Give me a moment:

Optimal leverage in the presence of volatility, skewness, and kurtosis

Lars N. Kestner
John O'Dowd
March 2022

Summary

This paper discusses the role that volatility, skewness, and kurtosis play in the distribution of wealth for a series of returns with reinvestment. We begin with a simplistic model to explain how volatility shapes the wealth distribution of investments. We then look at how changing leverage varies an investor's returns, solving for an amount of leverage that maximizes this wealth. Lastly, we introduce a new method to determine optimal leverage to maximize wealth when investment returns are not normally distributed, as is the case in most financial markets.

STATISTICALLY APPLIED TRADING
www.satquant.com
New York

Introduction

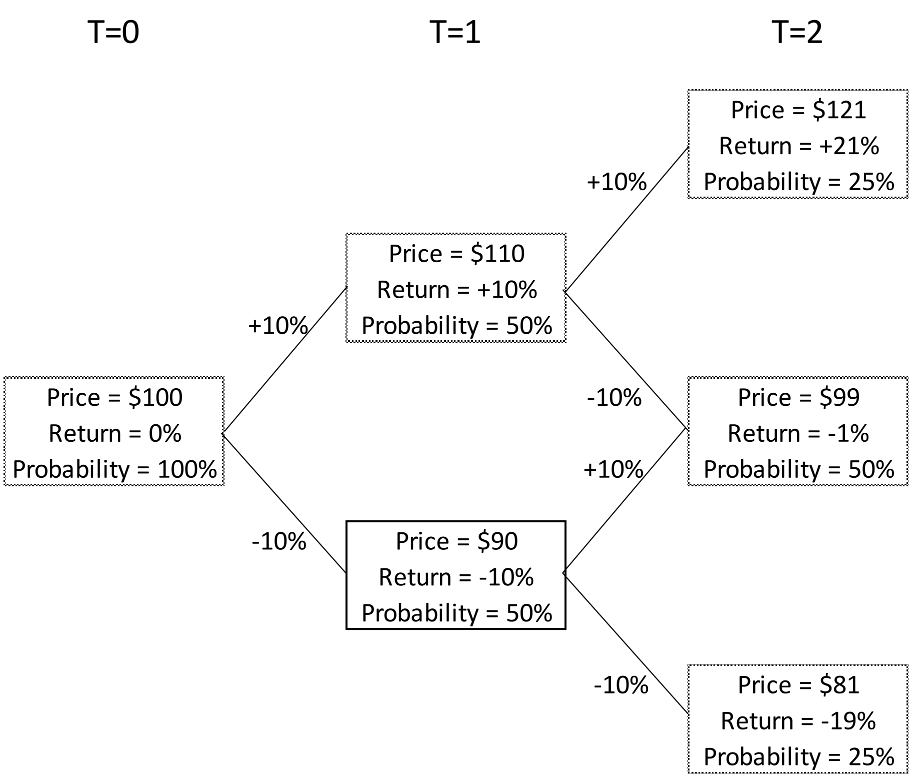
This paper discusses the role that volatility, skewness, and kurtosis play in the distribution of wealth for a series of returns with reinvestment. The magnitude of each measure shapes the amount of leverage an investor can apply to that investment. Part I of this paper is intended to explain how volatility and leverage shape the results of an investment strategy. We apply theory to practice and set the table for Part II of this paper where we explore special cases of investment returns where the 3^{rd} and 4^{th} moments are non-normal in nature.

Taking a page from derivatives pricing

Derivatives pricing often uses a binomial tree to simplify the price change of an asset. The asset price starts at today's level, say $100 per share, and can move in two states only: up or down by a return increment with a probability p. The value of the asset at each time T is the probability weighted average of future outcomes, discounted by an interest rate.

Here we will employ a tree with simple parameters, where the rise and fall are of equal percentages (+/-10%) and the probability of both rise and fall is 50%. With each step in time, the tree expands possible paths. The asset price on T=0 is $100.

After one time step, there is a 50% chance of a 10% rise to $110 and a 50% chance of a 10% decline to $90. The probability weighted average price is $100, the starting price.



