



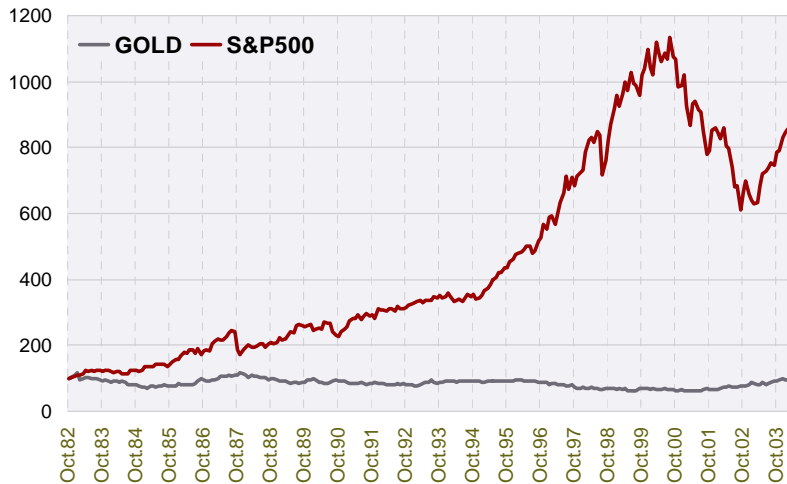
## **SORTING THE SHEEP FROM THE GOATS: DETERMINATION OF GOOD FROM BAD INVESTMENTS USING QUANTITATIVE MEASURES**

**David Harding**  
**Winton Capital Management**  
**1-5 St Mary Abbot's Place**  
**London W8 6LS**  
[www.wintoncapital.com](http://www.wintoncapital.com)

'Past Performance is Not an Indication of Future Results': like most boilerplate text, this warning is often consigned to the margins of investors' consciousness, its ability to pack a punch eroded by constant repetition. We occasionally need to be reminded that the truth contained in this statement is both profound and complex. Often the most accessible means of assessing an investment comes in the form of past performance, and some valuable information may be gleaned from it through cautious assessment. However, potential investors need to remain aware of the limitation and pitfalls of some of the most commonly used assessment techniques. Here we will be taking a fresh look at some of the most frequently used quantitative measures, exposing some of their commonly overlooked logical weaknesses and suggesting some ways of rehabilitating them into the investor's toolbox.

### **FORECASTING RETURN**

Most investments that investors are called upon to evaluate can be assessed using the representation of their history in the form of a time series. Both amateurs and professionals are inclined to evaluate certain quantitative characteristics of such time series for the purpose of judging whether the underlying investment is a "good" one or not; that is, whether it will produce positive returns over future years at a level of risk they can understand and accept. The statistics that they choose to calculate and evaluate need to be as well suited to this purpose as possible; they need to not only capture information about the past but also to be capable of providing a means for forecasting the salient features of the future of the time series, a point that is often forgotten, overlooked or wilfully misrepresented.



**FIGURE 1.** Which is the better investment?

Let us say we would like to compare the investment potential of two venerable investments, the US stock market and gold based on their historical performance (Figure 1). The return of each time series can be measured in many different ways; last year, 2 years, 3, 5, etc, or average rate of return (compound or otherwise) over the investment's life. Each of these measures can give very different comparative results (Figure 2).

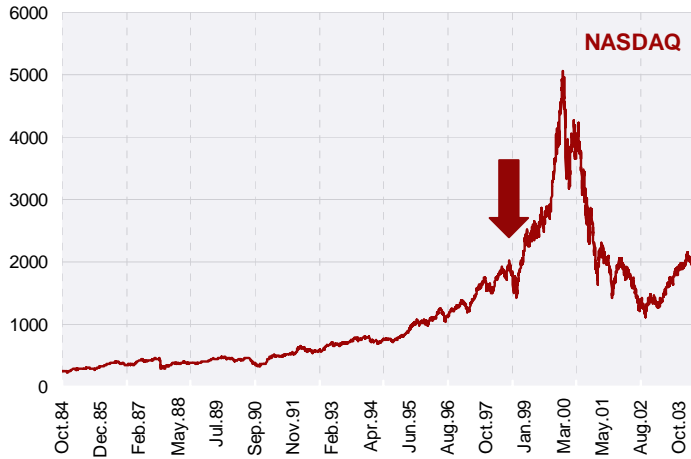
	<b>GOLD</b>	<b>S&amp;P 500</b>
Total Return	<b>-9%</b>	<b>728%</b>
Last 1 Year Return	<b>14%</b>	<b>21%</b>
Last 3 Year Return	<b>25%</b>	<b>3%</b>
Last 5 Year Return	<b>35%</b>	<b>-17%</b>
Average Annual Return	<b>0.5%</b>	<b>10.9%</b>
Compounded Annual Average Return	<b>-0.4%</b>	<b>10.3%</b>

