REPRINTED FROM

Risk

RISK MANAGEMENT | DERIVATIVES | REGULATION

RISK.NET AUGUST 2012

Momentum strategies offer a positive point of skew

Cutting edge: Asset management

Momentum trading: 'skews me

An attractive characteristic of momentum strategies is that they generate a distribution of trading returns that is positively skewed. This can mean fewer winning than losing trades, but still positive expected returns. ***Richard Martin*** *and* ***David Zou*** *derive explicit formulas for the skewness of the trading returns, for certain simple types of model, and show that it has a characteristic term structure.*

In a very interesting paper, Potters & Bouchaud (2005) point out that the question: "what fraction of your trades are winning?", often asked of fund managers, is largely irrelevant to fund performance. For trend-following strategies, it is well known that small trading losses are common, but occasional big gains are produced when the strategy levers itself into a trend. The longevity of trend-following funds suggests that this characteristic has served them well over the years, and this leads us to suppose that the fraction of winning trades is positively misleading. Studies on the subject have generally been empirical (for a good overview, see, for example, Till & Eagleeye, 2011, and references therein).

Potters & Bouchaud (2005) take a particular trend-following system, based on a trending signal equal to an exponentially weighted moving average (EMA) of past returns, or equivalently the difference between the spot price and an EMA of past prices. The signals to buy and sell are based on the passage of this signal through pre-defined levels. For a market that follows a geometric Brownian motion they derive the distribution of profit and loss analytically in terms of the confluent hypergeometric function, and show that it is positively skewed.

In this article, we take a slightly different approach. We consider a more general form of trending indicator, and rather than using it as a trigger, we simply take a position in the market that is directly proportional to it. We do not claim that this strategy is in any way optimal, and nor do we claim that it is better than the one considered by Potters & Bouchaud (2005). The construction here gives analytical advantages, but in practice, and as pointed out by a referee, it has the disadvantage of putting on too large a bet when momentum is strong; capping the signal removes that problem, but makes the subsequent analysis less tractable.

We directly evaluate the third moment of the distribution of trading returns, without making any assumption about the market returns beyond their first three moments – in particular, we are not assuming normality. For pure trend-following systems, the skewness is always positive. We are also able to analyse mixed systems that have trending and counter-trending behaviour.

We show that the skewness of the returns depends strongly on their period, so that even if the one-day returns have no skewness, longer-period returns may be skewed. This may at first seem curious, but arises because successive daily trading returns are not independent. So it is possible to obtain a skewed distribution by adding non-skewed random variables, if those variables are appropriately dependent.

The skewness of a distribution does not immediately address the question of whether it exceeds its mean more or less than half the time. However, one can see diagrammatically that a distribution with positive skewness is likely to have the property that its median is less than its mean, or equivalently that the probability of exceeding its mean is less than ½. More substantially, this property can be formalised with the Gram-Charlier expansion, which estimates the probability as $\frac{1}{2} - \kappa_3/(6\sqrt{2\pi})$ where κ_3 is the coefficient of skewness. Take, for example, the exponential distribution: the exact probability of exceeding the mean is $e^{-1} \approx 0.368$; the Gram-Charlier approximation gives $\frac{1}{2} - 2/(6\sqrt{2\pi}) \approx 0.367$, which is very close. On the other hand, there may be considerable divergence for other distributions: for example, the third moment may not even exist, or one may have a right-skewed distribution with negative third moment but positive higher odd moments. Nonetheless, the skewness is potentially much easier to obtain and analyse, so we concentrate on it here, deriving some new analytical formulas, in equations (2) and (5).

The expected return of momentum strategies arises from serial correlation of market returns, and is perhaps best described as a result of the way information is disseminated into markets or of behavioural characteristics of market participants. It is, however, an unrelated mechanism to the skewness: even if market returns are independent and identically distributed, during which period the strategy will produce no expected return, the trading returns will still have positive skewness, as we show here. This is an important effect because it causes momentum strategies to hold on to most of their previous profits during the periods where they are not making money. As pointed out by Till & Eagleeye (2011), this 'long-option behaviour' distinguishes them from other strategies that tend to have higher Sharpe ratios, the implication being that this is a form of remuneration for negative skewness. Incidentally, it is possible to be mathematically precise about the sense in which momentum strategies are 'long options', rather than just falling back on a phenomenological argument, and we explain this later on.

