with a high dose of uncertainty. Prediction with certainty is impossible. To quantify the uncertainty, we need to be probabilistic; and to be probabilistic means knowing something about the distribution of outcomes at intermediate steps and using this information to update our knowledge. Data allows us to do this.

[11] See, for example, chapters 9 and 10 of Antti Ilmanen's book "Expected Returns", Wiley, 2011.

In exhibit 7 and 8 I show the distribution of time to arrivals at various intermediate aid stations at the Western States 100[12]. Any experienced veteran of the race will tell you that there are four out of 28 "aid stations" that are more important than the others, since they break up the race into four neat sections. From the start to Robinson Flat (RF) is a 29 mile stretch also called the "high country". From RF to Devil's Thumb (DT) is called the "canyons", which are notorious for quad busting downhills and steep 1000+ climbs over short distances, not to mention the summer heat that is reflected off the canyon walls and can feel like 110 degrees or more. Then the relatively short section from DT to the village of Michigan Bluff, where runners first re-enter civilization (for many it is just getting dark), and where many decide to stop. Finally the "race" starts around here, as they say, for the survivors, with 55 miles done, and about 45 miles to the finish line. When we look at the distribution of times (from the 2009 race where the author also ran for the first time), we find an interesting pattern. The distribution is fairly normal up to Robinson Flat. Runners are still pretty fresh after 30 or so miles, and the non-physical factors are not really in play yet. As you evolve to further checkpoints, the distribution starts to get more left-skewed. At the finish, there is a pronounced double hump right around 1440 minutes (24 hours), and 1800 minutes (30 hrs). The 24 hour hump occurs because runners who finish in under 24 hours receive the coveted "silver" buckle (and recent sports medicine suggests that the brain has ultimate control of the race, and making this cutoff essentially allows tapping into physical resources). Those under 30 get a "bronze" buckle, and those above 30 are not counted as finishers officially. The law of arrival times is clearly not normal (it is a decent power law fit, however). This is not surprising, since in the tails we find universal power law behavior (we will talk more about scaling phenomena later). The data tells you that the physical part only explains part of performance.

[12] Maps and other details are at www.wser.org, the website for the Western States 100.



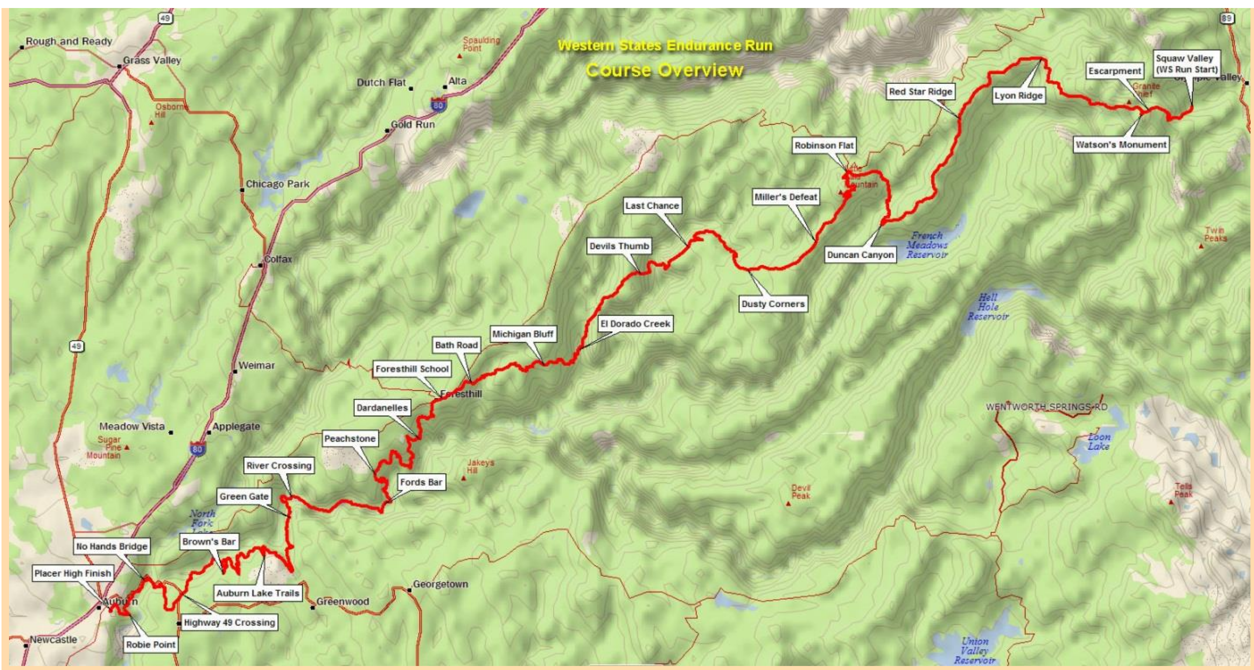


Exhibit 7: Elevation profile and distance chart for the Western States 100 race. (Source: www.wser.org)

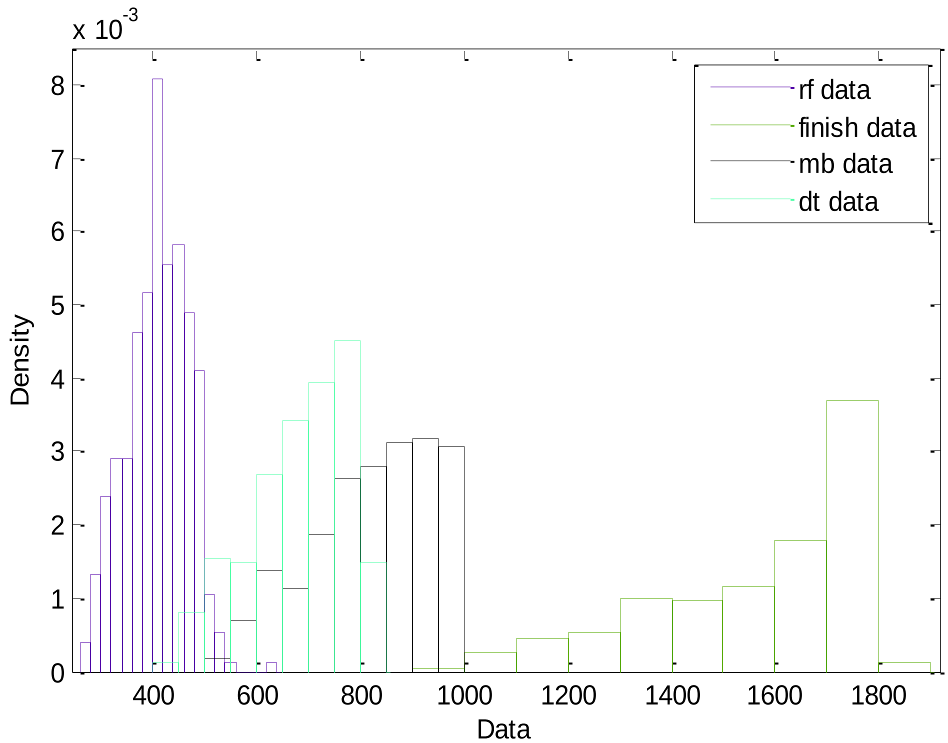


Exhibit 8: Distribution of arrival times at the 2009 Western States 100 race at four major aid stations. (Source: author)

Experiments in the 19th and 20th century made physicists listen. First, there was the absence of the "Aether" that the Michelson and Morley experiment tried to find, and ultimately led to the discovery of special relativity and the constancy of light. Listen to the data. Then there was the discovery that probabilities do not add, quantum amplitudes do (Young's double slit experiment), which called into question the whole idea of particles as granular objects. It would be all too easy to discard the anomaly in the pattern observed as "noise", and in the process of doing so throw away the proverbial baby with the bathwater.

