Portfolio skew and kurtosis

by David Buckle

There have been many occurrences of extreme returns from the fund management community, but of course only the negative ones make the headlines. The explanation for such events is frequently that financial asset returns exhibit extreme values more often than expected by the normal distribution.[1]

In studying the assets themselves, and making inferences about the likelihood of extreme values of portfolio returns, the implicit assumption is that the portfolio positions are constant through time.[1] In practice they are not, and regardless of the distribution of the financial asset returns, the portfolio returns might exhibit more extreme values than the financial assets themselves purely because the positions are changing.

In this article, we examine the skew and kurtosis of portfolio residual[2] returns, emphasising the difference between asset and portfolio. With the assumption that the asset residual returns are normally distributed, when portfolio weights are fixed there is no difference between the skew and kurtosis of assets and portfolios (that is, skew = 0 and kurtosis = 3). But when portfolio weights vary through time, as they do for actively managed portfolios, portfolios can exhibit high kurtosis even if the assets themselves do not – solely because of the weight volatility. This is most pronounced in portfolios of few assets, and is also affected by the quality of the position taking.

Buckle (2004) showed that, assuming the Treynor & Black (1973) security selection model at each investment period with positions derived from modern portfolio theory[3], the skew and kurtosis of the resultant residual portfolio returns over multiple investment periods are given as:

$$S = \Delta^{\frac{1}{2}}\left(8\Delta+6\right)/n^{\frac{1}{2}}\left(2\Delta+1\right)^{\frac{3}{2}} \qquad (1)$$

and:

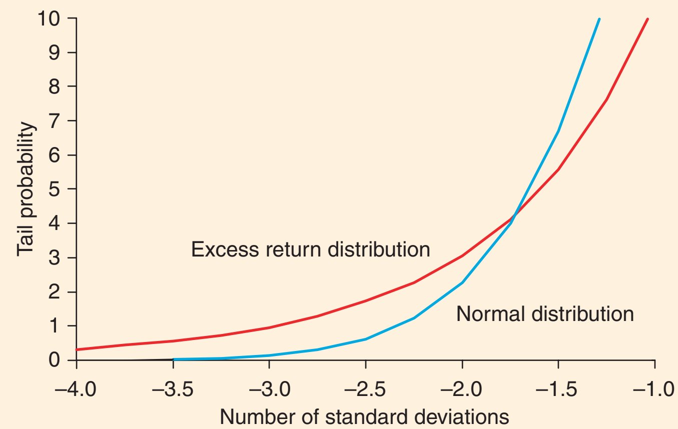
$$K = \left(12\Delta\left(\Delta+1\right)\left(1+4/n\right)+3\left(1+2/n\right)\right)/\left(2\Delta+1\right)^{2} \qquad (2)$$

where $\Delta = IC^2/(1 - IC^2)$, with IC denoting the information coefficient[4], and n denotes the number of assets in the universe (not including the market itself, which we assume to be an asset available to the investor).

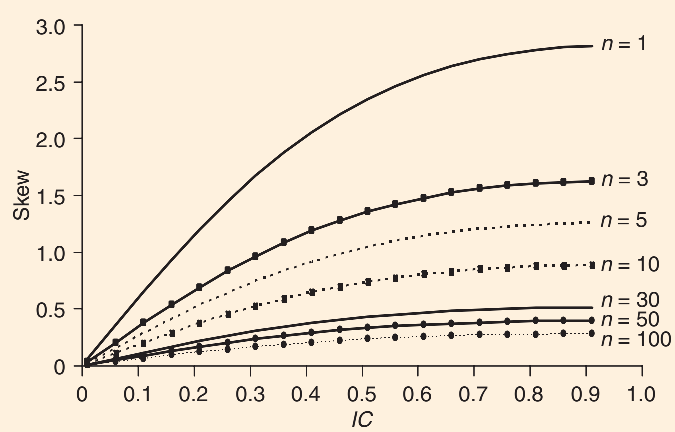
Buckle does no more than cite these results so we further that work with a brief study here. Equations (1) and (2) are not readily interpreted as they stand, but there are some interesting special cases. We observe that as the number of assets increases, skew tends to zero and kurtosis tends to three. This result holds for any level of IC. It is an appealing result because, if we have enough assets in our portfolio, the portfolio returns should exhibit no skew and kurtosis in excess of that expected by the normal distribution.

