The volatility paradox: When winners lose and losers win

Consider the following hypothetical investment opportunity, a strategy that has a long, reliable data series of past returns covering all manner of cycles and market conditions. You are confident that the series is representative of what is likely to happen in the future. Over many years, this strategy has delivered a positive return, 20%, exactly half the time, and a negative return, -18%, exactly half the time.

Is this a money-making strategy or not?

The case for: if in any given year you have a 50% chance of adding 20% to your wealth (turning each $100 into $120) and a 50% chance of subtracting 18% (turning each $100 into $82), then this is a 50/50 proposition with the payoffs stacked in your favor. That adds up to an expected gain of 1% each year on average.

So, based on past return patterns, the strategy appears to have a positive expected return and ought to increase your wealth over time.

***The case against*:** A closer look at the exact same past return patterns shows that this strategy didn't work in practice: it actually lost money. Time and again it turned $100 into $120 and then into $98.40 (or $100 into $82 into $98.40). The longer the period, the greater the loss: 1.6% every two years, on average.

Can a strategy have a positive expected return, but actually tend to lead to a loss of wealth?

So can you really expect to add value in the future with this strategy, which lost you money in the past? Can a strategy have a positive expected return, but actually tend to lead to a loss of wealth?

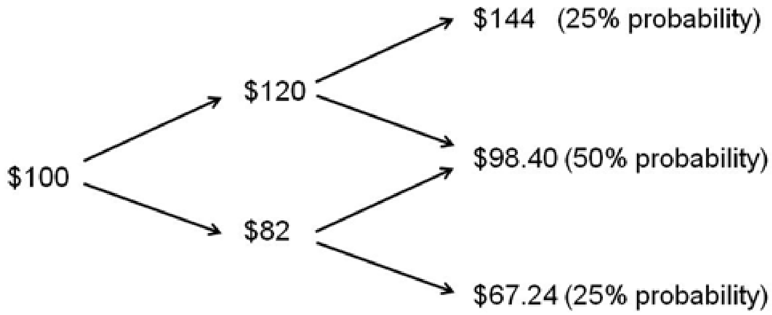
That's the volatility paradox.

The volatility paradox: A question of perspective

Let's start by disentangling the math above; then I will turn to some of the implications.

Like most paradoxes, the volatility paradox really flows from a subtle shift in perspective. So, first we must be clear that the case for and the case against our hypothetical investment opportunity are looking at two different things, and two different versions of what it means to *expect* something. The case *for* looks at a distribution of possible outcomes, and the case *against* at a single path drawn from that distribution. Just because our beliefs about the distribution are based on that single path doesn't alter the fact that these are two different things.

So the actual experience – the case against – represents a single path, with gains and losses in exactly equal numbers. The distribution – the case for – is based on probabilities (which *might* lead to an exactly equal number of each outcome over any given time period, but won't necessarily do so). We can see the impact on wealth below:



The average (mean) outcome after one year is ½ x (120 + 82) = $101. After two years it is ¼ x (144.00 + 98.40 + 98.40 + 67.24) = $102.01, a 2.01% gain (or 1% annualized). You can continue to build out the tree over more years, and the average wealth continues to grow at 1% compounded each year: 3.03% (or 1% a year) after three years, 10.46% (still 1% a year) after ten years. And so on.

But only one individual path will actually be experienced. So, looking forward two years into the future about what we can expect to experience from this strategy, there are four possible paths leading to three different possible outcomes. Even this very simple model of markets gives room for confusion: while the average return is positive, only one of the four paths makes a profit; the other three lose money. The *average* expected outcome is positive but the *median* is negative, i.e. more than half of the outcomes lose money.

The difference between the two versions of "expectation" is that one looks only at the probability of a gain or loss, and the other takes into account also the size of those potential gains and losses.

Even the volatility that you actually experience is not known: the standard deviation of the distribution of possible outcomes after two years is about 27%, while the actual path-wise volatility has a 50% chance of being zero at that point (if the same outcome is experienced twice). So neither the actual return nor the actual volatility of the path that is experienced in practice will necessarily bear any close resemblance to the expected return or the underlying volatility of the full distribution of possible outcomes that existed initially. The distribution of possible outcomes is not the same thing as the distribution of the experience along a single path drawn from that distribution.

Don’t expect the expected

Mathematically speaking, the expected value of a distribution is the average. Using this definition of expectation, our strategy is a money-making strategy. But we really expect the strategy to lose money (in the sense that it is more likely to do so). The difference between the two versions of “expectation” is that one looks only at the probability of a gain or loss, and the other takes into account also the size of those potential gains and losses. In other words, it’s because the potential outcomes are skewed that the word “expectation” becomes ambiguous.