For investing, knowing the distribution of asset returns is central in many different ways. In a piece of research I published last year, I compared the consequences for investing when the underlying distribution is multi-modal as opposed to uni-modal shown below in exhibit 10 (which is the usual assumptions underlying portfolio optimization). The results are quite interesting and intuitive.[13] First, the optimal asset allocation to risky assets falls drastically within the bimodal construct. Second momentum becomes a more important investing risk factor than value. This can be traced to the fact that in a bimodal distribution the most likely habitat for markets is at one of the two tails, and this results in trend dynamics and emergence of momentum as a key factor that you don't want to fight (more on this later). For a unimodal distribution the density is maximum in the middle, so any deviation from the middle meets mean-reversion forces. Finally, option prices and "skews" are underpriced relative to the true distribution. The point is that the underlying probability distribution of asset returns is not stable; in fact we can say very little about the qualitative shape of the distribution with absolute precision; so to be distribution aware is to be aware of the reality of investing.

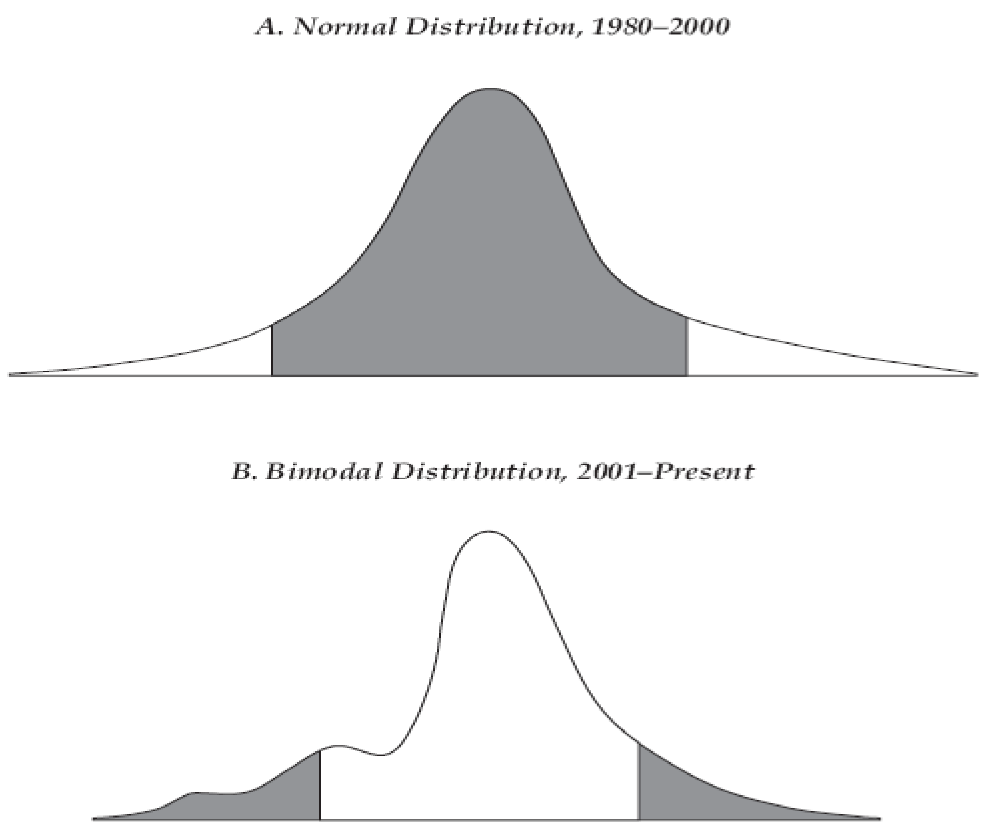


Exhibit 10: Unimodal vs. Bimodal Distributions. (Source: author)

[13] "Asset Allocation and Risk Management in a Bimodal World", FAJ, March 2013, Vol. 30, No. 1:67-78.

An important corollary to knowing your distribution is to be Bayesian. Bayes rule simply says that the posterior probability is proportional to the prior probability times likelihood. In the Western States 100, I knew after the first two of my six tries that for me to finish the race in under 24 hours, I would need to be out of the Devil's thumb checkpoint latest by 3:30 p.m. Now it is entirely possible for one to leave there at 4 p.m. and even later and still come in under 24 hours, but the probability of this joint result is very low. While results from year to year can vary due to temperatures, in 2013 (a "hot" year), 94 people finished in under 24 hours (out of 277 finishers and 383 starters). There were only 2 out of these 94 who left DT after 4 pm (< 2%), and 8 who left DT after 3:30 pm. 31 runners (33%) left after 3 pm and still made it under 24 hours. So if you want to make it to the finish in less than 24 hours, there is a one in a third chance you can do it if you leave DT after 3 pm. Which drops to less than one in ten if you leave after 3:30 pm. But there is a catch: you cannot try to approach this problem as a partial optimization to DT by 3 pm…almost 25 runners who actually left DT at 3 pm or earlier did not finish or finished in more than 24 hours. In Bayesian terms what this means is that hitting each checkpoint in a particular window of time increases the likelihood of finishing in under 24 hours and updates the prior probabilities, but at the same time you have to do this with the ex-ante conditional calculation of being able to finish!

In the investment context, I have found egregious violations of Bayes rule. The simplest version of this is the false estimation of many conditional probabilities as higher than the unconditional probabilities. It is easy to think of arbitrages with options to take advantage of this (conditional trades). One extreme example of this from the European crisis period. I found in many cases that the market's estimation of probabilities of default of various countries individually and jointly did not satisfy the axioms of probability! But there are also notable successes. Approaches that combine priors with views such as the Black-Litterman approach for asset allocation are examples of systematically tilting asset allocation.

3. Use proper coordinates and units

While fundamentally there is no difference in the objective of running a short sprint or an ultramarathon (the objective in both cases is covering the distance in the shortest amount of time), casting the problem in the right units and coordinates make the task easier to execute. When running a sprint, the effort is all about intensity, or in purely physiological terms obtaining the maximum muscle output per second (over 10 seconds in the case of a 100m sprint). In an ultramarathon, the effort is all about sustainability, i.e. the maximum muscle output that can be sustained over the distance and terrain. To understand this, note that a typical ultramarathon is over climbs, flats and descents. If you try to sustain maximum muscle output on the climbs, you will use up a lot more energy per unit time (or per unit effort), than you would use on flats and downhills. Since the total amount of energy expenditure is finite (over a 100 mile race the expenditure is approximately 16000 kCal), it makes sense to optimize over the whole course rather than per unit time. This is a prime reason why shorter distance track stars need a few races to adapt to long, hilly trail races. Thus the right units are not distance covered per unit time, but distance covered per unit effort.

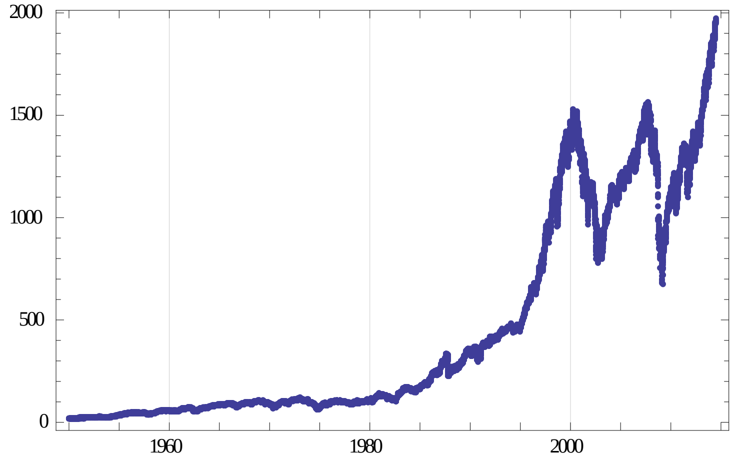
The coordinate system of flat-land (straight lines) has to be modified when we describe locations on curved objects (like latitudes and longitudes). When moving to the description of space-time we simply cannot make sense with flatland coordinates – we need to use curvilinear coordinates and tensors. This observation is much more than the rectilinear to polar coordinate transformation on a plane, where the change of coordinates makes the description more intuitive, but might not add any analytical value; in our example it is actually impossible to describe the geometry without the more sophisticated coordinate system.