**FIGURE 2.** What measures to use?

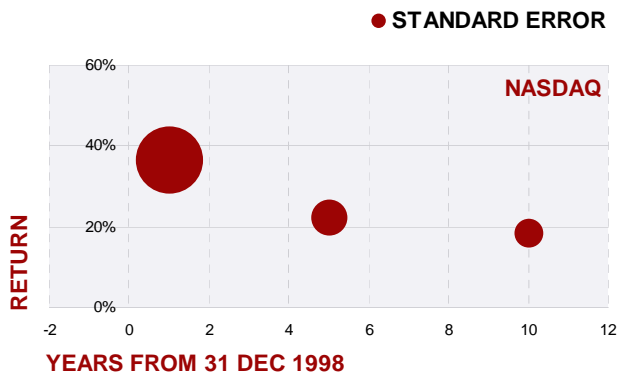
### RETURN, ERROR & RISK

For historical return to be a useful statistic for forecasting future return the time series should, as a rule, be as long as possible. That is because return as a statistic often has a high error associated with it; this error is what is often described as the risk. High returns in particular are usually inextricably linked with higher risks and ironically it is higher risks that are more predictable from high returns than are future high returns. This problem is particularly great with shorter term returns. The error in forecasting the future return is proportional to the inverse of  $\sqrt{t}$  where  $t$  is the length of the time series. Thus, a time series history of 4 times the length will halve the error in the return forecast. For typical investments, e.g. stock funds, a reasonable long run estimate of expected return and risk might be 10% annually with a standard deviation of 20%. However, the last year's return from a particular fund might be +80% with an annualised monthly standard deviation of 100%. A four year series would reduce that standard deviation to 50%, a 9 year series to 33% and a sixteen year series to 25%; starting to approach the reasonable long run estimate. Clearly the 80% return is the last statistic that one should reasonably extrapolate into the future and yet both professional and amateur investors regularly make this crude mistake.

Let us take the case of an investor tempted to invest in the tech stocks at the end of 1998 by the streak of recent returns (Figure 4). Figures 5-6 show the kind of information they would have derived from analysing different time windows, and, in retrospect, the bearing it might have had on the fate of the investment.