Model setup

We define U_{n+1} as the return per unit volatility for the asset X between time n and $n + 1$, that is:

$$U_{n+1} = \frac{X_{n+1} - X_n}{\hat{\sigma}_n}, \quad \hat{\sigma}_n = \mathbf{E}_n\left[\left(X_{n+1} - X_n\right)^2\right]^{1/2}$$

with $\mathbf{E}_n$ denoting an expectation conditional on $\mathcal{F}_n$. The reason for dividing by $\hat{\sigma}$ is that we wish U_n to be dimensionless and appropriately normalised.

Following the general principle that one should bet a number of contracts (or contract notional) inversely proportional to the contract volatility[1], we define the position in the asset to be $\varphi_n/\hat{\sigma}_n$ at time n, where $\varphi_n \in \mathcal{F}_n$ is a function of any or all of the U's up to and including U_n, so it cannot cheat by looking ahead. Clearly,

[1] *This is essentially to keep a reasonably constant level of risk on: if the same position is held while the volatility rises substantially, one is likely to break one's market risk limits. More formally, this follows from stochastic control theory (see, for example, Björk, 1998, chapter 14)*

the profit and loss arising from the period between time n and time $n+1$ is $\varphi_n U_{n+1}$.

For much of this article, we assume that the risk-adjusted returns U_n are independent and identically distributed and of zero mean, though the raw returns may not be because of stochastic volatility. So, we are studying the behaviour of the strategy under the assumption that it is not generating any expected return (Potters & Bouchaud, 2005, do the same). We also assume that U_n has zero third moment, so that its first three moments are 0, 1, 0. So, although U_n has no skewness, we are about to show that the trading return may have. We make no further distributional assumptions about the U_n.

The M-period trading return is defined as:

$$Y_n^{(M)} = \sum_{k=0}^{M-1} \varphi_{n+k} U_{n+k+1}$$

The first moment of the trading return is clearly zero. The second moment is given by:

$$\left\langle \left(\varphi_0 U_1 + \cdots + \varphi_{M-1} U_M\right)^2 \right\rangle$$

Let us consider this expectation as a sum of products. The cross-terms all vanish because each contains a U_{n+1} term multiplied by a term in $\mathcal{F}_n$. This leaves the squared terms, which give simply:

$$M\left\langle \varphi^2 \right\rangle$$

(as $\langle U^2 \rangle = 1$). The proportionality in M is a consequence of the trading returns being uncorrelated (note not necessarily independent).

The third moment of the trading return is given by:

$$\left\langle \left(\varphi_0 U_1 + \cdots + \varphi_{M-1} U_M\right)^3 \right\rangle$$

Expanding this, we obtain four types of terms: (i) $\varphi_{n_1} U_{n_1+1} \varphi_{n_2} U_{n_2+1} \varphi_{n_3} U_{n_3+1}$ with $n_1 < n_2 < n_3$; (ii) $\varphi_n^2 U_{n+1}^2 \varphi_m U_{m+1}$ with $n < m$; (iii) $\varphi_n^2 U_{n+1}^2 \varphi_m U_{m+1}$ with $n > m$; and (iv) $\varphi_n^3 U_{n+1}^3$. The independence of the U's and the assumptions about their moments show that (i), (ii) and (iv) all vanish. We are therefore left with (iii), which can be written:

$$3 \sum_{0 \le m < n < M} \left\langle \varphi_n^2 \varphi_m U_{m+1} \right\rangle \tag{1}$$

As φ_n^2 may depend on U_{m+1}, this expression is not necessarily zero, though it must be when $M = 1$, as the sum in (1) is empty because of the symmetry of the market returns. The factor of 3 comes from the three ways of permuting the indexes in (iii).

Linear systems

We now specialise these results to linear systems, by which we mean:

$$\varphi_n = \sum_{j=0}^{\infty} a_j U_{n-j}$$

Linear systems have several advantages. They are easily constructed, for example through EMAs (for which $a_j \propto \alpha^j$), which can be implemented recursively. They are also easily added together, so that one can combine the momentum of different periods, or weight momentum of certain periods, negatively to capture counter-trending behaviour. Finally, analysis is reasonably straightforward, and the moments of the trading returns can be captured using the coefficients a_j alone.