The volatility drain: Not as simple as it sounds

This important distinction is easily overlooked when a picture of expected future experience of an investment strategy is built from past experience. One effect of this has been the prominence of a concept that is variously called “volatility drain,” “volatility drag,” or other similar names. This concept rests on the money-losing part of the analysis above. To illustrate the volatility drain, we need only note the following: even though a 20% gain followed by an 18% loss would leave you with less money than you started with ($98.40), if each of those returns were to be cut in half (10% and –9%), then you start to turn a small *profit* instead ($100.10 after two years). So a reduction in volatility automatically adds value, right?

Not at all.

The volatility drain argument doesn’t hold when it is applied to distributions: it is *not* the case that reducing the variation in the ending distribution also increases the expected return. Take the example of a half-size version of our original strategy described above: when we consider the full set of possible outcomes, it is easy to see that the effect of halving volatility here is also to cut expected returns in half. The expected gain of 1% a year that we discussed above turns into 0.5% a year: our two outcomes after one year are $110 and $91; the four possible paths over two years lead to outcomes of $121, $100.10, $100.10 and $82.81, and so on.

Variation along a path in the past does not equate to the variation among the different paths that could be experienced in the future.

So the volatility drain affects individual paths and the probability of certain outcomes, but it does not affect the average ending wealth. I would quickly note that there are very good reasons for wanting to reduce the volatility and the uncertainty of your outcomes, including that this would increase the median wealth value (which is, in general, a desirable thing.) But increasing the expected accumulation of wealth (in the average-value mathematical sense) is not one of those reasons, and the volatility drain argument should not be taken to imply that it is.[1]

Looking forward and looking backward

The distinction between a single path and the full distribution translates to a difference between looking backward and looking forward. Looking backward, there is only one path – no range of possible outcomes – so there is only one way to think of uncertainty: as the variation in the experience along that one path.[2]

---

[1] And while some might think, “Well, since I will experience only one path, the volatility drain argument will improve my actual outcome along that path, so that is a good thing” – that too is a flawed argument: you don’t in practice get to choose the volatility of the individual path, any more than you get to choose the return of the individual path. The decisions you make affect the distribution, not the choice of path.

[2] For example, the impact of rebalancing and volatility on backward-looking return experience in the specific instance of commodity markets is explored in “The Strategic and Tactical Value of Commodity Futures” (Claude B. Erb and Campbell R. Harvey, 2006: Financial Analysts Journal 62 [2]), work that is expanded in “Commodities Futures Returns: Reconciling History with Expectations” (Leola Ross, 2010: Russell Research).

But the future is less simple. Viewed from today's perspective, there are many possible paths. So there is not only the variation in experience along any one path to deal with, there is also the variation among the different paths.

Variation looking backward – the volatility that was experienced over the long history of this strategy – would normally be expressed as an annualized standard deviation (19% a year in this example.) But the variation in possible outcomes looking forward is not the same thing. And the variation along any single path chosen from that distribution is something else again. Variation along a path in the past does not equate to the variation among the different paths that could be experienced in the future. (And I'm leaving aside here all those real-world considerations about whether the underlying distribution is stable over time: I'm just talking here about a simple theoretical world in which it is.[3])

This distinction draws out some of the stranger aspects of the volatility paradox. It *is* possible for our experience with a particular strategy to have been negative, yet still support the belief that it has a positive expected average outcome in the future. If we take the single path we have experienced as the most representative expression of the return distribution, then it describes the median of what might be expected from the future – and medians can be negative, even if true averages (means) are positive. Indeed, if the past were the most representative experience of any given distribution, it can be argued that every strategy can be expected to generate, on average, greater wealth in the future than it did in the past![4]

The longer the time horizon, the more significant the volatility paradox becomes.

Time diversification doesn't work, either

Another oddity lies in the effect of time on the probability of making a gain. Investment professionals have long been schooled to believe that if a strategy is on average a winning strategy, then the more often you get to play, the more likely it is that its winning nature will shine through. You may be behind to date, but given enough time, you expect the favorable odds to assert themselves and a profit to result. So we might expect that if you made a series of 100 of these bets you'd be more likely to end up ahead than if you played just once. Again, this is not so: the longer the time period you look at, the more likely it is that this strategy will show a loss – even though the expected (average) return is positive.