We see the relevance of using the right units in investing in a very explicit way given the access to high frequency trading data on many liquid markets. We collected millisecond data on liquid futures contracts on the S&P 500 and other markets. When we look at data at such small intervals, it is not continuous in the sense daily or weekly data might be. There are jumps. As a matter of fact more than half of the clock time the markets do nothing when we look at such high frequency. In the example summarized in Exhibit 11 from October 9, 2008 (during the crisis), 50% of the 100 millisecond time stamps were “dead”. However, at such small time scale the natural units for measurement are not calendar (or clock) time, but trading time (or volume time).[14] When cast in terms of volume buckets or “trading time” (we take equal volume chunks and take the price changes or returns on those chunks), the series of returns looks a lot less jumpy and can actually help us apply low frequency statistical techniques. Clearly the units matter, and in the right units you can connect the high frequency data to lower frequency phenomena. One application of this approach is to use the high frequency data to increase the set of rare event observations. Then one can say statistically more believable things about how rare events in one market influence rare events in other markets. For instance, the increase in correlations between markets during the “Twitter” flash crash (on April 23, 2013, which lasted only about an hour displayed in Exhibit 15) gave us a microscopic view of where common trades were positioned and where they were exposed to a quick bout of correlated delivering. This allows us to build better tail hedging portfolio strategies.[15] By looking in the proper units, we can also see (see the Hidden Markov Model decomposition summarized in Exhibits 12 and 13), that higher frequency data shows very different regimes than lower frequency data.

[14] See Marcos M. Lopez de Prado, “Advances in High Frequency Strategies”, Ph.D. dissertation, Complutense University, 2011.
[15] See Bhansali, V.,”Stress Points: What High Frequency Data Tell us About Hidden Tail Risks”, PIMCO Viewpoints, May 2013.

	Positive	Unchanged	Negative
Raw Returns	30%	43%	27%
Time sampled (one-minute)	40%	4%	56%
Mean Filtered (1 minute window)	41%	3%	56%
Volume Bucketed (50,000 shares)	41%	1.5%	57%

Exhibit 11: Probability of returns of raw data and under different filtering algorithms for SP500 futures on October 9, 2008. (Source: author)



$$\begin{pmatrix} 1.263 \times 10^{-15} \\ 1. \end{pmatrix}$$

a : $\begin{pmatrix} 0.962281 & 0.0377189 \\ 0.0113121 & 0.988688 \end{pmatrix}$

$$\begin{pmatrix} 0.230713 \\ 0.769287 \end{pmatrix}$$

$$\begin{pmatrix} -0.000681488 \\ 0.000586808 \end{pmatrix}$$

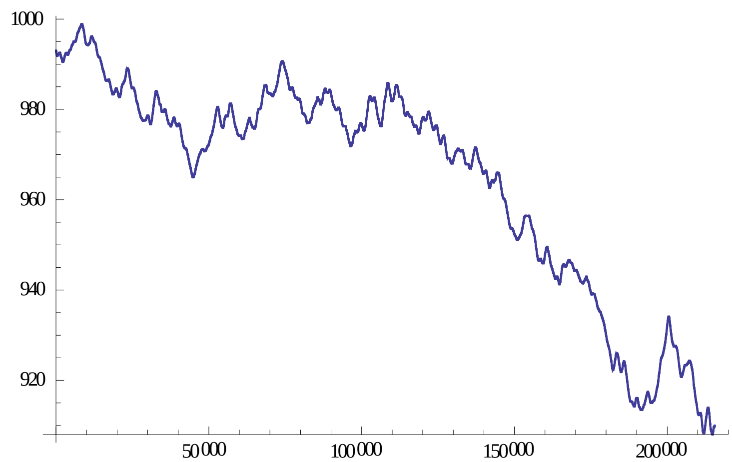
$$\begin{pmatrix} 0.0165848 \\ 0.00635681 \end{pmatrix}$$

X : 54822.1

8.13648×10^{-8}

18

BIC : −109576.



$$\begin{pmatrix} 1.06864 \times 10^{-58} \\ 1. \end{pmatrix}$$

a : $\begin{pmatrix} 0.997453 & 0.00254663 \\ 0.00217277 & 0.997827 \end{pmatrix}$

$$\begin{pmatrix} 0.460392 \\ 0.539608 \end{pmatrix}$$

$$\begin{pmatrix} 2.91593 \times 10^{-6} \\ -3.23379 \times 10^{-6} \end{pmatrix}$$

$$\begin{pmatrix} 2.19542 \times 10^{-6} \\ 2.13761 \times 10^{-6} \end{pmatrix}$$

X : 2.50846×10^{6}

8.79595×10^{-8}

8

BIC : -5.01683×10^{6}

Exhibit 12: Hidden Markov model fits of a two state system on S&P 500 for weekly data from 1952 and 100 millisecond data from October 9, 2008. (Source: author)

[illegible] : {1., 2.0266 10^{19}}

a : {{0.428128, 0.571872}, {0.424558, 0.575442}}

[illegible] : {0.426079, 0.573921}

[illegible] : {0.00590371, 0.00225707}

[illegible] : {0.00225761, 0.00332167}

[illegible] X : 269.023

[illegible] : 8.50524 10^{8}

[illegible] : 23

BIC : 508.51

Exhibit 13: Regime Shifts in Volume Time for October 9, 2008. (Source: author)

Another example of using the proper units is the use of forwards vs. spots. Whether its bond math (and related term structure models), volatility trading (local vs. terminal vol models), it makes a huge difference whether we describe the dynamics of the system in terms of yields or forwards. In finance the whole approach to evaluating in a different "measure", such as the ones used in term structure modeling, are examples of using the proper units in which computation not only becomes simpler, but also intuitive. The very concept of the "risk-neutral" measure makes it possible for us to derive option pricing mathematics. Moving into the "forward" measure allows us to derive bond option pricing formulas as simple extensions of Black-Scholes.

A wonderful advantage of moving to the right system of coordinates is that the underlying problem simplifies into mathematics that probably has already been solved. While there are numerous examples of applications of mathematics from physics to financial problems (for example the path integral formulation), there are also examples of the mathematics of finance and investing enlightening computation in science and other fields.[16]

4. Know the types of errors (and correct for them)

Trial and error is a key part of learning. This is especially critical when the learning is by doing rather than by "reading up" (at least I have not learnt anything ever without making a large number of errors in the process). But it is important to not make random and repeated errors. The first thing to recognize is the type of error we are prone to making. As a reminder, the two "classic" types are Type I (reject true null) and Type 2 (accept false null) errors.

Rejecting the hypothesis in the presence of strong and growing evidence to the contrary is likely to be a Type I error. The classic example of this was the null that a human could not run a mile in under 4 minutes. Of course, Roger Bannister did it and just to show that it was not a physical limitation, a half dozen other runners did the same within a few months. Today, any contender for one of the shorter distance races in the Olympics has either run, or can run a 4 minute mile. The fastest marathon is still over 2 hours. Will this null be proven wrong?

An especially dangerous Type II error is one when a flight instrument does not provide the information it was designed to. For example, a blocked pitot tube would show no airspeed, and in an instrument flying environment, would not provide the necessary input to the pilot to fly within the structural limits of the aircraft. To prevent dependence exclusively on false inputs from one instrument, every pilot is required to know how to "cross-check" against other instruments that can provide secondary information. In the case of the blocked pitot tube, a rapid change in the vertical speed indicator, or the attitude indicator (which provides reference relative to the horizon), communicates information on whether the aircraft is climbing or descending.

[16] See, for example, the application of Hidden Markov Models in astrophysics in a recent paper, "A Formal Method for Identifying Distinct States of Variability in Time-Varying Sources: SGR A* as an Example", L. Meyer, G. Witzel, F.A. Longstaff and A.M. Ghez, the Astrophysical Journal, 791:24, 2014.

This would under normal circumstances be accompanied with declining or rising airspeed. This "partial panel" flying is part of every pilot's training.

In investing, I have observed both Type I and Type II errors. An example of a null is "stocks always go up", or another example is the frequent use of level regressions of related securities to infer causality. A type II error is for a risk management system to fail, for example the use of value at risk as a sole measure of risk being taken. Since a VaR model can obtain the same result with a distinct set of inputs (and risks), and because of variations in the measurement of the inputs such as asset covariances, it is easy to miss risks until it is too late.

