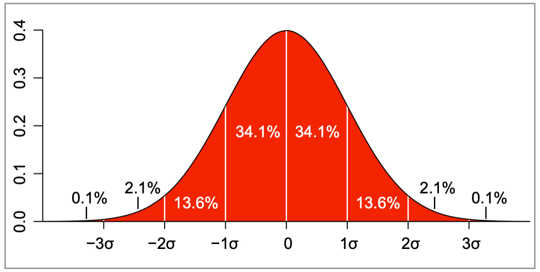
The Futility Of Volatility As Risk Metric

Christoph Horvay

There are few enough risk metrics that have become popular within financial services. Arguably there are only two: ‚volatility', and ‚beta'. In this essay, I will deal with volatility only, and show how dysfunctional this popular statistic is as a risk metric. There are four principle reasons why volatility does not serve the function it has been assigned.

Volatility is the investment community's synonym for a term borrowed from statistics: one standard deviation. According to the laws of normal distribution, and based on a large enough sample set, predictions are made in what section under the distribution curve any additional data can be expected. A schematic normal distribution is shown below. The peak of the curve is the median frequency. It divides all data in two halves: better, and worse than the median.



Each half is subdivided by one, two, and three standard deviations. Frequency (vertical axis) drops off sharply toward the extremes. It approaches zero, but only reaches it in infinity.

With the help of standard deviation, certain portions of the total sample population can be identified: 99.8% of data will be found inside three standard deviations, 95.6% inside two, and 68.2% within one standard deviation. These percentages refer to both sides of the median. The diagram shows Z-scores. With actual data, the median value will rarely be zero. When the x-axis uses percent as unit, care must be taken not to confuse portions of data with percentage points as unit.

Volatility Is Too General To Identify Risk

In finance, the term volatility is equal to the value of one standard deviation. While rarely written in that manner, it is bi-directional. Volatility is calculated from rates of change and expressed as a percentage value. Volatility may refer to daily, weekly, monthly, quarterly, or yearly changes. There is also no general rule on how many units of time are reflected. Since there is no uniformity, this information should be given. Very common is monthly volatility. Annual volatility is rarely encountered because that requires a fairly long history. Sometimes monthly volatility is annualised. But beware, to mathematically convert volatility from one time increment into another will generate a result different from directly measuring in that increment.

If an investment is shown to have a volatility of 1.3%, then we must remind ourselves that volatility is bi-directional, it is one standard deviation higher, or lower than the median and must be read as +/-1.3%. The median value itself is usually not shown.

Without the median, all we know from that volatility is that 68.2% of data are contained within a band that has a width of 2.6% (2x 1.3%). We do not know where on the axis that band is found. It could be located between +5.0% and +7.6%, or between -10.0% and -12.6%, or anywhere else. The same applies to two, and three deviations. Showing volatility while omitting the median value is a half-truth at best. It is useless. Worse than that, it is either rooted in a lack of understanding of the methodology and thus self-deceptive, or applied against better knowledge and thus malicious.

To illustrate, Table-1 contains four hypothetical values for volatility, calculated from 36 rates of change. If you embrace the doctrine that volatility equals risk, then you will conclude from Table-1 that:

Table 1

	A	B-1	B-2	C
Volatility	0.9%	0.4%	0.4%	0.4%

1.) A carries the most risk,
2.) B-1, B-2, and C have identical risk.

Forget numbers for a moment and think of bungee jumping. The risk in jumping is that you may fall to your death. You like to do it anyway, because you need a boost to your ego. A trusted friend theorises that the longer the rubber band to which you will be tied, the more you are going to bounce around and the scarier, and riskier, the jump will be. You may have guessed that the rubber band represents ‚volatility'. Following your expert friend's advice, you go for the safe option. From a selection of rubber bands of 30ft, 40ft, and 60ft length respectively, you opt for the safest: a mere 30ft. Next, and in anticipation of an enjoyable rush of adrenaline, you jump head first off the nearest bridge. You have little opportunity to consider subtle flaws in your own, and your friend's reasoning. Time flies, as do you, but you are headed in the wrong direction. Compared to your rubber band, the bridge was just low enough for you to not having to listen to that friend's advice again, ever. The next time an expert tries to convince you that volatility is a measure of risk, be mean. Insist on the median.