We shall need the autocorrelation function of the impulse response:

$$R_k^a = \sum_{j=0}^{\infty} a_j a_{j+k}, \quad k \ge 0$$

and the system function, that is, the z-transform of the weights:

$$A(z) = \sum_{j=0}^{\infty} a_j z^{-j}, \quad z \in \mathbb{C}$$

The function must be bounded for $|z| \ge 1$. A system built from EMAs always has a rational system function and its poles are usually a key part of the design and analysis. For a general account, refer to Haykin (1989).

The simplest example is the single-EMA case, which we will call EMA1: $a_j = \alpha^{j+1}$ and $A(z) = \alpha/(1-\alpha z^{-1})$. This arises as the difference between the spot price and an EMA of past prices, and is used by Potters & Bouchaud (2005). The decay-factor α is linked to the effective period of the EMA, N, by $\alpha = 1 - N^{-1}$, so that the EMA becomes progressively slower, or more highly smoothed, as $\alpha \to 1$.

Another important example is the difference of two expressions of this form, which we call EMA2: $a_j = (\alpha^{j+1} - \beta^{j+1})/(\alpha - \beta)$ and $A(z) = 1/((1-\alpha z^{-1})(1-\beta z^{-1}))$. This arises as the difference between two EMAs of prices, a common device in technical analysis (see Herbst, 1992, chapter 9). It has less day-to-day variation than EMA1, because it is the difference of two smoothed prices.

It is convenient to define a class of systems that we call SPRZ ('simple poles, regular at zero'). The precise conditions are: $A(z)$ bounded for $|z| > 1 - \varepsilon$ for some $\varepsilon > 0$; the only singularities to be simple poles; and $A(z)$ regular at the origin. These should be thought of as mild analytical conditions that enable the ready application of residue calculus; systems with multiple poles can be understood as limiting cases of systems with single poles as the poles coalesce.

Linear systems: second and third moments

It is immediate that the second moment of the M-period trading return is MR_0^a. For the third moment, we have to find the $U_{m+1}U_j$ term ($j \neq m + 1$) in φ_n^2 in the expression (1). This is:

$$2\sum_{k=0,k\neq n-m}^{\infty} a_{n-m}a_k U_{m+1}U_{n-k+1}$$

This now has to be multiplied by $3\varphi_m U_{m+1}$ and the expectation taken. Thus it is necessary to look for any overlap between U_{n-k+1} and φ_m, and so in the k-summation we only need terms with $k \geq n - m$, and exclude the others. The resulting expression emerges as:

$$\begin{aligned} &3\sum_{0\le m<n<M} 2\sum_{k=n-m}^{\infty} a_{n-m-1}a_k a_{m-n+k} \\ &= 6\sum_{0\le m<n<M} a_{n-m-1}R_{n-m}^a \\ &= 6\sum_{k=1}^{M-1}(M-k)a_{k-1}R_k^a \end{aligned} \qquad (2)$$

If we understand a pure momentum, or trend-following, strategy to be one in which all the a_j are positive, then by (2) the trading returns must be positively skewed. By the same token, a counter-trending strategy, with the a_j's negative, has a negatively skewed return distribution even if the market returns are symmetrical.

We can z-transform (2), that is, multiply the kth term by z^{-k} and sum from $k = 1$ to ∞, to get:

$$G_3(z) = \frac{6z}{(z-1)^2}\sum_{k=1}^{\infty} a_{k-1}R_k^a z^{-k}$$

From the presence of a double pole in $G_3(z)$ at $z = 1$ we deduce that the third moment is asymptotically $6M\Sigma_{k=1}^{\infty} a_{k-1}R_k^a$ as $M \to \infty$. Recalling that the second moment is linear in M, we deduce that as a function of the return period M the skewness starts from zero, reaches a maximum somewhere and decays as $M^{-1/2}$.

The $M^{-1/2}$ asymptotic form is, intriguingly, the same as that observed in Lévy processes. However, the origin of the skewness is completely different. With a Lévy process, it arises because the one-period returns are asymmetrical but independent – here, they are symmetrical but not independent.