An especially dangerous situation can arise when strategies are back-tested for profitability. If the analyst selects a specific strategy from an ensemble of strategies, then Sharpe ratios (and t-stats) for the strategy cannot be trusted without making appropriate adjustments.[17]

Why are the same type of errors so prevalent across different types of activities? I can think of two main reasons: (1) Biases and (2) Excessive dependence on one measurement methodology. The biases themselves can be quantified in many cases. For example, in behavioral finance theories, the biases themselves originate from asymmetries in loss and gain functions as well as in the subjective weighting of probabilities. These can be further traced to systems (e.g. Kahneman's system I and system II)[18]. We spoke already about using VaR or perhaps another system as a risk management tool in isolation.

But this leaves open the question of how to avoid making both Type I and Type II errors. Against the behavioral biases one almost failsafe technique I have been able to use in practical applications is to use good coaches or mentors, and if possible, get paid advice from someone who is held to a higher standard of performance as a professional. Someone who has had experience going through their own trial and errors and been able to use their intellectual and other skills and resources to deliver performance professionally (in whatever area of application), implicitly understand the importance of separating the signal from the noise. Good coaching is never free! In my own experience, however, the trade is usually quite cheap for the mentee in retrospect. And as discussed above in the context of flying safely even with failed instruments, we also see the advantage of redundancy and multiple points of view. Cross-checking results against redundant systems can be used to alleviate mistakes that might be made by excessive reliance on just one measurement apparatus. Finally, and to be discussed more below, simulating possible outcomes in advance of investment can reveal important hidden assumptions.

5. Simplify

In any foot race, there are three key overarching variables that matter: physiological, psychological and environmental. Of course these can be further distilled: physiological factors

[17] See Harvey, C. R. and Y. Liu, "Backtesting, SSRN, October 25, 2013.
[18] D. Kahneman, "Thinking Fast, Thinking Slow".

that matter are the ability to use oxygen (VO2 max), biomechanics etc. However, equally physically talented athletes can have very different performances due to their psychological approach. A race is a highly optimized effort, and those who can use their minds efficiently usually end up performing better. The environment matters a lot too. If it is uphill and windy, the performance will obviously suffer. If it is hot, speeds will go down as the body will not be able to maintain the hot effort for long. So the different combinations of factors can result in differential performance.

Of course, research in physics relies on simplification (everyone has heard the joke about a physicist describing an elephant as a sphere). A large class of complex phenomena can be described by studying a simple model which has spins parallel or anti-parallel, i.e. the Ising model. It explains magnetization, spontaneous symmetry breaking, etc. etc. The understanding of very complex natural phenomena can be achieved by finding a simple model that has all the essential factors of the more complex application and then doing a deep dive into the behavior of the simpler system.

In investing applications, there has been a renaissance in factor modeling, and for a good reason. There are simply too many types of investable assets, and they can each masquerade as something new. But digging deeper into the risks that drive the returns, we find a beautiful simplicity. A handful of factors matter – in fact two factors, the equity beta and duration, which are related to economic growth and inflation, are structurally the only factors that should matter over long horizons, especially for personal investment in liquid assets. Further, and as discussed rigorously in the paper “Where do alphas come from?”, by Andy Lo takes the factor based approach a step further and decomposes expected returns into those arising from harvesting risk premia, factor timing and security selection. This fits in nicely with most investor’s time horizons – the secular investment horizon and the shorter, more cyclical adjustments.

The second aspect of simplification is to minimize waste and frictions. In running this means simply movement without waste. For example, a close tuck, arms back and forth (and not transversally), minimal vertical movement (the latest generation of Garmin running watches can measure this), correct stride length. In short, “good mechanics”. In aviation, efficiency is absolutely critical. Aluminum frames, round, streamlined fuselage, low weight materials are examples of minimizing waste and friction. In physics, the reason why analytical techniques and toy models are so important is because they teach one to think efficiently. When listening to the dialogue in a movie, notice next time the efficiency - a few elements lifted from full conversations are sufficient to get the message and keep moving the story forward. The very design of functional programs is one of simplicity. Each function should do one task, and do it well, so that by a logical nesting any more complex program can be evolved out of the simple programming elements. In a program like Mathematica, even a loop is a function (the “Table” function). In a world of unlimited computational power, having powerful conceptual ideas, free of extraneous details, forces us to look at what is the central point, and build more complex models on the simple models.

Transactions cost minimization plays the same role in finance. Many great investment models that look fantastic on paper fail miserably when applied to real markets. Transactions costs can be minimized in numerous ways: (1) using factor models the economic risks that are being taken can be directed towards the most liquid combinations of securities; (2) the frequency of trading can be optimized so that there is minimal amount of trading required to achieve the same end results; (3) when possible investments should be implemented as a provider of liquidity (and getting paid to do so), than as a receiver of liquidity.

6. Look for scaling rules (and anticipate the possibility of sharp/sudden transitions)

Scaling laws are also sometimes known by the more fancy name of dimensional analysis, but they are more than dimensional analysis (since you can have scaling in dimensionless quantities). For example, a simple rule is that area scales like the square of the length of a side. In many problems, you can almost get the answer by creating a combination of variables on the left and right hand side and solving for unknown exponents until they balance.

Physics is full of scaling rules. In the case of a simple spin spin interaction model, the scaling rule that is relevant is the asymptotic behavior of correlation functions between spins as a function of their separation. It falls off as an exponential of the separation ($e^{-r/\xi}$). However, at the critical temperature ($T = T_c$) the asymptotic behavior is very different and implies long range correlations. Of course, what this means is that at the critical temperature the neighbouring spin interactions talk to each other at infinite separation, and the system exhibits collective behavior at all scales.

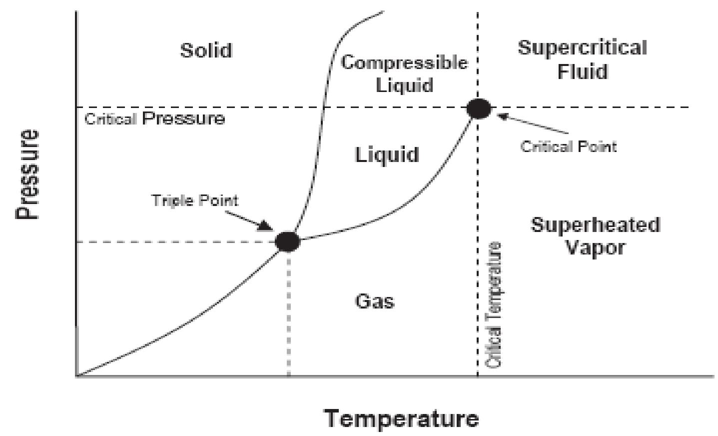
In finance, the most basic rule, the Brownian motion, is that the distance from the starting point scales like the square root of time. Note that try as much as we can, we cannot describe this in the traditional language of speed (distance per unit time), since the scaling is on square root of time. So an analogy from ordinary mechanics would not work for financial time series. A different set of scaling rules are required for financial time series. As we go to lower and lower time scales and higher and higher frequency, we find that the scaling rules do not necessarily act in a Brownian fashion. We can still say intelligent things about exceedances, or about tail behaviors, in terms of scaling laws, however. In many cases these “laws” follow power law behavior. As we have discussed and written about extensively, markets exhibit collective behavior when risks and leverage exceed a critical parameter (see Exhibit 14). When such events happen, uncorrelated markets become correlated, and shocks can propagate through systems at all frequencies and time scales.[19]

A more practical example of this phenomenon that we have studied comes from the sleepy world of municipal bonds.[20] Due to a particular feature of the US tax code which treats the gains from

[19] “Market Crises - Can the physics of Phase Transitions and Symmetry Breaking tell us anything useful?”, JOIM, 3rd quarter, 2009.

[20] A. Ang, V. Bhansali and Y. Xing, “Taxes on Tax-Exempt Bonds”, Journal of Finance, vol. 65(2), pp 565.

the purchase of a municipal bond either as capital gains or income tax (depending on purchase price and time left to maturity), there is a sharp impact on the price of the bond as it approaches this so called “de-minimis” threshold (see Exhibit 16). In today’s low yield environment, what this means is that many low coupon municipal bonds are likely to hit there thresholds if yields were to rise substantially, resulting in a sharp “phase transition” in their prices and liquidity.



Source: PIMCO.

Figure 1 Phase diagram for water.