The situation is less comforting when the number of assets is small. At the limit when there is just one asset (in addition to the 'market asset'), that is, $n = 1$, as IC decreases so skew tends to $6 \times IC$ and kurtosis tends to nine. With IC typically less than 0.1 in practice, we can see that the residual portfolio return distribution will not be particularly skewed, but will exhibit kurtosis three times as large as that expected by the normal distribution. In fact, in a simulation of Buckle's framework, we found that the frequency of a negative three-standard-deviation event or worse is expected by the normal distribution to be one per 1,000, but was observed to be one per 100. Put plainly, the normal distribution tells us to expect a three-standard-deviation event once every four years, where the actual expectation is once every four months. The normal distribution says we should

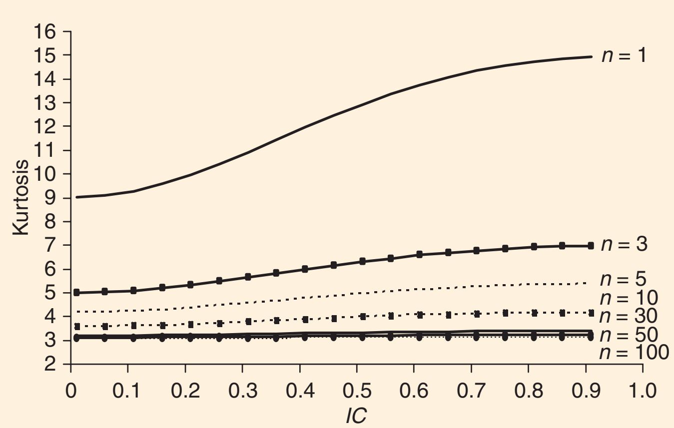
1. Tail probabilities of excess returns relative to the normal distribution



2. Skew in active portfolio returns (*n* assets)



3. Kurtosis in active portfolio returns (*n* assets)



expect a negative four-standard-deviation event once every 100 years, where the actual expectation is once a year. In figure 1, we illustrate the tail probabilities for a range of events. The increase in probability of an extreme event once we get beyond 1.75 standard deviations can clearly be seen to be greater than that expected by a normal distribution.

As IC increases, both skew and kurtosis increase. Increasing skew should

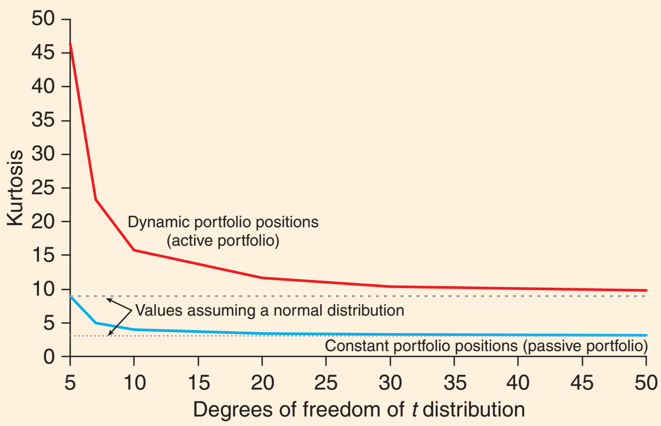
[1] See, for example, Simons (2000). This is a general summary article on value-at-risk, including an example fund analysis. We use it as a reference of our statements throughout this article. We also use the data tabulated in it in our empirical example

[2] By residual return, we mean the non-market component of return in a capital asset pricing model-style factor model

[3] The underlying mathematical model that Buckle (2004) used to derive equations (1) and (2) is summarised in the appendix. The key references are Buckle (2004), Markowitz (1952) and Treynor & Black (1973)

[4] IC is defined as the correlation between forecast and return, and therefore a measure of forecast ability. See Grinold & Kahn (1999)

4. Active and passive portfolio kurtosis with Student-*t* distributed asset returns



be expected because portfolio returns clearly move toward a chi-squared distribution with perfect foresight (which exhibits skew). At its limit, kurtosis gets as large as 15. To some extent, we are less concerned with kurtosis when *IC* is high, because most of the extreme events will be positive ones. However, it is worth noting that for any level of *IC*, we never have less kurtosis than when *IC* is zero, and therefore kurtosis is always greater than $3(1 + 2/n)$. Figures 2 and 3 show how skew and kurtosis both decrease with the number of assets, but increase with forecasting ability.

There is some evidence that asset returns are more fat-tailed than a normal distribution[1], so we simulate Buckle's framework with the normal distributions replaced by Student-*t* distributions. As is evident in figure 4, the same characteristic of active portfolios exhibiting much larger kurtosis than assets still prevails. As we know, when the degree of freedom is small, asset kurtosis is larger than the normal distribution, but that characteristic is exaggerated in the active portfolio. Additionally, as the degree of freedom gets large so the Student-*t* distribution tends to the normal distribution, and we recover our previous results. Perhaps more importantly, when we simulated portfolios consisting of many securities, the skew and kurtosis tended to zero and three respectively, regardless of the fact that the individual asset returns had higher kurtosis. This is expected because under the model assumptions, it is clear that the excess returns for each asset are independently identically distributed and therefore the central limit theorem (where the sum of the returns from any distribution tends to a normal distribution) comes into effect for large asset universes.