Further analysis of the third moment

The asymptotic third moment (without the $6M$ pre-factor) is $\Sigma_{k=1}^{\infty} a_{k-1}R_k^a$, which is also equal to:

$$\frac{1}{2\pi i}\oint_{|z|=1} A(z)A(z^{-1})^2\,dz$$

To see this, write the integrand as a product of Taylor series; the integral pulls out the z^{-1} term. This expression can be calculated using residue calculus, if we restrict ourselves to the SPRZ case, as:

$$\sum_j \rho_j A(\alpha_j^{-1})^2$$

Meanwhile, the second moment is MR_0^a, and:

$$R_0^a = \frac{1}{2\pi i}\oint_{|z|=1} A(z)A(z^{-1})z^{-1}dz = A(0)^2 + \sum_j \rho_j \alpha_j^{-1} A(\alpha_j^{-1}) \qquad (3)$$

Collecting the results together, we deduce that the skewness of M-period trading returns, for large M, is:

$$\kappa_3^{(M)} \sim \frac{6\sum_j \rho_j A(\alpha_j^{-1})^2}{\left(A(0)^2 + \sum_j \rho_j \alpha_j^{-1} A(\alpha_j^{-1})\right)^{3/2} M^{1/2}} \qquad (4)$$

In the EMA1 case, we immediately obtain:

$$\kappa_3^{(M)} \sim \frac{6\alpha}{(1-\alpha^2)^{1/2} M^{1/2}} \sim 3\sqrt{2}\left(\frac{N}{M}\right)^{1/2}$$

where the right-hand expression is obtained by assuming that $N = (1 - \alpha)^{-1}$ is large. In the EMA2 case, the poles are at α, β and are of residue α^2, $-\beta^2$, and the result is, after a little algebra:

$$\kappa_3^{(M)} \sim \frac{6(\alpha+\beta)(1-\alpha\beta)^{1/2}}{(1-\alpha^2)^{1/2}(1-\beta^2)^{1/2}(1+\alpha\beta)^{1/2} M^{1/2}} \sim 3\sqrt{2}\left(\frac{N_\alpha + N_\beta}{M}\right)^{1/2}$$

We can also return to (2) to get the exact third moment, not just the long-term asymptotic form. To do this, we write (2) in terms of $A(z)$, as:

$$\frac{6}{(2\pi i)^3}\sum_{k=1}^{M-1}(M-k)\sum_{j=0}^{\infty}\oint\oint\oint A(y)y^{j-1}A(z)z^{j+k-1}A(w)w^{k-2}dw\,dy\,dz$$

in which the contours for w- and y-integrals are of radius $1 - \varepsilon$ and the contour for z is just $|z| = 1$. The j-summation and the y-integral can be done immediately – the placement of the contours means $|yz| < 1$, which is necessary for convergence of the sum; in doing the y-integral, we expand the contour out to ∞ and pick up the residue at $y = 1/z$ on the way. Next, do the k-summation using the identity:

$$\sum_{k=1}^{m-1}(m-k)r^{k-1} \equiv \frac{r^m - 1 + m(1-r)}{(1-r)^2}$$

to arrive at:

$$\frac{6}{(2\pi i)^2}\oint\oint A\left(z^{-1}\right)A(z)A(w)\frac{(wz)^M - 1 + M(1-wz)}{(1-wz)^2}w^{-1}dw\,dz$$

The term that is linear in M exactly generates the large-M result we have already obtained, once the w-integral is done (again, by expanding the contour out to ∞ and picking up the residue at $w = 1/z$). The remaining part can be calculated by collapsing the w-contour around all the singularities inside the unit circle. Note that no singularity arises from the $(1 - wz)^2$ term and then collapsing the z-contour in the denominator, as $|wz| < 1$. In the SPRZ case, we finally obtain the third moment as:

$$\begin{aligned} &6M\sum_j \rho_j A\left(\alpha_j^{-1}\right)^2 - 6A(0)\sum_j \rho_j A\left(\alpha_j^{-1}\right) \\ &-6\sum_{j,k}\rho_j\rho_k\alpha_j^{-1}A\left(\alpha_k^{-1}\right)\frac{1-\alpha_j^M\alpha_k^M}{\left(1-\alpha_j\alpha_k\right)^2} \end{aligned} \tag{5}$$

For EMA1, the exact expression for the skewness is therefore:

$$\kappa_3^{(M)} = \frac{6\alpha}{\left(1-\alpha^2\right)^{1/2}M^{1/2}}\left(1 - \frac{1-\alpha^{2M}}{1-\alpha^2}M^{-1}\right) \tag{6}$$

This rises from zero to a peak and then rolls off as $O(M^{-1/2})$ (see figure 1a). The maximum skew[2] is roughly 2.1–2.4, and occurs for the period $M \approx 1.1N$ (recall $\alpha = 1 - N^{-1}$).