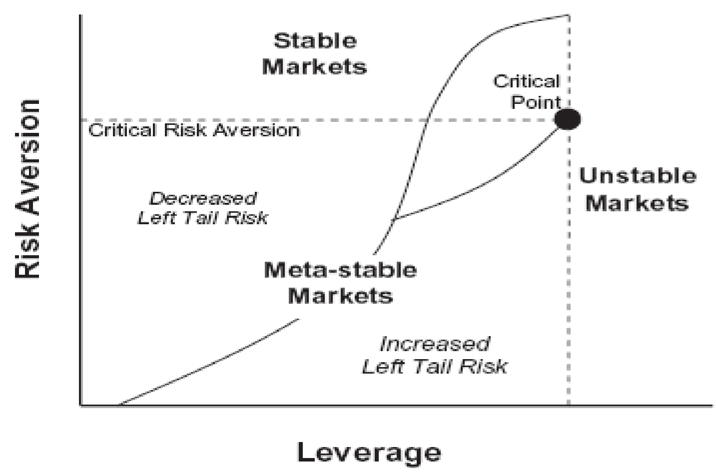


Exhibit 14: Phase transitions in physical systems and markets. (Source: author)

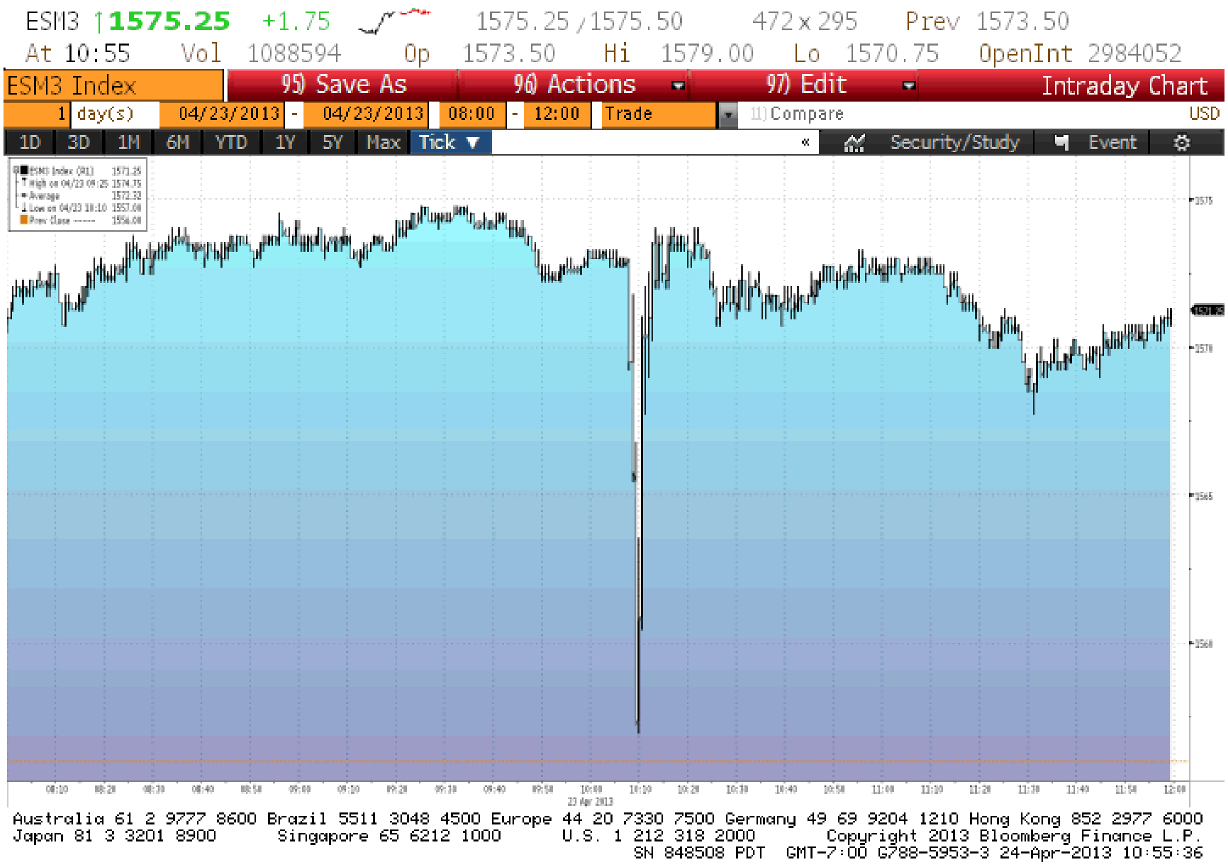


Exhibit 15: S&P 500 prices on April 23, 2013 (the “Twitter” flash crash). (Source: author)

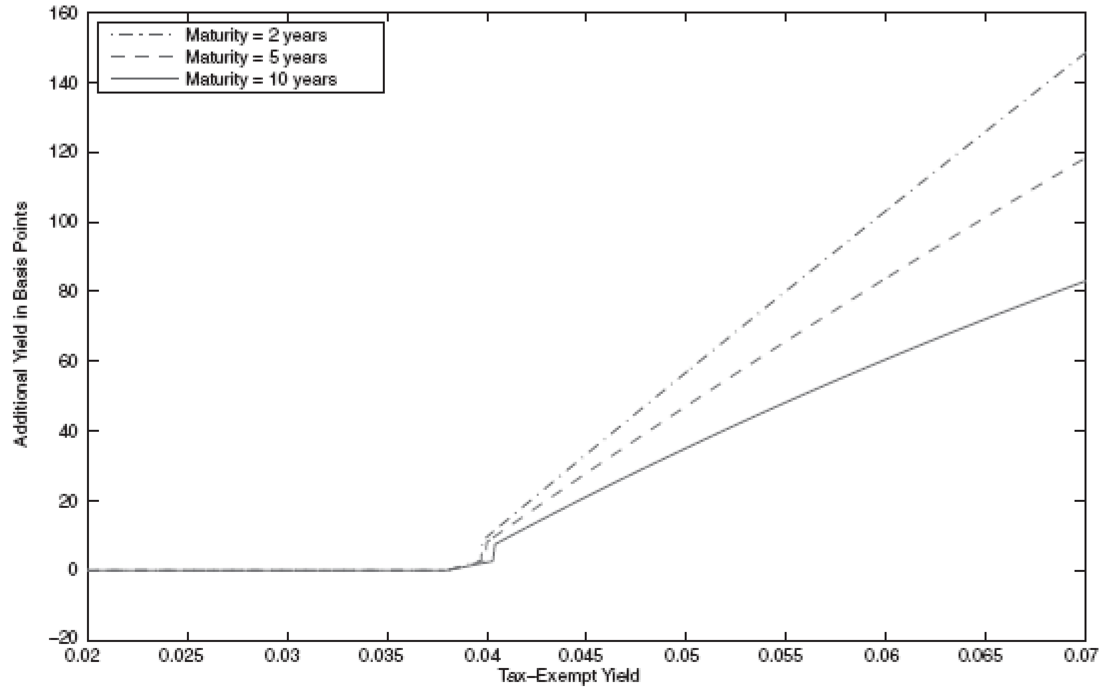


Figure 1. Additional yields required by below de minimis bonds. We consider par bonds of different maturities paying semiannual coupons of 3.8%. For a given tax-exempt yield y (on the x-axis), we compute the additional yield $\bar{y} - y$ (in basis points on the y-axis) above the tax-exempt yield required to obtain the same IRR y on the after-tax cash flows as the tax-exempt yield y. We assume that the income and capital gains tax rates are $\tau_I = 0.35$ and $\tau_C = 0.15$, respectively.

Exhibit 16: S&P 500 prices on April 23, 2013 (the "Twitter" flash crash). (Source: author)

7. Simulate freely (especially if you suspect path dependence) and use Gedanken (thought) experiments

Experience in ultra-running teaches us the value of knowing the course, knowing the equipment, and knowing how we perform under conditions that simulate the real thing. Most runners do not go out and run their "A" race without running something similar under similar conditions. For me, before I get to the starting line of my A race each season, I have run at least a few shorter races (typically 5 to 6 of them in increasing order of distance, from 50K to 100K before a 100 mile race), with enough time to correct for any gross errors. Simulating the real thing allows us to go through all the elements – pace, nutrition, gear, etc. before the actual thing.