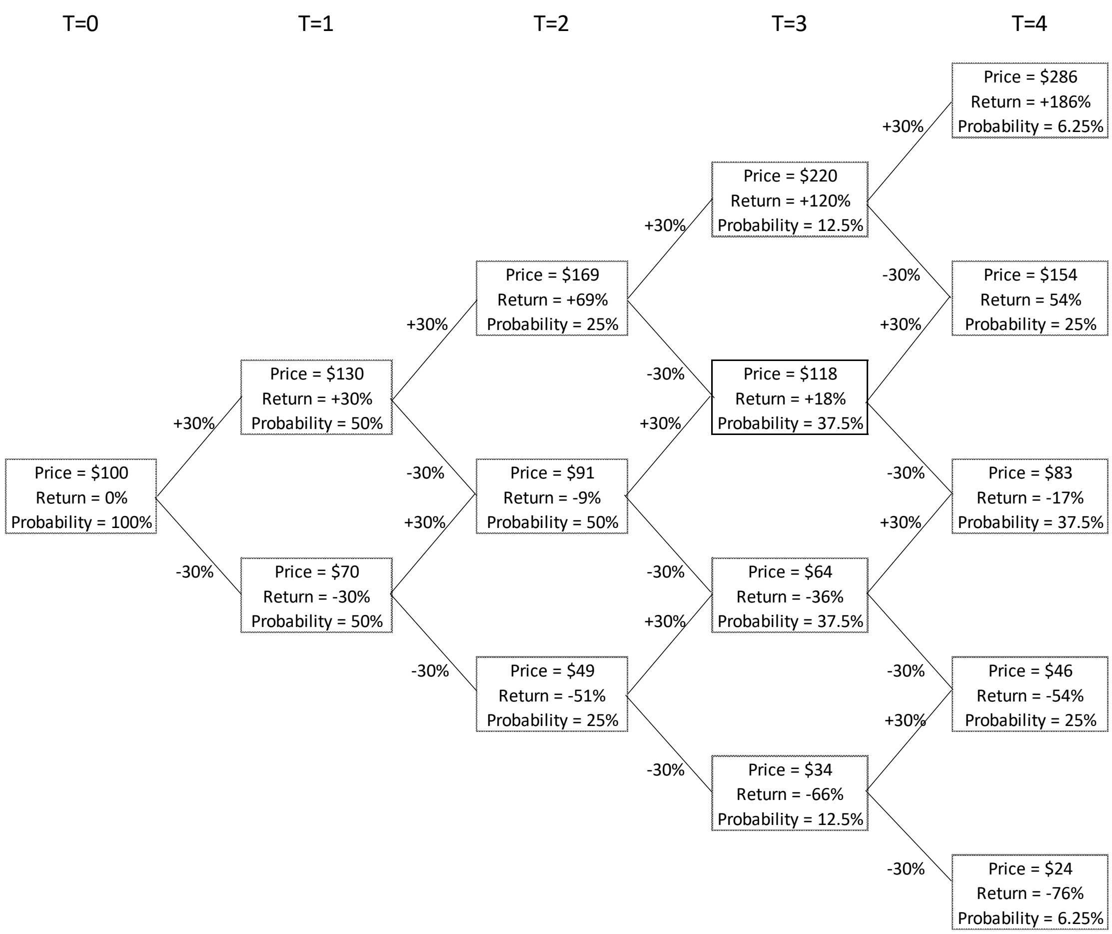
Binomial tree where each step is +/-10% return with probability=50%.

Adding another step in time at T=2, we see that the +21% return of the 2 win path is slightly larger than the 2 loss path return of -19% is negative. The 1 win/1 loss path, occurring with probability=50%, leads to a small -1% negative return[1].

[1] A 1 win/1 loss path will always lead to a price lower than the starting price as $(1-x)(1+x)=1-x^2$.

Now, let's increase the volatility of the asset by increasing the return size on each step to +/-30%.

The same pattern holds but magnifies with the increase in volatility. The asymmetry of the 2 win versus 2 loss scenarios widens (+69% versus -51%). The loss associated with the 1 win/1 loss scenario also becomes more negative (-9% in the higher volatility world versus -1% in the previous lower volatility world).