Volatility Is Not Designed To Identify Risk

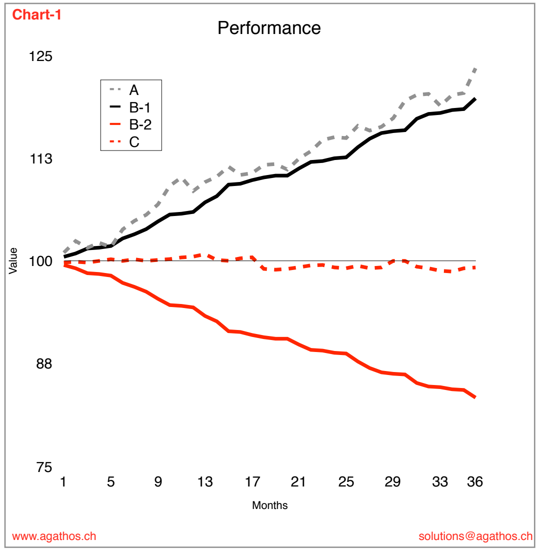
Volatility includes all rates of change and that makes it a perfectly sensible measure of the variability of returns. For the same reason it cannot also be a measure of risk. Chart-1 (overleaf) shows performance of the same hypothetical investments already seen in Table-1. For the sake of transparency, the underlying rates of change are shown in Appendix-1. Two of the examples, B-2 and C, (both shown in red), have not managed to rise above the inception value, while the other two, A and B-1, have produced a meaningful increase in value. According to the volatility paradigm, B-1, B-2, and C are supposed to have identical risk.

When comparing B-1 (solid black) and B-2 (solid red), you may notice their symmetry. Indeed, the rates of change are identical in value, only the signs have been reversed.

B-1 never suffers any loss, a case of extremely low risk, or the very successful control of risk. By contrast, B-2 only makes losses, an equally extreme example of risk affinity.

If you accept volatility as a measure of risk, you embrace a risk metric that is blind to even the most glaringly obvious distinction.

There is a conceptual problem with this metric. Volatility shows the portion of outcomes that have a particular, predetermined probability. But volatility does not identify probabilities of a particular, predetermined outcome.



Volatility Misrepresents Risk

My comments to this point are based on an opportunistically selective, not to say semi-educated use of normal distribution. The median value has already been mentioned. To include it in the assessment of data makes a big difference. Only with knowledge of the median is it possible to locate the band of one standard deviation on the x-axis.

At this point it may help to return to the schematic representation on page one. Median minus one standard deviation yields the lower inflection point (-1σ). Median plus one standard deviation arrives at the upper inflection (+1σ). An inflection point divides roughly ⅙ from the remaining ⅚ of the data.

Table-2 refers to the same examples used earlier. Only this time additional data is provided. Cast your eyes to median values and the inflection points.

Table 2

	A	B-1	B-2	C
Volatility	**0.9%**	**0.4%**	**0.4%**	**0.4%**
Median	0.8%	0.5%	-0.5%	0.1%
LIP	-0.2%	0.1%	-0.8%	-0.3%
UIP	1.7%	0.8%	-0.1%	0.5%
Higest	2.5%	1.3%	0.0%	0.8%
Lowest	-1.5%	0.0%	-1.3%	-1.4%

In B-2, one half of data will be worse than -0.5%, ⅙ even less than -0.8%. Approaching it from the other end, from the upper inflection down, ⅚ of changes will be worse than -0.1%. In B-1, ⅚ of values will be better than +0.1%. In the same vain, comparing A and C will reveal that in spite of the higher volatility, the LIP of A is slightly better than that of C, all thanks to the much higher median value, whereas in B-2 the median is negative.

These are very different qualifications of our four examples when you recall that, according to the volatility paradigm, B-1, B-2, and C all were supposed to contain equal risk. They don't!

Legitimacy, Or Illegitimacy Of Assumptions

To make deductions based on the laws of normal distribution can only be a productive exercise if a most fundamental assumption about the data holds true: that it is normally distributed. For it to be normally distributed, it must be random in origin. But is that actually the case?

Let's try to be courteous, and answer with yes. Then volatility would be a legitimate statistic, if not necessarily also a useful one. It's popularity seems to imply not only legitimacy but utility too. By implication, those who use it must believe the data random.

There is more than a hint of irony, plus a great deal of logical inconsistency found in that assumption: A multi-trillion-dollar sized, global industry deploys statistics that can be legitimate only if the results generated by that industry are as sophisticated as coin flipping. Can such a group of people be trusted to grasp the concept of value in any form at all? By using volatility as risk metric, the financial services industry arguably puts itself on par with dressed-up hyenas, scavenging for fees, and multitudes of revenue in other guises.

Mostly out of diligence, partly out of diplomacy, and partly because I prefer not to rest my own argumentation on leads taken from hyenas, I want to consider the alternative response: ‚No', financial market data is not random. That statement could indeed be the death knell for volatility.

But perfect normal distribution may not be needed. Then the case for volatility will rest on two factors: the degree of imperfection, and the knowledge of that degree, not generalised or assumed, but case by case.