The brilliant work of many theoretical physicists resulted in an exact solution of the two dimensional Ising model almost 75 years ago. However, the solution of the three-dimensional

Ising model in exact form has still eluded researchers. However, by simulations we know with a high degree of accuracy what the behavior of the three dimensional model is. The most impressive demonstration of simulation in physics did not even require a computer. The discovery of special relativity purely by the thought experiment of trying to ride a beam of light by Einstein shows that even mental simulation can be sufficient to discover simple organizing principles. Such "Gedanken" experiments are early precursors to the more sophisticated simulation techniques on our desktops. But the history of simulations precedes even Einstein. In 1638, Galileo proved by a thought experiment[21] that falling objects in vacuum must fall at the same rate regardless of their masses.[22] Feynman famously said that all of quantum mechanics could be understood by carefully thinking through the implications of the two-slit experiment mentioned earlier, since it results in phenomena that are impossible to explain in any classical way.

All professional pilots (all airline pilots included), are required to pass flight tests at the level of an ATP (airline transport pilot) in a simulator (the "actual thing – flight in a real airplane does not count"). The way this is done is through an intense initial training course anywhere from a week to a month (this is after a 1000+ hours of basic training as a private or commercial pilot with instrument rating in the type of the aircraft to be flown). This is followed by usually an annual (if not more frequent) recurrent training session in a full motion flight simulator (the one I use is in Wichita, Kansas). Once inside the simulator, each switch is in the same place as in the actual airplane. The instructor can control weather, runway conditions, day vs. night at the flick of a switch. More importantly, they can create emergencies on whim and when least expected. The reason this is so critical is that during a typical trading career, emergencies or rare events happen, only rarely. So we are not used to reacting to them. By compressing time into an intense set of events in a simulator, the pilot can practice the rare event over and over again, until it becomes almost as normal as a normal event. In other words, the aid of full motion simulator realistically allows the pilot to focus on the tail of the distribution, which is impossible to replicate in a real airplane with high enough frequency.

The special thing about investing is that we only see one path of the markets, hence we cannot perform real life experiments – we are forced to use thought experiments and simulations. One example of this is the path dependence of asymmetric monetary policy – what the Fed will do in response to the market's response to its own behavior. This, as we have shown, by simulations, says a lot about the yield curve.[23] More practically for you and me, when saving for retirement,

[21] Discorsi e dimostrazioni matematiche (1638).

[22] The logic goes as follows (in the book it's a discussion between Simplicio and Salviati): if a heavier object is expected to fall faster than a lighter object, then the combination of the two objects should fall faster than either of the original objects. But if the two objects were to be joined, we should expect the lighter object to retard the rate of fall of the combination. The two assertions are mutually contradictory unless all objects fall at the same rate.

[23] V. Bhansali, M. Dorsten and M.B. Wise, "Asymmetric Monetary Policy and the Yield Curve", Journal of International Money and Finance, Vol. 28, 8, 2009.

the fact that we can only observe one history becomes critical, since we are apt to react not probabilistically to adverse events, but deterministically (selling out of we get too fearful, too early, or at the wrong time). This is precisely one area where simulation of future outcomes can be extremely powerful, since it allows imagining scenarios that have not happened in the past. Exhibit 17 shows the expected maximum drawdown of a retirement portfolio with different levels of equity beta, and indeed points the investor to prudent risk management. With powerful simulation tools such as Mathematica at your disposal, there is no excuse for imagining a world like never seen before, and quantify the tradeoffs in a fully probabilistic sense. In the world of investing, simulations the all too prevalent bias of trusting the back-test too much – simulations allow one to imagine alternative universes and to practice our responses to them. In a sequence of papers and a book I have authored on this topic, the benefit of tail risk hedging becomes apparent.[24] Another direct application of the strategy of simulations is to world of systematic, quantitative investing. In a portfolio management approach similar to risk-parity, where allocations are made on the basis of risk equalization, we can carefully isolate scenarios where the approach makes sense and where it fails.[25] There are also brilliant examples of solutions of real-world finance problems with "thought experiments". Indeed the Black-Scholes formula is derived by simulating a locally hedged portfolio. Or look at the solution of any path dependent option pricing problem (almost all exotic options are path dependent, and only in certain cases can we replace the computation with an analytical solution. The arithmetically averaged "Asian" option to this day does not have an analytical solution, but it trades! So simulation is the only way to solve many real world problems. A great application of this technique is Fung and Hsieh's approach to explaining the "factor" that explains the performance and risk characteristics of trend followers (a "lookback" straddle).

[24] V. Bhansali, "Tail Risk Hedging", McGraw-Hill, 2014.

[25] Bhansali et. al. " The Risk in Risk-Parity", "Active Risk Parity", and "Beyond Risk Parity", papers in the Journal of Investing, 2011 and 2012.

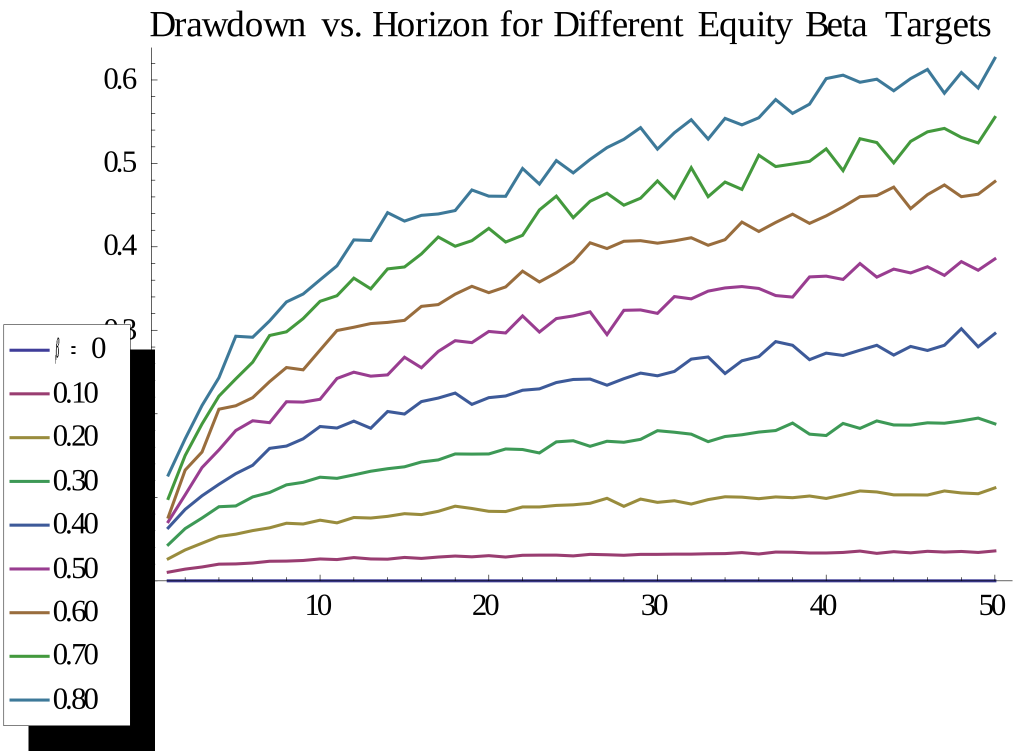


Exhibit 17: Expected maximum drawdown from unit value vs. time horizon for different equity beta exposures for a retirement portfolio. (Source: author)

8. Identify tradeoffs (and deal with imbalance by owning free or cheap options).

Think of the tradeoffs. In both running and aviation there is a clear tradeoff between speed and endurance. If you increase your speed, beyond a particular speed the endurance drops off, usually precipitously. This, as we spoke about, has to do with the finiteness of some variables, such as energy.

Statistical physics is based fundamentally on the tradeoff between energy and entropy. In the classic model for magnetization, two different forces are in competition. The first one is minimization of energy – it makes sense for all spins to align. The other one is maximization of entropy – the state of the systems in which all spins are aligned is also a lower probability state. As one varies the temperature, the system undergoes a phase transition, where the two forces are balanced, and spontaneous magnetization occurs.[26] And of course the critical moment in any movie is when the protagonist has to make a choice between a rock and a hard place. The choice sets the path of the movie and its ultimate resolution.