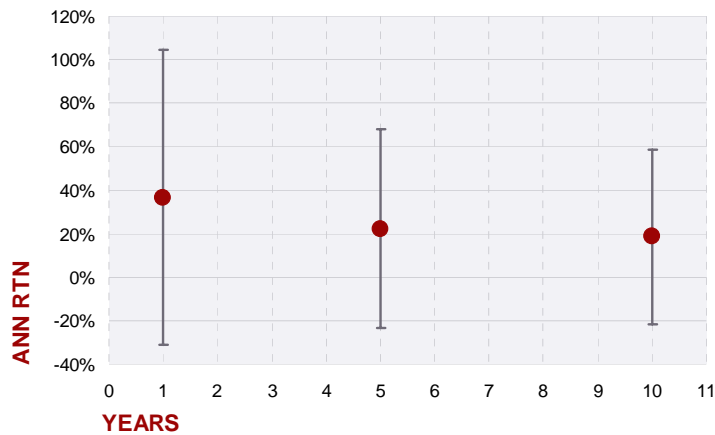


**FIGURE 4.** Point of assessment.



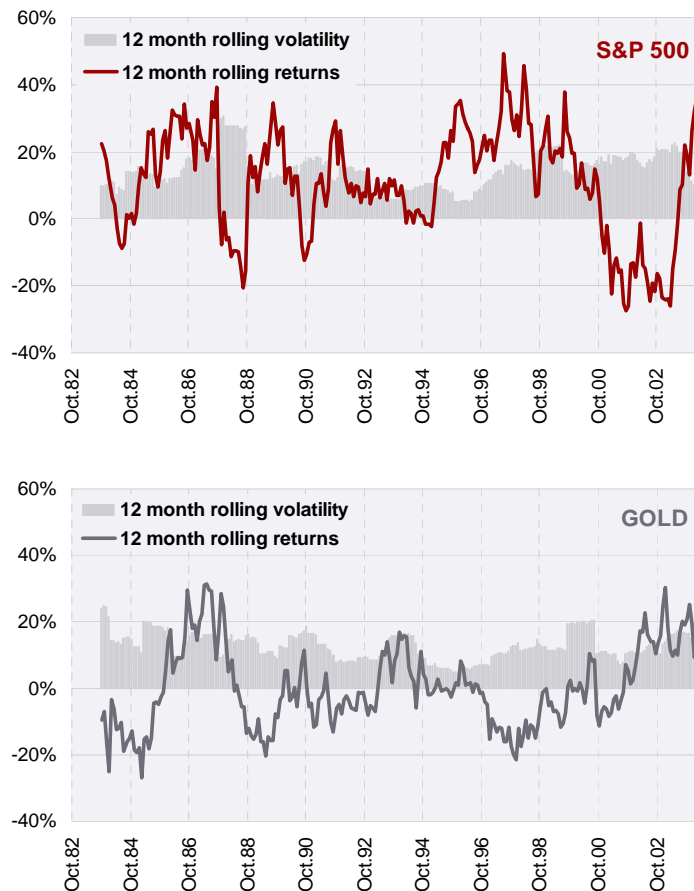
	Mean	StDev	SE
1 Year	37%	26%	26%
5 Years	22%	18%	8%
10 Years	19%	16%	5%

**FIGURE 5.** The longer the time series the better the estimate (1): Mean return and error.



**FIGURE 6.** The longer the time series the better the estimate (2): Mean and 99% confidence intervals.

It is also worth bearing in mind that an estimate of the error on returns is likely to be more accurate than an estimate of the returns themselves (Figure 7). The error associated with a return forecast coming from a historic time series remains substantial even when the return is estimated from 10-20 years of historic time series data. In sum, it is safe to assume that large positive historical returns imply large negative future returns as well as large positive ones, i.e. one can expect high volatility in both directions, rather than exclusively large positive future returns.

