

A Critique of the Sharpe Ratio

By David Harding,
Winton Capital Management
1a St Mary Abbot's Place
London W8 6LS, United Kingdom.
Tel. +44 20 7610 5350
Fax. +44 20 7610 5301
www.wintoncapital.com

"Be careful what you wish for, as you may just get it..."

The Sharpe ratio is a statistic which aims to sum up the desirability of a risky investment strategy or instrument by dividing the average period return in excess of the risk-free rate by the standard deviation of the return generating process. Devised in 1966 as a measure of performance for mutual funds, it undoubtedly has some value as a measure of strategy "quality", but it also has several crucial limitations (see Sharpe 1994 for a recent restatement and review of its principles). Furthermore, its widespread and often indiscriminate adoption as a quality measure is leading to distortion of proper investment priorities, as investment firms manipulate strategies and data to maximise it.

The most basic problem with the Sharpe ratio is that whilst return is a definite and meaningful quantity (an "observable"), risk is not. It is true that standard deviation can be calculated from any time series of return data, but it is not at all true that its "meaning" will be the same for all time series. For the standard deviation to be a meaningful statistic at all the return time series must be generated from a process that is both stationary and parametric.

Stationarity. This means that each daily change in a time series is being generated by a process with broadly constant statistical characteristics (variance, skew, kurtosis). A common violation of this in track records occurs when a manager's style or leverage changes over the course of their history. It is common for a manager's early track record to differ significantly from later performance, due to his ability, for example, to exploit a particularly good anomaly with a small amount of money, or because when he had little to lose he could afford to "swing for the bleachers". An increase in assets under management can enforce constraints on such opportunities, leading almost inevitably to a change in investment strategy. A simple test of rolling standard deviation through time (or even average period change) may be sufficient to uncover the manifestations of the resulting non-stationarity. It is probably too demanding to ask that a manager's record show no style drift at all; markets are dynamic things and some change over time is unavoidable. However, efforts should be made to establish reasonable stationarity in any time series over which such statistics are going to be calculated.

Parametricity. This means that the process generating the returns should be capable of being characterised by a distribution with parameters which have a known meaning. In a normal distribution, for example, the standard deviation conveys quantitative information about dispersion (risk). In a non-parameterisable distribution, it conveys no information about anything. Non-stationary distributions are non-parameterisable. Some parameterisable distributions have curious characteristics. Cauchy-Paretian distributions have infinite variance – a Cauchy-Paretian return generating process would not be well assessed using the Sharpe ratio. Many investment strategies produce odd return distributions; some will produce many small profits and the occasional large loss - before the loss has appeared such strategies will have a high Sharpe ratio!¹ Other strategies like merger-arbitrage may well produce similar

¹ In a very instructive article on this subject, Goetzman et al. explored the use of option-like strategies for engineering high Sharpe ratios without the need for manager skill (2002).

sorts of distribution, vulnerable, like LTCM, to a major cataclysm. Some trading approaches which are swift to take profit but run losses may also twist the statistics to their advantage, while the serial correlation of monthly returns common in the more illiquid hedge fund strategies (e.g. convertible arbitrage) can lead to a substantial overstatement of Sharpe ratios (Asness et al. 2001; Lo 2001).

Sharpe Justice? Rewarding the right and punishing the wrong thing. Even if stationarity and parametricity criteria are met, the Sharpe ratio can have some perverse attributes. The Sharpe ratio appears at first blush to reward returns (good) and penalise risks (bad). Upon closer inspection, things are not so simple. The standard deviation takes into account the distance of each return from the mean, positive or negative. By this token, large positive returns increase the perception of risk as though they could as easily be negative, which for a dynamic investment strategy may not be the case. Large positive returns are penalised, and thus *the removal of the highest returns from the distribution can increase the Sharpe ratio*: a case of “reductio ad absurdum” for Sharpe ratio as a universal measure of quality! We might suggest an improvement by considering only the negative semi-standard deviation for the denominator, a measure known as the Sortino ratio. However it still remains vital that the semi-standard deviation used is meaningful, in the sense that it is calculated from a sufficiently well-understood return distribution, where the assumptions of stationarity and parametricity can be made.

A second “reductio ad absurdum” scenario is one in which the return distribution produces a string of very small but consistent profits which will produce a very high Sharpe ratio with very little return. An example would be investing in short-term AAA commercial paper, which should consistently produce a little over risk-free (e.g. government paper), but where the Sharpe ratio would not be a good measure of “skill” at all. While this weakness might be picked up in other tests, it may not be, even in quite a long data sample. Indeed, when faced with these sorts of time series, the attraction of raw *return* as a statistic reasserts itself. Who would prefer to make 0.01% over risk-free each month rather than 1% with a standard deviation of 1%?

These “reductio ad absurdum” scenarios may seem like extreme cases but they do point out the danger of relying slavishly on any single measure for assessing product quality for risky investments. Furthermore, practitioners will know that none of the dangers that are pointed out here are theoretical. They are all mirrored in the (faulty) decision-making processes of investors who overvalue short run consistent returns and biased return-generating processes. Suggested solutions include paying more attention to stationarity and parametricity in making estimates of future return and risk from manager track records; flexibility in design of appropriate statistics for measuring quality for investment strategies, more use of semi-variance than variance as a risk measure; and a revival of return as a quality statistic (albeit *long term* return subject to some risk constraint).

References

- Asness, C., R. Krail & J. Liew 2001 ‘Do Hedge Funds Hedge?’ Working Paper, AQR Capital LLC.
- Goetzman, W., J. Ingersoll, M. Spiegel & I. Welch 2002 ‘Sharpening Sharpe Ratios,’ Working Paper.
- Lo, A.W. 2001 ‘The Statistics of Sharpe Ratios,’ forthcoming in *Financial Analysts Journal*.
- Sharpe, W.F. 1994 ‘The Sharpe Ratio’, *Journal of Portfolio Management*.