In investments, there are also clear tradeoffs; for example between return and risk. In most cases, tradeoffs can be simply encapsulated in terms of a ratio. For example, the Sharpe ratio, the Treynor ratio and many others are simply ways to encapsulate the tradeoff in one quantity. When dealing with catastrophic financial events, the ratios are typically specified in terms of multiples; e.g. the risk premium compensation per unit of actual risk. For instance, corporate bond spreads for high rated bonds are many times larger compensation than the actuarial or realized risk of these types of bonds. Equities deliver a large premium over the actuarial value of their losses, and traditional re-insurance of course delivers a high multiple over the realized value of their losses. This excess compensation, as we discussed in the first part of this paper, can be used to build diversified carry portfolios. On the other hand, there are times when the multiples, or excess compensation is low, zero, or even negative. These are times when it pays to purchase insurance for little premium or in fact get paid and get to own insurance. Credit derivatives based hedges prior to the 2008 Global Financial Crisis were a great example of this dynamic. The trick is to balance the tradeoffs optimally in the face of all sorts of uncertainties (some of which are tabulated in Exhibit 18).

[26] See, for example, my JOIM paper "Can the physics of phase transitions tell us anything useful?", JOIM, 2010.

Event	Category	Indicator 1	Indicator 2	Indicator 3	Prob
Unexpected inflation or deflation print	ECONOMIC SURPRISE	Core CPI	PCE		25%
Fed Tapering Talk	MONETARY POLICY	Libor	MBS basis	5y 5y	20%
New neutral realization/equity derating	ECONOMIC SURPRISE				20%
Middle East/Syria/Iran	GEOPOLITICAL SURPRISE	News			15%
China economic crash	CHINA SURPRISE	China econ data			15%
Continued Equity market rally forces short covering melt up	POSITIONING/LEVERAGE	Flattening Skew	Vol rising on rallies		15%
Japan/China territorial dispute	GEOPOLITICAL SURPRISE	News			10%
Failed Auction	MARKET PLUMBING	Auction tail			10%
Large muni default	CREDIT MARKET SURPRISE	MMD curve	Muni OAS blowout	Ratios sharply widen	10%
Surprise negative earnings from US companies	ECONOMIC SURPRISE	Earnings data			10%
Europe peripheral problems	EUROPE SURPRISE	Peripheral spreads			10%
Hedge fund failure	POSITIONING/LEVERAGE	Increased correlations	Sharp Skew increase	Treasury bid	10%
ECB QE surprise	MONETARY POLICY	EONIA	Euribor		10%
Japanese QE failure	MONETARY POLICY	Japan sov CDS	Yen libor		10%
Energy shock	SUPPLY SHOCK	CL	CO		10%
Vol targeting gone awry (aka 1987 in a new guise)	MARKET PLUMBING	SPX			10%
Washington surprise	GEOPOLITICAL SURPRISE	News			10%
Ukraine gets worse	GEOPOLITICAL SURPRISE	News			5%
China sells treasuries or rebalances fx	CHINA SURPRISE	Sharp yield rise	Large dollar move		5%
North Korea going fully rogue	GEOPOLITICAL SURPRISE	News			5%
Housing re-bust	LEVERAGE DRIVEN	Housing data			5%
Repo market failure	MARKET PLUMBING	10 y special Repo rate	GC spike		5%
Hi frequency flash crash	MARKET PLUMBING	Sharp intraday market swings			5%
ETF redemption/cash problem	MARKET PLUMBING	ETF discounts			5%
Fed Surprise Rate Changes	MONETARY POLICY	Fed funds			5%
Fed Botched Operations	MONETARY POLICY	FRFA issues			5%
New regulation	POLICY SURPRISE	News			5%
Failed M/A deal and financing problems	MARKET PLUMBING				5%
Brazil shock	ECONOMIC SURPRISE	CDS	BRL		5%
Sovereign Default	GEOPOLITICAL SURPRISE	Sovereign CDS	News		5%
Bank Failure	MARKET PLUMBING	VIX	Bank CDS		5%
False "tweet" or news	GEOPOLITICAL SURPRISE				2%
Large scale terrorism on US soil	GEOPOLITICAL SURPRISE	News			1%
Cyberattack on US infrastructure	GEOPOLITICAL SURPRISE	News			1%
Bird flu	GEOPOLITICAL SURPRISE	News			1%
Nuclear reactor meltdown	LARGE SCALE ACCIDENT	News			1%
Meteor	NATURAL DISASTER	News			0.50%
Natural Super Cat	NATURAL DISASTER	News			0.50%

Exhibit 18: Things that can go wrong nd result in increased volatility. (Source: author)

9. Don’t fight momentum (look for it and ride it)

One of the mantras when running a long race is to build momentum, and not do anything that stops the momentum. What this means in practice is to take what the trail gives you, and not fight it. On uphills, you want to slow it down to conserve energy. On downhills you want to run like a “marble in a groove”. This is also the secret behind pacing. You want to get stronger as the race goes on.

Of course, flight, as almost all physical systems having to do with movement, is based on momentum. The whole point of aircraft engines is to maintain forward momentum so that the laws of physics can take over and create lift (and to even avoid hazardous wind shear shown in Exhibit 19). Similarly in any good movie, you will observe that the momentum of the story seemingly starts to increase somewhere near the middle. Events start to get compressed, and time seems to move faster as the story starts to take a more singular direction. Good screenplay writers know how to evoke this momentum in the story. Urgency and momentum are created by introducing a “ticking clock”. A time limit is given to the characters for achieving their wants (notice this next time – most movies have some sort of time limit within which something needs to be accomplished by the protagonist). The other way to achieve momentum is to “advertise”, i.e. letting the audience know something important is coming up, so they can be focused on where the story is going and keeping them involved.

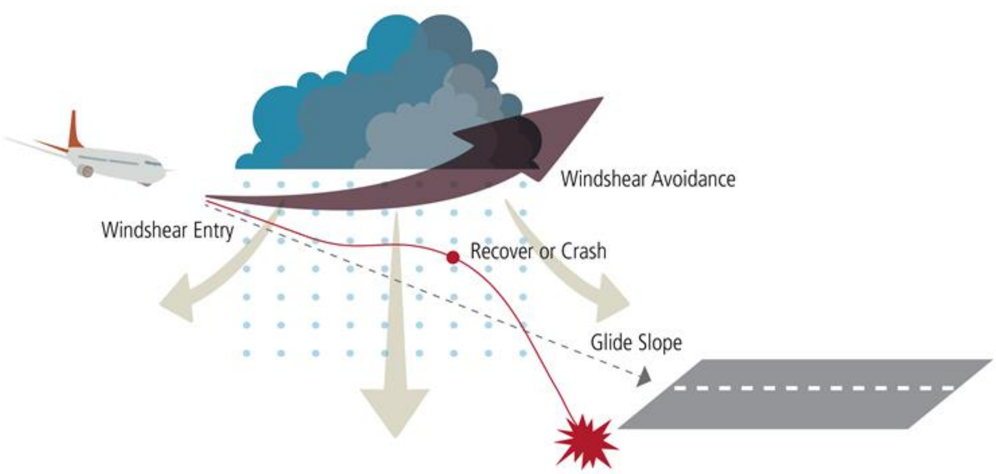


Exhibit 19: Maintain momentum to avoid shear. (Source: author)

In financial markets, however, there is a tendency for participants to want to pick bottom (“catch a falling knife”), or pick a top, since, as behavioral finance would tell us, we get satisfaction relative to a known and familiar reference point. We also expect to be generally better than the average, and expect that analysis always yields value. So it makes sense to trust the results of "thoughtful" analysis than the result of the madness of crowds in the markets; the simplest representation of which is price momentum. Recently there has been very powerful theoretical research (which explains momentum as a consequence of behavioral biases), empirical validation (which looks at momentum as a proxy for option replication), and historical cataloguing (which collects data over the last 200 years!) that supports momentum (especially time series momentum) as a critical risk factor for investment portfolios.[27]

10. Pay heed to the environment (and be flexible in adjusting to it)

One of the biggest mistakes one can make (I made it in 2013), is to run every race with the same strategy, regardless of environmental conditions. This is a classic example of ignoring the initial conditions and/or the boundary conditions. In 2013 the temperatures at the WS100 were recorded to be the second highest in the 40 years of the race (reaching a peak of 108 degree F). This was well known before the start of the race, and the race directors warned the runners that this meant throwing out target splits and times and personal record goals out of the window (generally speaking running generates heat, and if the environment is over 100F there is very little heat dissipation, so the body has to slow down to conserve critical functions). This was a classic case of the need to modify running strategy in the face of very adverse environmental factors.