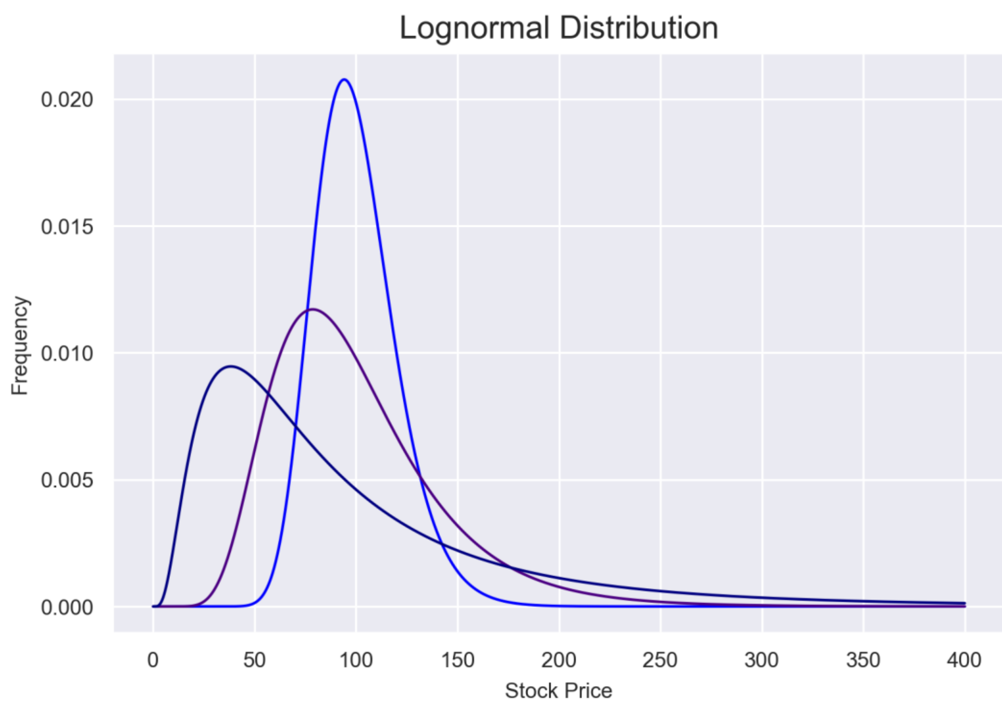


Binomial tree where each step is +/-30% return with probability=50%.

Going out further to T=4 in our higher volatility tree, we see the pattern even more pronounced. The most positive path (4 wins) yields a return of +186% while the most adverse 4 loss path yields a return of -76%. The 2 win/2 loss path yields a price of only $83, for a return of -17%. Despite having an equal number and equal size of wins and losses, this decline in investment value is often referred to as "volatility drag".

As we increase the number of steps and shrink the length of time between steps, our tree begins to converge to geometric Brownian motion. Geometric Brownian motion is used to model the price diffusion of many financial markets and leads to a lognormal distribution of asset prices.

Below is a graph of probability distributions of a standard lognormal for an asset with zero mean after 1 year when varying the annualized volatility at 20%, 40%, and 80%, respectfully.



Lognormal return distribution for an asset with initial price of $100, zero mean, and volatility of 20%, 40%, and 80%.

As we increase the volatility of our asset, we create paths with far-right tails but move more of the outcomes to the left-side of $100. This phenomenon shows the volatility drag concept in continuous return space and how it becomes more pronounced at higher volatility. As volatility increases, the distribution shifts to where more of paths are losers and a small outlier of very large winners emerge. All this occurs despite an expected value of $100 across all three distributions.

What does this “mean”

The goal of investing is to find positive expected returns and take risk. But as we saw previously, high volatility can create a profile where most outcomes are negative, while only a few lucky paths receive stellar returns.

Even with a positive expected return process, we can adversely affect our wealth by taking too much risk relative to our expected returns. If the volatility underlying our investment process is too large relative to expected returns, we risk pushing the bulk of our ending wealth possibilities to the left side of the distribution. When we do this, there will be a small chance of stellar outliers, but investors will need a lot of luck to participate in these lucky paths.

Consider a game where we flip a coin that has a 75% chance of landing heads and 25% chance of landing on tails. If the coin lands on heads we double our bet. That is, we receive our bet back plus the amount of the bet as winnings. If the coin lands on tails, we lose our bet. We can play the game as many times as we like.