Data distribution can deviate in two ways from normal: horizontally and vertically. Both types of distortion will alter the proportion of data found in any given area under the curve. A horizontal deviation (skew) is easily detected because it destroys the symmetry of the curve in favour of a slant. The simplest test to identify it is to compare median and mean values. If the two differ, then the data cannot be normally distributed. A vertical distortion (kurtosis) will likewise misrepresent any forecast made. A depressed curve has fewer data around the centre, a tall one more. Both types of distortion may be present in mild form and still permit application of the methodology. What seems odd then is that in financial services the value for volatility is never shown in conjunction with the values for skew, and kurtosis. The industry does not seem bothered by the absence of such due diligence.

One must consider the inner workings of financial markets, prior to articulating any expectation regarding patterns of distribution. While ultimately moved by facts, financial markets are also inherently bi-polar. They are driven by alternating cycles of greed and fear. Greed leads to an overly optimistic interpretation of fact (glass half full). Fear will make the same fact look bleak (glass half empty).

But facts change too. Sometimes the glass may be ⅞ full, at others ⅘ empty. When a nearly full glass is seen through the filter of optimism, things get quite euphoric. Except for the final stage of greed, bubbles tend to build slowly. Once a peak is reached, when there is no additional dope around to maintain illusions, any casual event will pierce the bubble. Then the excess tends to deflate more rapidly than it was built up. Obviously, the same applies to a fear scenario. But there is also such a thing as survivorship bias in financial markets.

Thus, a general expectation would be that negative rates of change are fewer and perhaps greater, positives changes frequent and perhaps moderate. There is no reason to expect financial market data to be randomly distributed. Even less so, if one considers that the durability of distinct cycles will hardly be uniform. Much depends on the length of time covered in any observation. If a data sample captures a full bull phase, together with a full correction, results will differ from another set of data, where portions of a preceding, or subsequent cycle are included. I think the distribution of data will be, above all, highly unpredictable in itself.

Even with financial markets moving according to fundamentals, daily rates of change will more likely be random in nature than annual ones. Over small increments of time, such as days, there will simply not be enough new information to cause changes other than random. Thus, given an equal sample size, short term data will more likely appear random, changes over longer increments of time less so, or not at all.

To illustrate that point with the help of life data, Table-3 contains daily, and quarterly rates of change for Pictet LPP40, a popular Swiss performance benchmark index used in the pension fund industry. For transparency, the underlying rates of change can also be found in Appendices 2 and 3, at the end.

Consider the skew evident from the data. Absolute distortion is arrived at by deducting the median value from the mean. Proportional distortion shows the same differential but expressed as percentage of volatility. How bothersome, or tolerable any distortions are, everyone must decide for themselves. Clearly, there is considerable uncertainty regarding the alleged accuracy, or lack thereof.

Table 3

Pictet LPP40 (2005)	50 Days	50 Quarters
Volatility	**0.2%**	**3.7%**
Median Rate of Change	0.1%	1.8%
Mean Rate of Change	0.1%	0.9%
Mean minus Median	-0.0%	-0.9%
Proportional Distortion	-6%	-24%

Conclusion

On balance, there are more and stronger reasons to discard volatility as a measure of risk than to keep it. It seems odd that volatility has been made so popular, considering the weaknesses this metric is prone to.

- The half-hearted, potentially semi-educated application of normal distribution under the label of ‚volatility' seems pointless.
- A responsible use of the laws of normal distribution may be considered a component of ex-post risk assessment. But even then, doubts remain that volatility is too general, too ambiguous and too unreliable.
- Volatility can only be seen as a measure of variability of returns, but not as a measure of risk. A host of other metrics are far better suited to that task.
- If the hypothesis of normally distributed rates of change in financial markets were true, then by implication, the entire financial services industry would be redundant.

If the ‚random origin' hypothesis is untrue, then the continued use of volatility as risk metric casts serious doubts on the proficiency, or the integrity of all those who embrace this metric.

The financial services industry may simply not be interested in greater transparency of the results it produces. For as long as clients accept mediocre standards, the industry has nothing to gain from the popularisation of any metric that permits a transparent qualitative analysis of performance.

Appendix 1

For the calculation of ‚performance' as seen in Chart-1, the rates of change shown in the tables below have been chained together vertically, using columns from left to right.

Data Example A

1.0%	1.0%	1.0%	0.1%	-0.1%	0.6%
1.4%	0.7%	0.6%	-0.6%	1.3%	0.1%
-0.8%	1.2%	1.1%	1.2%	-0.5%	-1.2%
0.6%	2.1%	-0.9%	0.8%	0.4%	1.1%
-0.5%	1.0%	0.2%	1.2%	0.9%	0.2%
2.1%	-1.5%	0.9%	0.3%	1.8%	2.5%

Date Example B-1

0.5%	0.5%	1.1%	0.2%	0.1%	1.2%
0.4%	0.6%	0.7%	0.0%	1.1%	0.5%
0.6%	0.9%	1.3%	0.8%	0.9%	0.1%
0.1%	0.8%	0.1%	0.7%	0.6%	0.3%
0.2%	0.1%	0.4%	0.1%	0.2%	0.1%
0.9%	0.2%	0.3%	0.3%	0.1%	1.1%