Suppose I start with a stake of $100 and run a series of 100 coin flips, each of which results in either a 20% gain or an 18% loss. As I've shown above, the average outcome at the end of the 100 flips is a gain of 1% on each flip compounded to produce $270 after 100 flips. However, when we look at the whole range of possible outcomes, the great majority leave me with less than $100. The single most likely outcome (which in this case happens to be the same as the median) is a meager $45; there's roughly a 50/50 chance I'd end up with not even that much money. The odds that I will end up with more than $100 are just 31%. The odds that I will equal or better the "expected" outcome of $270 are less than 1 in 5. The reason the average is so much higher than the median is that those 20% of outcomes show very substantial gains. I've got a big chance of a loss, and a small chance of an outsize gain. So the effect of time on this strategy is not to make its true winning nature shine through at all, but rather to make it feel increasingly like a lottery ticket purchase (albeit a

[3] Actually, as I've written elsewhere – see Bob Collie, "Don't Marry a Model" (2011), *Russell Viewpoint* – I would dispute that there is an underlying distribution to real-world investment returns at all. But that argument is an unnecessary complication here. In this article, I am just dealing with a theoretical world in which an underlying distribution not only exists but is obligingly stable.

[4] Bizarre as this statement sounds, it is really nothing more than a general extension of the example we have been using throughout – a strategy whose past returns led to a reasonable belief that the expected return is positive, even though the actual experience was a loss.

lottery with the overall odds stacked in my favor).[5] Indeed, the longer the time horizon, the more significant the volatility paradox becomes. This is something that has been noted in the academic literature: Hughson et al have noted, “Because cumulative returns are positively skewed, the mathematical expected cumulative return substantially overstates the future cumulative return that investors are likely to realize, and the problem grows worse as the horizon increases.”[6]

Arithmetic and geometric returns

Mathematically oriented readers may have recognized that the effect I have described is closely linked to the difference between arithmetic and geometric average returns.

The wealth that was generated by a strategy in the past is most accurately captured by the familiar geometric return calculation. So if $100 turned into $98.40 over two years, or $45 over a hundred years, we would say that the annualized average return was –0.8%. And this number also provides a guide to the median value that the same distribution of returns should generate in the future. But it’s not the right starting point for an estimate of the expected average wealth that the strategy will generate: for that, we would need to look not at the geometric average but at the arithmetic average.

This jump from a geometric average for past returns to an expected arithmetic average for future returns is the mathematical explanation for the volatility paradox. It is also the explanation for why the volatility drain does not mean that a reduction in volatility increases expected future wealth if the arithmetic average return stays the same. Indeed, in the case of a normal distribution, the difference between these – the geometric and arithmetic averages – is proportional to the variance of the distribution.[7] So it is impossible to hold constant two of the three statistics (geometric average, arithmetic average and variance) while varying the third.[8]

An unthinking reliance on the past is an insubstantial basis for judgments about the future.

A demonstration of why it is that future wealth accumulation is driven by the arithmetic – and not the geometric – average is provided by Jon Christopherson, David Cariño and Wayne Ferson in a recent book.[9] With the authors’ permission, the relevant section is quoted in the appendix.

Conclusion: Look at the full picture

The volatility paradox exists even in a very simplified world: a strategy with a known, stable underlying return distribution and a long data series. No real-life strategy can supply past return data that comes even close to the quality and reliability of what I assumed in my example. Even with a complete description of the pattern of returns, we saw that allowing

[5] Behavioral economists have long since demonstrated the existence of loss aversion, i.e., that we tend to place greater importance on a loss than on an equal gain. That this strategy feels less attractive over a long time period can probably be ascribed at least in part to this effect.

[6] From “The Misuse of Expected Returns” by Eric Hughson, Michael Stutzer and Chris Yung, Financial Analysts Journal 62(6) (2006).

[7] The same relationship holds in our example strategy: the arithmetic – forward looking – average of 1% a year is 1.8% greater than the geometric average of –0.8%, while the difference between the arithmetic and geometric averages in the half-size version – which returns either 10% or –9% each year – is one-quarter as large: 0.45% a year.

[8] Yet another complicating factor lies in the fact that when historical experience is used as the basis for the assumed distribution of future returns, a systematic bias is introduced. This bias is a result of the skewness we have described. This leads to a tendency to overstate likely future wealth, and this becomes more significant when the data history is short or the projection period is long. See “Geometric or Arithmetic Return: A Reconsideration,” by Eric Jacquier, Alex Kane and Alan J. Marcus (Nov./Dec. 2003 *Financial Analysts Journal*), who conclude: “For typical investment horizons, the proper compounding rate is in between the arithmetic and geometric values.”

[9] Christopherson, J., Cariño, D., Ferson, W. “*Portfolio Performance Measurement and Benchmarking*,” McGraw-Hill, 2009.

one path to define our beliefs about the future –making the unspoken assumption that the past represents the most representative, or median, outcome – can lead to surprising conclusions.

This serves to remind that an unthinking reliance on the past is an insubstantial basis for judgments about the future. If you really want a sound basis for your decision, you need to do some serious thinking about the fundamentals of the proposed strategy. The less you know about an investment opportunity beyond past return data, the less confident you should be about your ability to model future behavior.