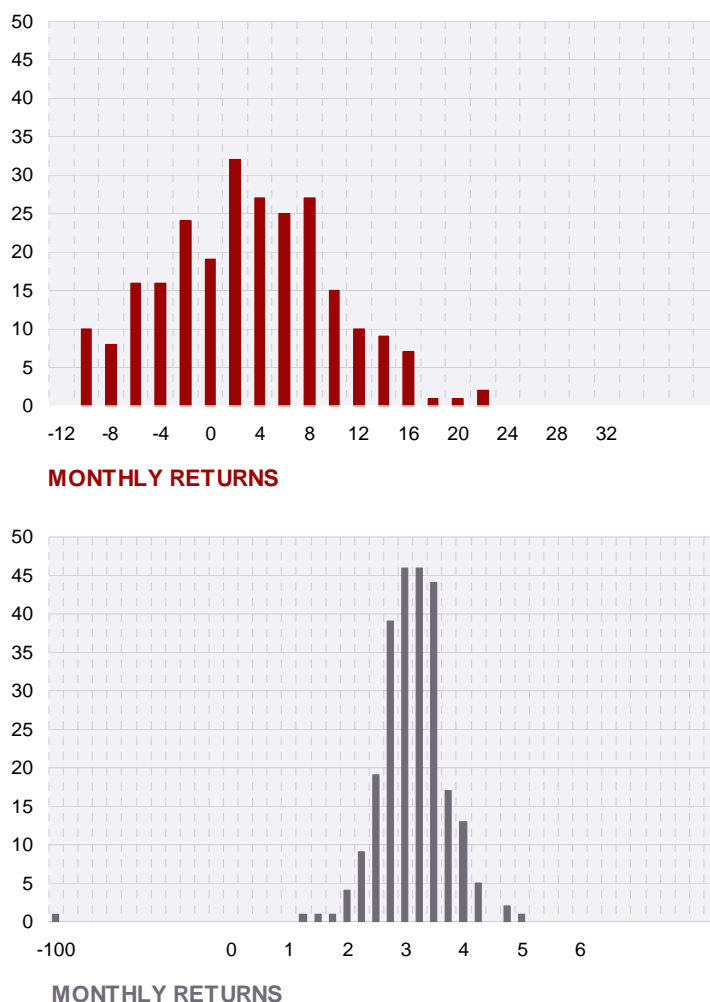


**FIGURE 7.** It is easier to estimate error than return.

## RISK ADJUSTED RETURN

All of this is known to professional money managers which is why they have devised better time series statistics for measuring investment quality. The key step forward was to calculate the amount of return per unit of risk taken, or, put differently, to standardise the return in risk units. The most popular derivation of this concept is the Sharpe ratio. The Sharpe ratio is calculated for a time series by dividing the mean period return (daily, monthly, yearly), in excess of the risk free rate, by the standard deviation of such returns. The Sharpe ratio overcomes some of the problems inherent in the pure return statistic.

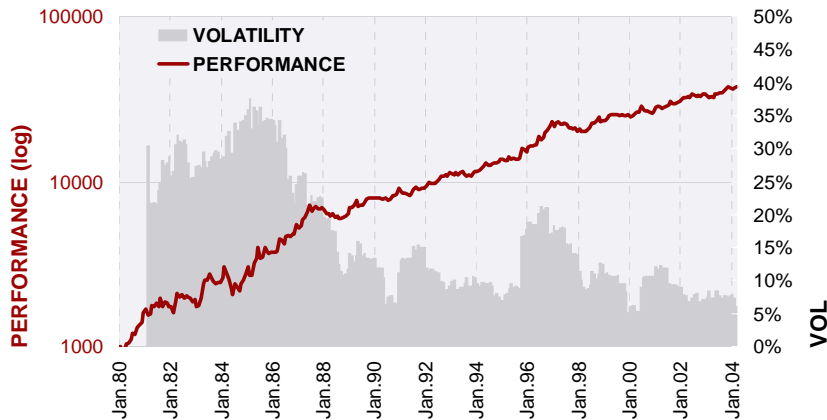
The Sharpe ratio, however, suffers from a number of drawbacks as a statistic of which it is wise to be aware (see Harding 2003 for a critique). Firstly, the denominator is standard deviation which is only a reliable and meaningful statistic for time series where the distribution of the first differences (price changes) is both parametric and stationary. "Parametric" implies that the distribution can be characterised by a known and meaningful distribution (e.g. normal, binomial, T, etc) with finite variance. Some financial time series do not satisfy this criterion, e.g. option granting strategies which produce lots of small profits and occasional large losses. In such cases, the Sharpe ratio would not give an accurate representation of the investment's risk/return profile. (see e.g. Figure 8).



**FIGURE 8.** These distributions have the same Sharpe ratio.

Other return distributions might be bi- or multi-modal, and their standard deviation may give a very misleading impression of the probability of certain events.

“Stationary” implies that the time series volatility, and higher moments, remain constant through time. This criterion would typically not be satisfied where the investment strategy or assets underlying the time series have changed too much through time. A common example would be a hedge fund starting with a high leverage in order to produce impressive returns, then gearing down on maturity to ease liquidity constraints and collect management fees (Figure 9).



**FIGURE 9.** This is not a stationary return process.

A simple test for stationarity is to ensure that there is no major trend in the rolling volatility. A related problem is that the underlying form of many distributions is unknown. Lower credits hedged with higher ones will tend to pick up a steady excess over risk free resulting in very high Sharpe ratio, presumably at the expense of occasional very large losses, such as incurred by LTCM.