This game is clearly in our favor to play as:

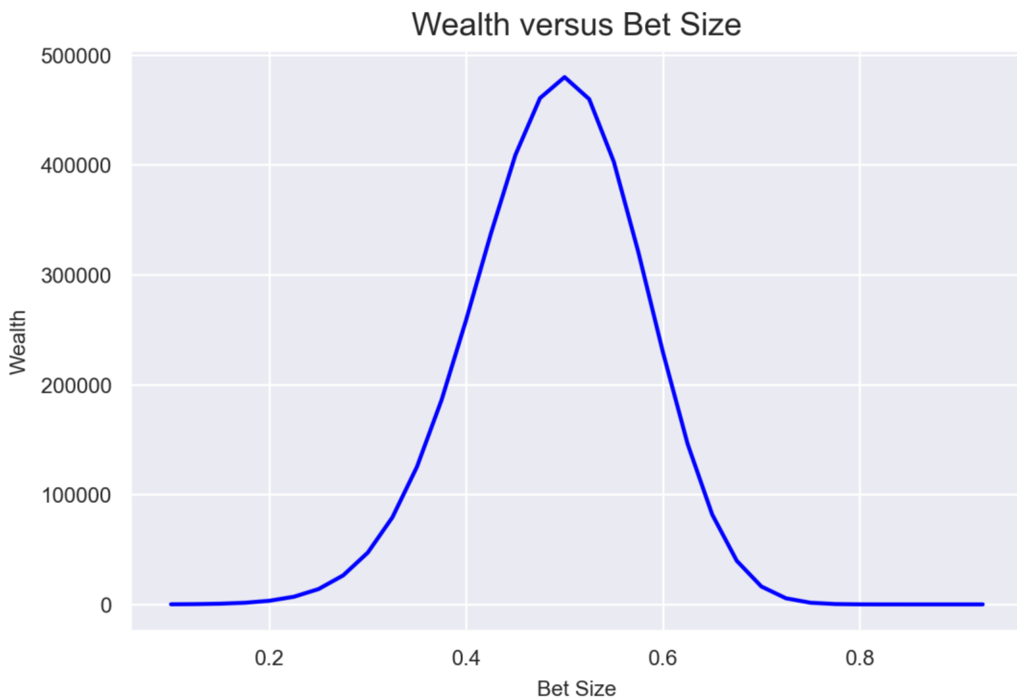
P(Win)*Return + P(Loss)*Return = (75%*100%) +(25%*-100%) = +50%

Despite the positive expected value, it would be unreasonable to bet all our accumulated wealth on each coin flip. If we did, except in incredibly remote cases of extremely good luck, eventually tails will appear and all our wealth would be lost with no possibility of continuing to play the game. But if we seek to maximize our expected dollar value on each coin toss, then we should bet everything each time we play. So what gives? The solution is that we need to maximize the return while also defending our capital base for future games.

If we know the return and volatility characteristics of a game or investment with certainty in advance, there is a solution to determine how much risk investors should take. The concept dates to the 1950s when J. L. Kelly Jr. published "A New Interpretation of the Information Rate". The paper utilized Kelly's background in Information Theory and focused on its application to risky bets. Gamblers and investors picked up on the work's unique usefulness in games that repeat. Kelly's paper highlighted that to maximize the long run growth rate, we should maximize not the expected value of wealth, but rather the expected value of the logarithm of wealth.

For the special case where the winning payoff equals the losing payoff (as in our coin toss game), the optimal Kelly bet is 2p-1, where p is the probability of winning the bet. We should bet 2*75% - 1 = 50% of our accumulated wealth on each successive flip of the coin.

To see this in practice, assume we play this game 100 times consecutively with the actual flips following the expected results of 75 heads and 25 tails. We start with $1 of wealth and bet varying percentages of that wealth on each flip of the coin. The graph below shows how increasing our risk as we bet leads to larger ending wealth. But once we bet more than 50% on each coin flip, the increased risk taken has a negative effect and leads to lower wealth as the drag from larger losses is more difficult to overcome.



Wealth curve after playing the coin toss game at varied bet size L.