For EMA2, the exact expressions for the second and third moments are:

$$\mu_2^{(M)} = \frac{M(1+\alpha\beta)}{(1-\alpha\beta)\left(1-\alpha^2\right)\left(1-\beta^2\right)} \tag{7}$$

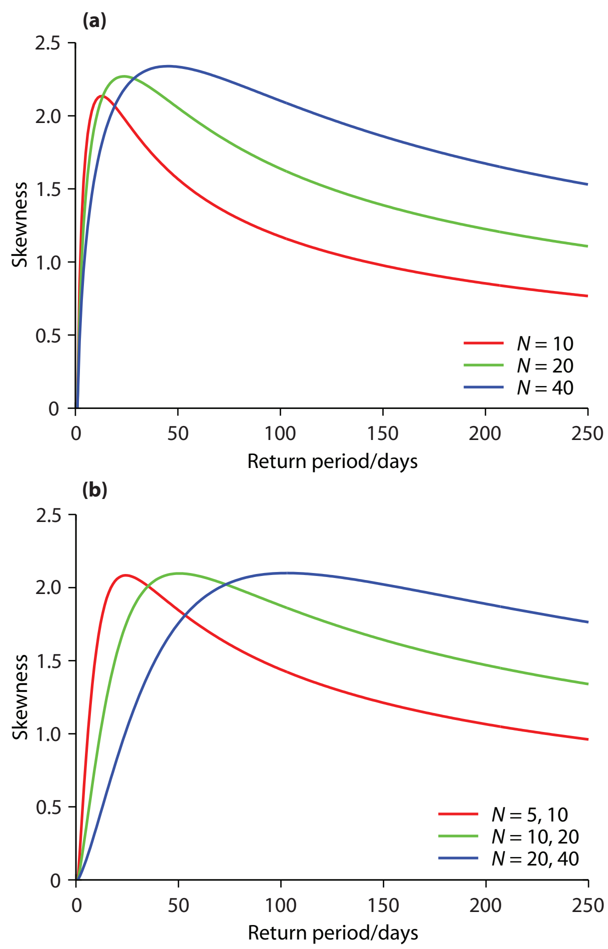
$$\begin{aligned} \mu_3^{(M)} = &\frac{6M(\alpha+\beta)(1+\alpha\beta)}{(1-\alpha\beta)\left(1-\alpha^2\right)^2\left(1-\beta^2\right)^2} \\ &+\frac{6\alpha^3\left(1-\alpha^{2M}\right)}{(\alpha-\beta)^2\left(1-\alpha^2\right)^3(1-\alpha\beta)} + \frac{6\beta^3\left(1-\beta^{2M}\right)}{(\alpha-\beta)^2\left(1-\beta^2\right)^3(1-\alpha\beta)} \\ &-\frac{6\alpha\beta^2\left(1-\alpha^M\beta^M\right)}{(\alpha-\beta)^2\left(1-\beta^2\right)(1-\alpha\beta)^3} - \frac{6\alpha^2\beta\left(1-\alpha^M\beta^M\right)}{(\alpha-\beta)^2\left(1-\alpha^2\right)(1-\alpha\beta)^3} \end{aligned} \tag{8}$$

and $\kappa_3^{(M)} = \mu_3^{(M)}/(\mu_2^{(M)})^{3/2}$ as usual. This is qualitatively similar to EMA1 (see figure 1b). The maximum skewness is around 2.1, and occurs at period $M \approx 1.7(N_\alpha + N_\beta)$, provided N_α and N_β are not too far apart. In the extreme case where either of the N's is equal to one, we recover EMA1. The limit $\beta \to \alpha$ is well behaved, but algebraically messy and omitted here. In essence, (5) telescopes the various geometric series that are implicit in the calculation of (2), and allows it to be done with an amount of computational effort independent of M.