Date Example C

-0.2%	0.2%	0.3%	-0.1%	-0.1%	-0.7%
0.1%	-0.2%	-0.7%	0.1%	0.3%	-0.2%
-0.1%	0.1%	-0.1%	0.2%	-0.3%	-0.3%
0.2%	0.1%	0.3%	0.2%	0.1%	-0.1%
0.2%	0.2%	0.1%	0.1%	0.8%	0.4%
-0.2%	0.1%	-1.4%	-0.3%	0.0%	0.1%

Data Example B-2

-0.5%	-0.5%	-1.1%	-0.2%	-0.1%	-1.2%
-0.4%	-0.6%	-0.7%	0.0%	-1.1%	-0.5%
-0.6%	-0.9%	-1.3%	-0.8%	-0.9%	-0.1%
-0.1%	-0.8%	-0.1%	-0.7%	-0.6%	-0.3%
-0.2%	-0.1%	-0.4%	-0.1%	-0.2%	-0.1%
-0.9%	-0.2%	-0.3%	-0.3%	-0.1%	-1.1%

Table-3. The data in the graph below was taken from the performance of Pictet's Pension Fund Benchmark Index. Daily and monthly data may be downloaded from https://www.am.pictet/en/switzerland/indices/lppindices?index-code=lpp2005 .

Source: Pictet

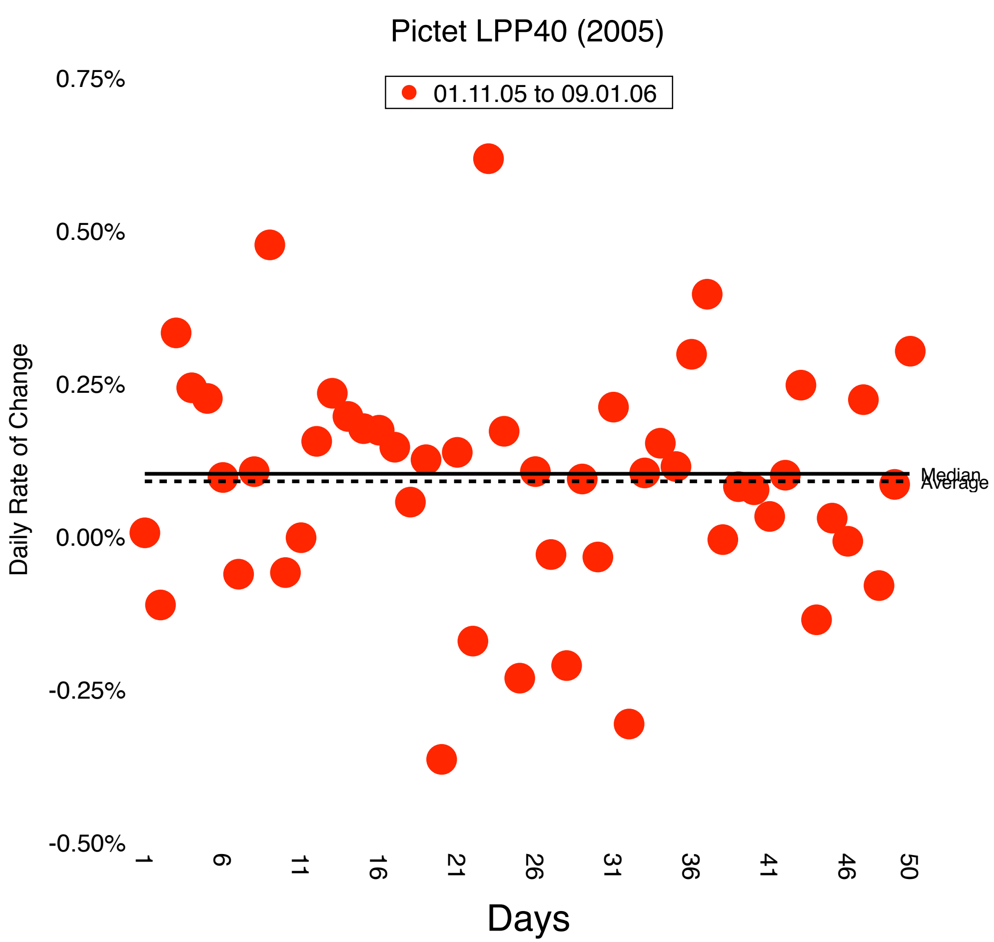


Table-3. The data in the graph below was calculated from the performance of Pictet's Pension Fund Benchmark Index. Daily and monthly data may be downloaded from https://www.am.pictet/en/switzerland/indices/lppindices?index-code=lpp2005 .

Source: Pictet