Translating coin flips to continuous returns

Unfortunately, we have not found participants willing to offer a coin flip game with such positive expected value. However, there are financial assets and strategies with positive expected returns that we can play many times in succession. We can scale the size of these bets using leverage to achieve maximum long run wealth in a Kelly type context.

Suppose we own a stock with a 10% expected return and 20% annualized volatility[2]. Also suppose we have the ability for near infinite leverage, we can trade at any time scale, transaction costs are zero, and leverage has no cost[3]. In the context of Kelly, how do we maximize the expected value of the logarithm of wealth?

Kelly's maximization of the expected value of logarithmic wealth is typically calculated by maximizing the expected geometric growth rate. Here we use a popular estimate of geometric average and apply a leverage factor L to returns.

$$W^* = r - \frac{\sigma^2}{2}$$

where r is the expected return and W* is our proxy for ln(wealth)

Applying leverage L to the above leads to:

$$W^{L*} = L\,r - \frac{L^2\sigma^2}{2}$$

[2] As we will see later in this paper, it is important that the 3rd and 4th moments follow a normal.

[3] These assumptions are for the sake of simplicity and are maintained throughout this paper. Leverage and rebalancing are certainly not without cost and are time varying in nature.

Differentiating W^{L*} with respect to L, we solve for the maximum W^{L*}.

$$\frac{dW^{L*}}{dL} = r - L\sigma^2$$

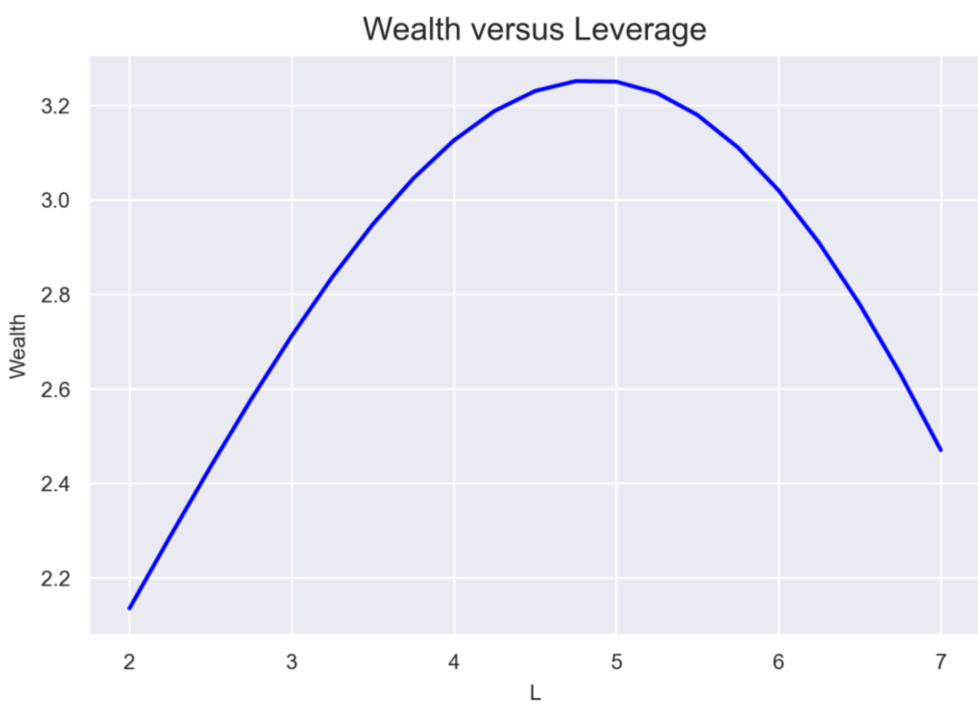
Using the descriptive statistics of 10% return and 20% volatility, we see optimal leverage of $r/\sigma^2 = 10\%/20\%^2 = 2.5$x.

Next, we use real world return data to confirm this property. We selected gold (proxied by the ETF: GLD) which between 2017 and 2021 had average annualized returns of 9.4% and volatility of 13.5%. If we knew these values for certain ex-ante, how much should we have been willing to risk?