While in flight, every pilot sooner or later meets turbulence. The traditional edict of maintaining "attitude, altitude and airspeed" goes out the window. You simply cannot maintain airspeed in the presence of turbulence, since this will stress the structure and possibly permanently damage the plane and its occupants. You cannot maintain altitude either, since the change of airflow will cause the aircraft to rise and fall. But you can try to maintain attitude, i.e. keep the plane straight and level.

The recent history of financial markets is one dominated by the impact of government. From active participation in deciding what institutions should survive to what shouldn't; to the intervention via regulation; to unprecedented policy action; to purchase of securities. These are distortions created in the aftermath of one of the largest periods of turbulence in investing. To maintain investment principles without modification (such as a belief in diversification as the sole method for portfolio risk control), is likely to result in major structural damage over time.

The low implied volatility across all assets in today's investing environment is thus likely a function of the excess liquidity and volatility suppression of the world's central banks. As mentioned earlier in the simulation section, we incorporated modified Taylor rules (that allow for asymmetry in the Fed's response function to deflation) and simulated how yield curve shapes would be affected (see Exhibit 20). Indeed we find that the markets today are reflecting

[27] See Moskowitz, T.J., Ooi, Y.H>, and Pedersen, L.H., "Time Series Momentum", Journal of Financial Economics, 2011. For a 200 year history, see Lemperiere, Y. et. al., "Two Centuries of Trend Following", Journal of Investment Strategies 3,(3), 41-61.

continued influence of the central bank for the next few years. Being aware of this fact and the inherent risks when the Fed's posture eventually changes, we can position our portfolios for an eventual rise in volatility and risk.

Inputs				
Initial inflation	Initial output gap	Min inflation	Max inflation	Monte Carlo iterations
1.5	-2.0	1.0	5.0	10000

pi*
2

rstar
1

Economic model involving inflation (pi) and output gap (y)

$$\begin{aligned} \mathrm{d}\pi(t) &= \mu_\pi \mathrm{d}t + \alpha_1 \pi(t)\mathrm{d}t + \alpha_2 y(t)\mathrm{d}t + \sigma_\pi \mathrm{d}w_\pi, \\ \mathrm{d}y(t) &= \mu_y \mathrm{d}t + \beta_1 y(t)\mathrm{d}t - \beta_2 \left[i(t) - \pi(t)\right]\mathrm{d}t + \sigma_y \mathrm{d}w_y. \end{aligned}$$

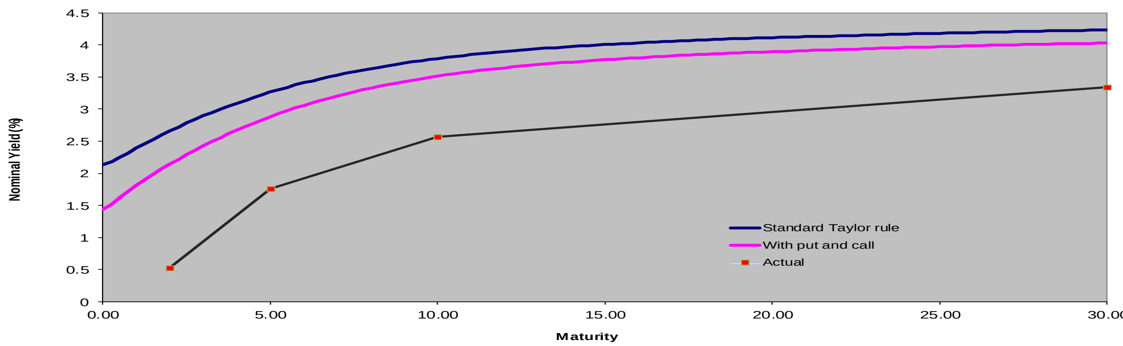
Standard Taylor rule for policy rate

$$i_0 = r^* + \theta_1(\pi(0) - \pi^*) + \theta_2 y(0) + \pi(0)$$

Taylor rule with put and call

$$i_0 = \hat{c} + \hat{\theta}_1 \pi(t_0) + \theta_2 y(t_0) + \frac{\alpha}{2}\operatorname{erfc}\left(\frac{E_{t_0}[\pi_{\mathrm{avg}}] - \pi_{\min}}{\sqrt{2\sigma^2_{\pi_{\mathrm{avg}}}}}\right) - \frac{\alpha}{2}\operatorname{erfc}\left(\frac{\pi_{\max} - E_{t_0}[\pi_{\mathrm{avg}}]}{\sqrt{2\sigma^2_{\pi_{\mathrm{avg}}}}}\right)$$

$$\hat{c} = r^* - \theta_1 \pi^*, \quad \hat{\theta}_1 = \theta_1 + 1.$$



Yields

	Model Yields No Put and Call	Actual Yields	Actual - Model No Put and Call	Model with Put and Call	Actual - Model with Put and Call
2	2.659	0.530	-2.129	2.144	-1.614
5	3.264	1.760	-1.504	2.880	-1.120
10	3.787	2.570	-1.217	3.514	-0.944
30	4.231	3.340	-0.891	4.030	-0.690
			3.012		2.286

Curve Shape

	Model Steepness No Put and Call	Actual Curves	Actual-Model No Put and Call	Model Steepness with Put and Call	Actual - Model with Put and Call
2/5s	0.605	1.230	0.625	0.736	0.494
2/10s	1.128	2.040	0.912	1.369	0.671
2/30s	1.572	2.810	1.238	1.885	0.925
10/30s	0.444	0.770	0.326	0.516	0.254

Exhibit 20: Stochastic simulations of Yield Curve from Taylor Rules with and without asymmetric Fed policy. (Source: author.)

11. Bonus tip: If its too good to be true it probably is (so don't abandon commonsense)

This tip does not really require too many examples, but note that even science can do a 180 switch over time. Sugar was once (1969), thought to be "diet food", controlling the appetite by suppressing your "appestat"! (Exhibit 21).

Exhibit 21: Sugar as diet food? Source: TIME magazine

The history of flight is full of examples of perpetual machines that promise levitation without the use of energy. Though I admit the possibility of many physical laws being incomplete or wrong, it is hard to see how they can violate conservation of energy in a blatant manner.

However, finance is full of "get rich quickly" themes and tricks that every investor should watch out for. Despite the enormous amount of regulation, we as humans are eager to believe that we can beat the laws of gravity, and get sucked into many of these schemes.

Let me give two quick recent examples.

The example of CYNK (see WSJ July 12, 2014) is a great example of how you can get fleeced. This company, with one shareholder and an address in Belize, and no business, was valued at $6BN (price of company = stock price x shares outstanding).

The second example is the ubiquitous "zero-cost" structure that purports to give you something for nothing. A common hedge in the aftermath of the crisis was the sale of upside equity index calls to finance the purchase of out of the money puts. The problem was that even though these looked free as a package, the package consisted of selling very cheap out of the money calls to finance the purchase of very expensive out of the money puts (due to the volatility skew, the "roll-up" of volatility and generally positive performance of the equity markets). One metric of the relative cheapness (see Exhibit XX) is the ratio of the strike "distance" of the puts to the calls. So if you bought a 5% out of the money put, you would sell a 2-2.5% out of the money call to make the structure zero cost.

If there is one thing discipline and the desire to excel in a hobby has taught me is that if its too good to be true it probably is.

Conclusions and Summary

In this admittedly "light and personal" piece I have intended to weld together what I know about quantitative investing along with lessons I have learnt from activities that I do outside of investing. I think of these handful of observations as additional tools (supplementing rigorous fundamental and quantitative analysis) to make intelligent investment decisions. I am sure that many readers will find similar analogies from their own experience that they have consciously or subconsciously utilized in their own investment decisions.