To provide some empirical support for our work, we use Simons' fund data.[1] Simons considered a US mutual fund that was benchmarked against the S&P 500, and observed that its returns between May 11, 1999 and October 10, 2000 had a kurtosis of 25 whereas the S&P 500 had a kurtosis of 4.9 over the same period. Unfortunately, we do not have details on what type of fund this is, but its kurtosis suggests that it is probably highly concentrated, so we take $n = 1$. Plugging this into equation (2) shows our model estimate of fund kurtosis should be equal to nine. In this example, we still underestimate the fund kurtosis. However, if we take the S&P 500 to be *t*-distributed with seven degrees of freedom, rather than normally distributed, thereby accounting for the fact that the S&P itself had a high kurtosis of 4.9 over the sample period, then our model estimates a fund kurtosis of 23. This is close to the sample kurtosis of 25.

In conclusion, we caution against the measurement of tail behaviour of assets, and infer from that the tail behaviour of active portfolios. We have seen that even if the asset returns follow a normal distribution the portfolio returns can exhibit substantially more extreme event risk than a normal distribution. When the asset returns follow a fatter-tailed distribution, active portfolios can have very high kurtosis indeed. Active portfolios with many assets are less likely to exhibit kurtosis than concentrated ones. ■

Appendix. The mathematical model

For reference, we provide the mathematical model that Buckle (2004) used to derive equations (1) and (2). Much of this is taken directly from Buckle's article.

Let the column vector $\mathbf{r}_t$ denote the returns (in excess of the risk-free rate) from our universe of assets over the single time period $(t, t + 1)$. We decompose these returns into a market component and an uncorrelated residual component, so that:

$$\mathbf{r}_t = \mu_t + \theta_t$$

where μ_t is the market component, and θ_t is the vector of single-period residual returns.

We suppose that the fund manager makes a forecast of the residual returns and denote the forecast pertinent to time t by the vector $\mathbf{m}_t$. We relate θ_t and $\mathbf{m}_t$ by the equation:

$$\theta_t = \mathbf{m}_t + \varepsilon_t$$

where ε_t is the error in our forecast.

☐ **Assumption.** $E[\theta] = E[\mathbf{m}] = E[\varepsilon] = 0$ and the variances of θ, $\mathbf{m}$ and ε are time-invariant and diagonal, and $\mathbf{m}$ and ε are independent and normally distributed.

For each time period, our framework is identical to the Treynor & Black (1973) security selection model. We have simply assumed a sequence of these models such that the average single-period alpha over time is zero (so it is not possible to beat the market unless positions are varied through time).

The information coefficient, denoted *IC*, is the correlation between the residual return of an asset and the forecast, thus:

$$IC = E[m\theta] / \left(V[m]V[\theta]\right)^{\frac{1}{2}}$$

where m and θ are an element of $\mathbf{m}$ and θ respectively.

☐ **Assumption.** IC is equal across all assets.

The active portfolio positions are denoted by $\mathbf{y}_t$ and are a function of our forecasts.

☐ **Assumption.** Modern portfolio theory is used to convert forecasts into positions, so that $\mathbf{y}_t = \mathbf{V}^{-1}\mathbf{m}_t$ where $\mathbf{V}$ denotes the covariance matrix of ε.

Using this notation, the active portfolio return is $\mathbf{y}^T\theta$, and it is the skew and kurtosis of this random variable that we are interested in. ■

REFERENCES

Buckle D, 2004
Some aspects of active portfolio management
Unpublished manuscript

Grinold R and R Kahn, 1999
Active portfolio management: a quantitative approach for providing superior returns and controlling risk
McGraw-Hill, second edition

Markowitz H, 1952
Portfolio selection
Journal of Finance 7, pages 77–91

Simons K, 2000
The use of value at risk by institutional investors
New England Economic Review, November–December, pages 21–30

Treynor J and F Black, 1973
How to use security analysis to improve portfolio selection
Journal of Business 46, pages 66–86

David Buckle is head of currency management at Merrill Lynch Investment Managers in London. The views expressed in this article are those of the author and do not necessarily represent those of Merrill Lynch Investment Managers. Email: david_buckle@ml.com