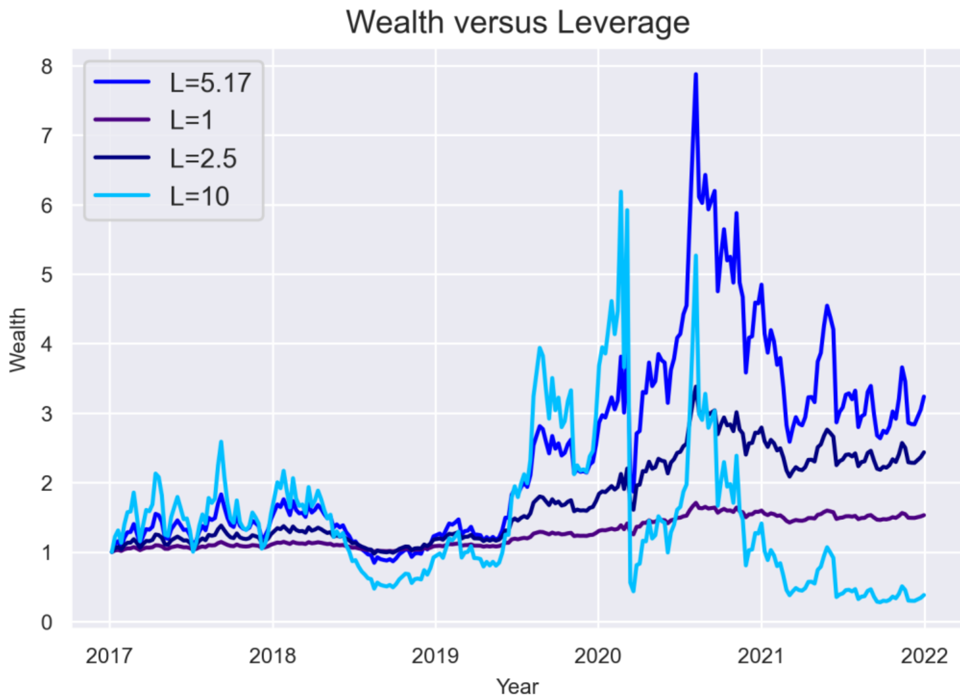
Solving for optimal leverage we calculate $r/\sigma^2 = 5.17$; that is, we should borrow 4.17x our wealth and own 517% of our wealth in the asset, rebalancing back to that leverage amount as the price of the asset increases or decreases.

In the graphs below, we simulate running L from 3.0 to 7.0 and calculating the ending wealth for each increment of leverage.

Applying various rates of leverage to returns we do see that L=5.17 is approximately peak wealth.



Wealth of long GLD at varying leverage of L. We continue to assume zero financing and rebalancing costs for the sake of simplicity.



Wealth curves at four levels of leverage.

Each of the paths in the graph above were generated using the same underlying return and volatility (or put another way, equal Sharpe ratios) but with varying amounts of leverage. We see that increasing leverage boosts ending wealth to a point, but not beyond. Trading above optimal leverage leads to more risk and lower wealth, just as Kelly suggested.

Preference for positively skewed return distributions

Part I of this paper noted the need for 3rd and 4th moments of the return distribution to be normal-like for the optimal leverage calculation to hold true. But what happens if returns have excess skewness or kurtosis? Because financial assets are prone to having excess tails, we need a more robust solution for determining optimal leverage that incorporates information about the return distribution beyond mean and volatility. This topic is explored here in Part II.

We show that, for investments with identical reward to risk, investors can apply more leverage to positively skewed returns and generate larger ending wealth. Meanwhile, negatively skewed distributions require investors to use less leverage, thus limiting wealth.