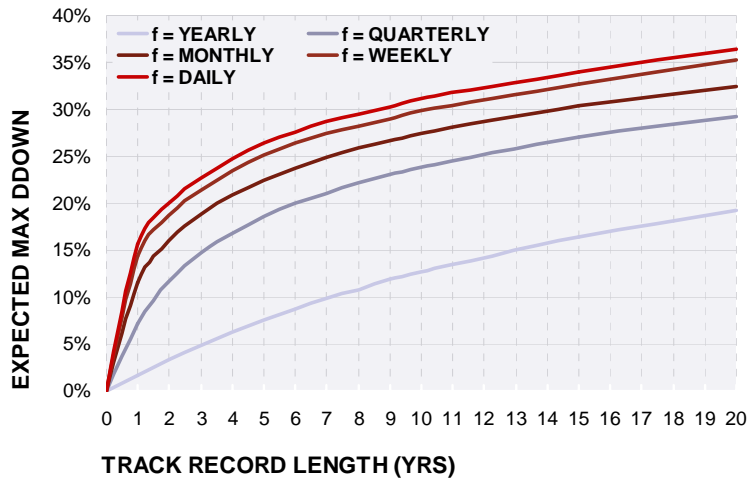
Another problem for Sharpe ratio is that it is symmetric regarding upside and downside risk. High returns have the effect of increasing the value of the denominator (standard deviation), and lowering the value of the ratio. Conversely, for a positively skewed return distribution such as that of a managed futures strategy, the Sharpe ratio can be increased by removing the largest positive months. This is patently absurd. Finally, as with return, there is an overwhelming issue of data bias. Bias can be introduced into a time series in all sorts of innocuous ways as well as deliberately; but regardless of its source, it has the effect of undermining the value of any statistic as a forecasting tool. Thus, all retrospectively constructed time series, portfolio backtests etc. must be viewed askance from the point of view of calculating forecast statistics. Similarly, short time series are more vulnerable to statistical bias; this is the well known point about coin tossing chimpanzee-like fund managers.

Finally, the Sharpe ratio is not invariant to its periodicity: there is of course a difference between the annualised monthly Sharpe ratio and the annual Sharpe ratio, which may be exacerbated by a variety of factors including the liquidity of the underlying investment (Lo 2001; Getmansky *et al.* forthcoming). High Sharpe ratios from relatively small data histories can be a result of entrepreneurs “gaming the system” rather than evidence of great future earning potential (Goetzman *et al.* 2002).

## DRAWDOWN

One statistic that is used to try to overcome the parametricity issue is drawdown. The drawdown is the maximum peak to subsequent trough fall in a time series; the maximum loss the investor has experienced from a previous high (which, in the worst instance could have been when he invested and thus represent his worst possible loss). This does not rely on the return process having any particular form and does have intuitive physical appeal. It also,

however, has important weaknesses of which it is necessary to be aware (Harding *et al.* 2003). First, for two time series with otherwise equal characteristics the longer will tend to have the greater drawdown. For most investable time series longevity would be presupposed to be a good thing implying survival, robustness, experience, etc. All other things being equal however, a longer track record implies a larger worst drawdown.

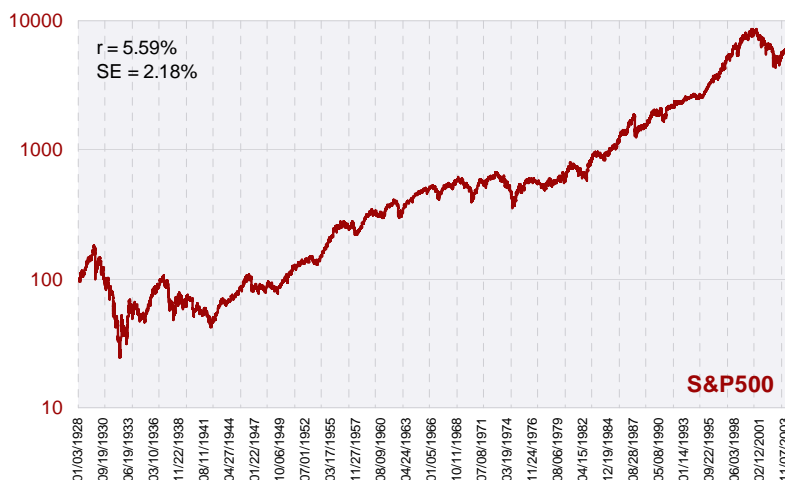


**FIGURE 10.** The longer the track record, the deeper the drawdown.

Secondly, maximum drawdown is a single number and will therefore have a large and uncertain error distribution. Thus we cannot be at all sure that a time series with a larger worst drawdown is being produced by a return generating process which will tend to produce worse drawdown! Essentially, by using a single number as the denominator we are balancing too much inferential weight on too slender a quantity of data.