But what the volatility paradox really serves to highlight is the importance of understanding the whole distribution of outcomes. There are times when a simple measure – whether the median outcome, the mean outcome, standard deviation or any other – can be *too* simple. In particular, compounding returns over long periods of time creates a significant skewness in the ending wealth distribution. Neither axis of the typical "risk vs. return" charts that we are so accustomed to can necessarily capture all of the relevant information. At times, a fuller depiction is required.

Simple reliance on either the average or the median outcome can be misleading. As Mark Kritzman noted almost 20 years ago on the subject of the choice between arithmetic and geometric averages, "The correct answer depends on what it is about future value that we want to estimate."[10]

The example on which we based this paper is a series of bets, each of which is stacked in your favor but yet is more likely to lead to a loss than to a gain in the long term. To decide whether to embark on that strategy, the full distribution of possibilities should be considered: it is relevant that the median outcome generates a loss; it is relevant that the strategy generates a gain on average; the full distribution of paths and outcomes is relevant. It is convenient and necessary to summarize these possibilities in some way, but, when summarizing, some information is inevitably lost, and sometimes that information can be important. From William of Occam to Albert Einstein, great thinkers throughout the years have emphasized that things should be made as simple as possible, but no simpler. That's the moral of this story.

Acknowledgements

I am grateful to David Cariño for his patient review of this paper and for providing so much of the background material. I am now more in awe than ever of the extent of his learning.

[10] Kritzman's "What Practitioners Need to Know About Future Value" (*Financial Analysts Journal*, May-June 1994), from which this quotation is taken, is evidence that the basic points made in this paper have been understood, by some at least, for a long time.

Appendix

Extracts from “Portfolio Performance Measurement and Benchmarking,” by Jon A. Christopherson, David R. Cariño, Wayne E. Ferson: pp. 24, 26.

Expected Return

In order to form useful assumptions about future returns, statistical models might be employed. A very common modeling assumption is that future returns will be independent draws from some probability distribution – an assumption known as *independent, identically distributed* (i.i.d.) returns. If an initial wealth amount V_0 grows by simple compounding of the periodic returns,

$$V_T = V_0 (1+r_1)(1+r_2) \ldots (1+r_T),$$

then the independence assumption enables us to calculate the expected wealth $E(V_T)$ by replacing the returns in the preceding formula by their expected values. If the expected value of the identically distributed returns is $E(r_t) = m$, then the expected wealth can be calculated as

$$E(V_T) = V_0 (1+m)^T. \qquad (4.5)$$

This formula looks suspiciously similar to the formula for ending wealth, given the geometric average of *past* returns:

$$V_T = V_0 (1+r_G)^T.$$

Perhaps because of this similarity, an analyst might think that the geometric average return r_G is a good estimate of the expected value m. Unfortunately, this is not the case. The arithmetic average return is a better statistic to use as an estimate of m, because it is unbiased in the sense that $E(r_A) = (1/T)\ \Sigma\ E(r_t)$ is equal to m. As shown in the preceding, the geometric average return is less than or equal to the arithmetic average return by an amount that depends on the volatility.

…

As shown in Equation (4.5), the expected future wealth of the portfolio is related to the arithmetic average, not the geometric average. The fact that the geometric average increases with decreasing volatility does not imply that expected wealth also increases. If the geometric average return is plotted on a graph…, volatility affects *both* axes of the graph. To keep the concept of expected return separate from volatility of return, an arithmetic average return should be used for the return axis of this graph.

Author: Bob Collie

For more information:

Call Russell Investments at **800-426-8506** or
visit **russellinvestments.com/institutional**

Important information

Nothing contained in this material is intended to constitute legal, tax, securities, or investment advice, nor an opinion regarding the appropriateness of any investment, nor a solicitation of any type. The general information contained in this publication should not be acted upon without obtaining specific legal, tax, and investment advice from a licensed professional.

Please remember that all investments carry some level of risk, including the potential loss of principal invested. They do not typically grow at an even rate of return and may experience negative growth. As with any type of portfolio structuring, attempting to reduce risk and increase return could, at certain times, unintentionally reduce returns.

Russell Investments' ownership is composed of a majority stake held by funds managed by TA Associates with minority stakes held by funds managed by Reverence Capital Partners and Russell Investments' management.

Frank Russell Company is the owner of the Russell trademarks contained in this material and all trademark rights related to the Russell trademarks, which the members of the Russell Investments group of companies are permitted to use under license from Frank Russell Company. The members of the Russell Investments group of companies are not affiliated in any manner with Frank Russell Company or any entity operating under the "FTSE RUSSELL" brand.

First used: July 2011 (Reviewed for continued use: March 2017)

AI-25269-03-20