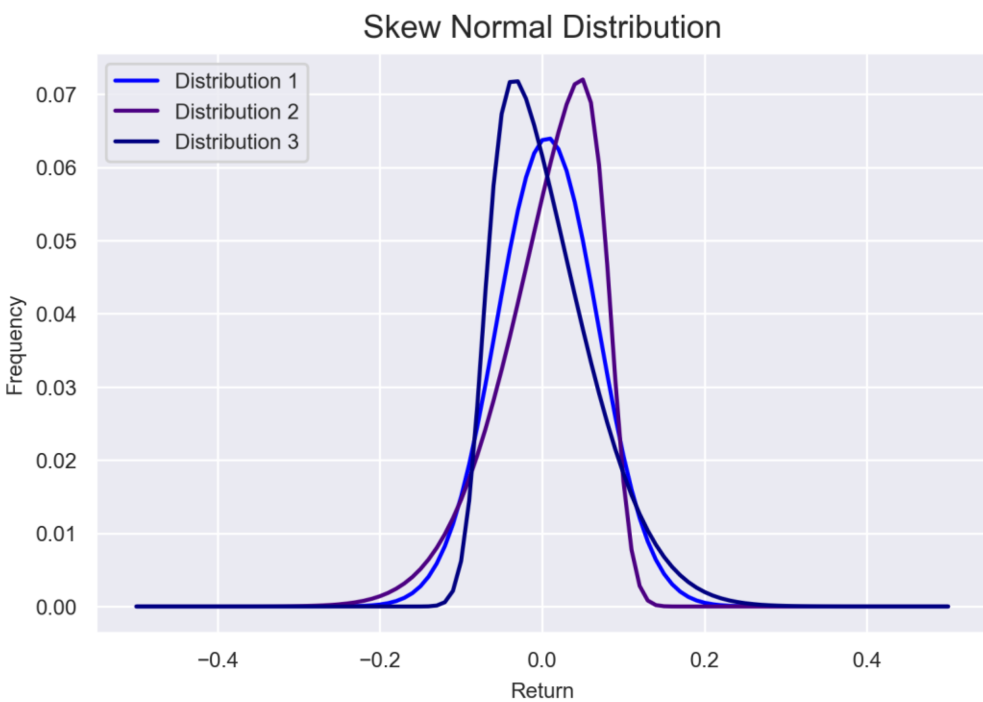
To illustrate this point, we create three return distributions, each with the same realized return and volatility. The first distribution has no skew, the second has negative skew, and the third has positive skew. To construct our return distributions, we employ the skew normal distribution[4]. We select three parameterizations. Each parameterization has the exact same mean and volatility (hence, equivalent Sharpe ratios), but differ in their skewness and kurtosis.

[4] A very detailed paper on the construction of skew normal distributions as it relates to investment returns can be found at Logica Capital Advisers.

	Distribution		
Parameters	1	2	3
ξ	0.0062	0.0844	-0.0720
ω	0.0623	0.1000	0.1000
α	0	-5	5

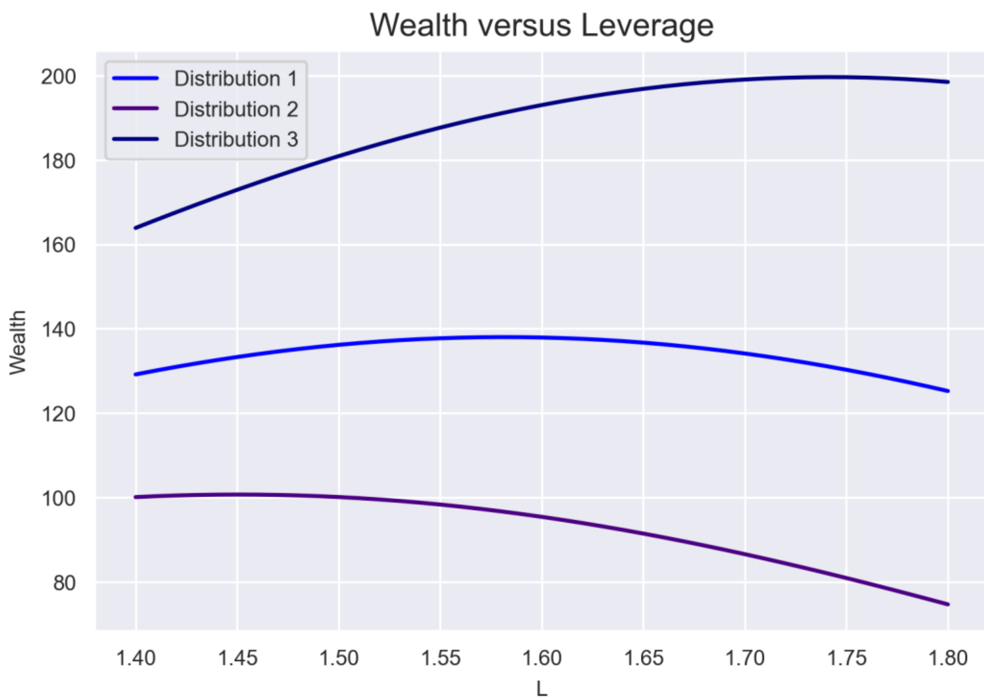
	Distribution		
Statistics	1	2	3
Mean	0.62%	0.62%	0.62%
Volatility	6.23%	6.23%	6.23%
Skewness	0.00	-0.85	0.85
Kurtosis	0.00	0.70	0.70

Skew normal parameters and their corresponding descriptive statistics.