Another statistic that is used to address the skewness problem is that of Sortino ratio. This is the mean period return divided by the semi-standard. This statistic does not demand symmetry of profits and losses and is a better measure than Sharpe ratio for time series resulting from dynamic investment strategies such as managed futures.

## BACK TO RETURN



**FIGURE 11.** Long term returns from US equities.

Even solidly built risk-adjusted statistics are not a panacea. They still leave us with the problem of underestimating the value of decent returns compared to strategies that consistently produce just a little over the risk free rate. We have recently suggested some amendments to traditional measures to remedy this problem. In the first instance, the Sortino ratio can be adjusted to incorporate a *minimum acceptable return*. This ratio utilises the mean return in excess of the minimum acceptable hurdle divided by the semi-variance with respect to that (minimum acceptable) return. This is a more useful statistic for real life situations, such as those faced by the promoter of an investment product bearing several percentage points per annum of fixed fees or the trustee of a pension fund that requires a certain level of returns in order to meet its liabilities. A more complex mathematical statistic named Omega tackles the issues of parametricity in a robust manner, but involves more complex computations (Keating & Shadwick 2002a,b; Winton Research Papers, 2003).

Both Omega and the modified Sortino ratio redirect much-needed attention to *pure return*, but in a more constrained fashion. In the end the market recognises the value of the very long term compound annual average rate of return (Warren Buffet's 26% for 40 years, the US stock market's 5% in real terms for 200 or so: Figure 11) in forming judgments. To produce 5% per annum for 20 years is not that exceptional or extraordinary; to produce 15% is very good. Yet the former process can easily produce much higher Sharpe and Sortino ratios and lower drawdown. Without an understanding of the statistical concepts of populations and samples, parametric distributions, statistical moments and their errors, investors are doomed to carry on being misled by their intuition into making one mistake after another.

## REFERENCES

- Getmansky, M., A.W. Lo & I. Makarov *forthcoming* 'An Econometric Model of Serial Correlation and Illiquidity in Hedge Fund Returns', *Journal of Financial Economics*.
- Goetzman, W., J. Ingersoll, M. Spiegel & I. Welch 2002 'Sharpening Sharpe Ratios,' Working Paper, Yale School of Management, International Center for Finance.
- Harding, D.W. 2003 'Sharpe Justification?' *Hedge Funds Review* (July 2003).
- Harding, D.W., G. Nakou & A. Nejjar 2003 'The Pros and Cons of "Drawdown" as a Statistical Measure of Risk for Investments' *AIMA Journal* (April 2003).
- Keating C. & W.Shadwick 2002a "A Universal Performance Measure", *The Finance Development Centre Limited*.
- Keating C. & W.Shadwick 2002b "An Introduction to Omega", *The Finance Development Centre Limited*.
- Lo, A. W. 2001 'The Statistics of Sharpe Ratios,' forthcoming in *Financial Analysts Journal* 58.4:36-52.
- Winton Research Papers 2003:  
'Assessing CTA Quality with the Omega Performance Measure'  
[www.wintoncapital.com/omega.htm](http://www.wintoncapital.com/omega.htm)  
'Case Studies of CTA Assessment Using the Omega Performance Measure'  
[www.wintoncapital.com/omega\\_casestudies.htm](http://www.wintoncapital.com/omega_casestudies.htm)

## Legal Disclaimer

This publication has been prepared by Winton Capital Management Limited (“Winton”), which is authorised and regulated by the Financial Services Authority (“FSA”). Winton is a London based global hedge and futures fund manager, which offers specialised investment management services to qualified investors.

The value of investments and any income generated may go down as well as up and is not guaranteed. You may not get back the amount originally invested. **Past performance is not necessarily a guide to future performance.** Changes in exchange rates may have an adverse effect on the value, price or income of investments. There are also additional risks associated with investments in emerging or developing markets

The information contained in this publication is subject to updating and verification and may be subject to amendment. No representation or warranty, expressed as to the accuracy of the information contained in this document and no liability is given by Winton as to the accuracy of the information contained in this publication and no liability is accepted for any such information.