Graphs of our three return distributions.

Next, we pick a leverage factor L and solve for wealth across L in each of the three distributions after 1,000 returns with reinvestment. We assume that realized returns follow each distribution's theoretical expectation perfectly.



Wealth versus leverage L with each of our three return distributions.

The graph above shows that despite the expected return and volatility being equivalent for all three distributions, the leverage L can be increased further before wealth peaks for the positively skewed return distribution. On the other hand, leverage for the negatively skewed return distribution peaks at lower L and lower wealth than either the zero or positive skewed distributions.

Estimating optimal leverage by including 3rd and 4th moments

In Part I, we presented an analytical solution to estimate optimal leverage using return and volatility estimates.

We pursue a similar path here while incorporating the 3rd and 4th moments of the return distribution. We use a Taylor expansion of expected log returns, following Wilcox (2000).

$$W^* = \ln(1+r) - \frac{\sigma^2}{2(1+r)^2} + skew\ \frac{\sigma^3}{3(1+r)^3} - kurtosis\ \frac{\sigma^4}{4(1+r)^4}$$

where r is the expected return and W* is our proxy for ln(wealth)

Applying leverage L to the above, we have:

$$W^{L*} = L\ln(1+r) - \frac{L^2\sigma^2}{2(1+r)^2} + skew\ \frac{L^3\sigma^3}{3(1+r)^3} - kurtosis\ \frac{L^4\sigma^4}{4(1+r)^4}$$

Differentiating W^{L*} with respect to L leads to:

$$\frac{dW^{L*}}{dL} = \ln(1+r) - \frac{L\sigma^2}{(1+r)^2} + skew\ \frac{L^2\sigma^3}{(1+r)^3} - kurtosis\ \frac{L^3\sigma^4}{(1+r)^4}$$

Using the descriptive statistics in the three skew normal distributions, we solve for L that maximizes our proxy for wealth.

Distribution	Solved Optimal L
1	1.62
2	1.49
3	1.80

Solved optimal L based on the 1st through 4th moments of each distribution.

Each of these solved optimal L's is very near peak wealth in each of the corresponding return distributions.

We can use this method to estimate the amount of leverage which maximizes wealth when given expected returns, volatility, skewness, and kurtosis.

Conclusion

Part I of this paper described the role that volatility plays in shaping the distribution of an asset or strategy over time. We saw that increasing volatility shifts the peak of wealth distributions to the left and creates a more skewed right tail. We then linked volatility's effect to Kelly's definition of optimal leverage, showing that too much volatility relative to return moves wealth sub-optimally.

Part II detailed the role that skewness and kurtosis have in determining how much leverage an investor can employ. Positively skewed return distributions allow more leverage, and thus higher wealth, than non-skewed or negatively skewed distributions.

Any estimate of optimal leverage is far from perfect as the non-stationarity of games played in the financial markets reduces their effectiveness. It is difficult enough to predict forward looking returns and volatility of an asset or strategy. Predicting skewness and kurtosis with true precision is an even taller task. Even if our crystal ball were perfectly accurate, trading at leverage multiples that maximize wealth will lead to exceptionally large volatility that likely exceed bounds tolerated by investors and professional allocators. Nonetheless, this paper describes the influence that third and fourth moments of the return distribution have on wealth and proposes a measure to estimate optimal leverage for those bold enough to trade there.

This document has been provided solely for information purposes. The information set forth herein has been obtained or derived from sources believed to be reliable, but it is not guaranteed as to its accuracy. Past performance is not a guarantee of future performance. This document is not research and should not be treated as research. The views expressed herein belong solely to the authors. The authors make no representations regarding the accuracy or completeness of this information. Readers of this document accept all risks in relying on the information within for any purpose whatsoever.