Demonstration

For a demonstration using real data, we use two datasets: the Swiss franc/US dollar futures and the S&P 500 futures. We use

1 Skewness of trading returns, as a function of period, for (a) EMA1-type model and (b) EMA2-type model. Note the characteristic shape



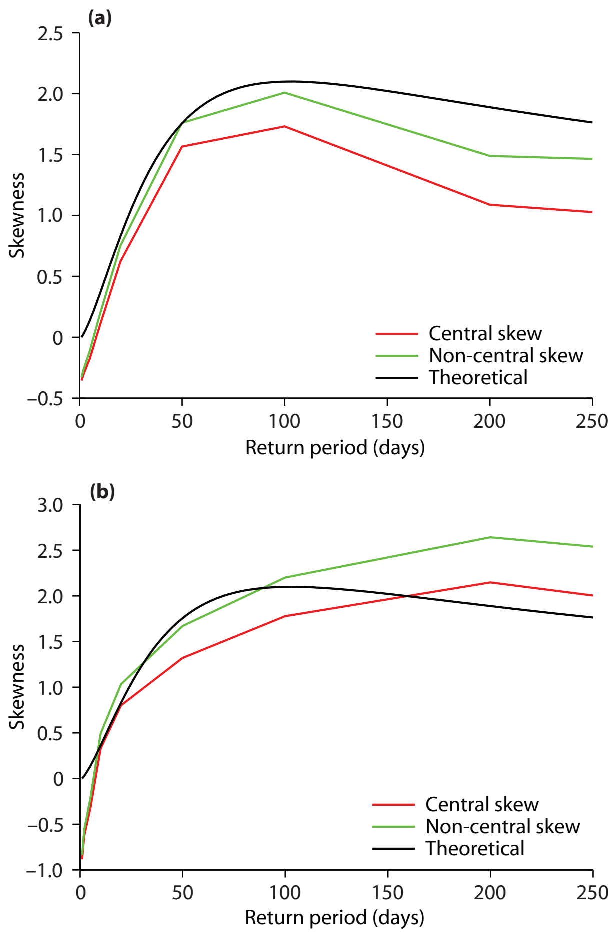
Source: Bloomberg

Bloomberg's SF1 Curncy and SP1 Index, rolled 10 days before expiry to create generic series. The data range is January 1, 1990 to December 31, 2009. For risk-adjusting the returns, we use a 20-day EMA of squared price changes to estimate the volatility ($\hat{\sigma}_n$ in the definition of U_n). We are using an EMA2 with $N = 20, 40$.

It is worth recalling the assumptions that we made in deriving our formulas: (i) independence of the risk-adjusted returns U_n; and (ii) symmetry of their distribution up to the third moment. In practice, the first clearly does not hold, because it implies that momentum strategies do not generate positive expected return, whereas the evidence is that on average they do. That means that when we examine real data, the observed skewness of returns may well not equal the theoretical result, by virtue of the mean being different. We therefore plot the central skewness (third central moment divided by the 3⁄2 power of the second central moment) and also the 'non-central' skewness (raw third moment divided by 3⁄2 power of the raw second moment). If the effect of trending is to generate a slightly positive expected return but keep the other moments roughly equal, then the non-central skewness will

[2] *Unless N is small, we can approximate (6) as* $(3/\sqrt{2x^3})(e^{-2x} - 1 + 2x)$, *with* $x = M/N$; *the maximum of this function is* ≈ 2.41 *and occurs at* $x \approx 1.07$

2 Skewness of trading returns, as a function of period, for (a) Swiss franc/US dollar and (b) S&P 500 futures. Theoretical result also shown. $N = 20, 40$



Source: Bloomberg

be fractionally higher than the central skewness. As to (ii), we know that equity markets occasionally have very negative returns.

Figure 2 shows the results for the two markets, superimposing also the theoretical result from figure 1b. In spite of the deficiencies in the modelling assumptions, the agreement is not bad and the general shape is right. The short-term skewness for the equity market is non-zero because of the asymmetry of the market returns; the higher long-term skewness is best ascribed to the particularly good trending behaviour in the mid-1990s generating high trading returns. The skewness of the trading returns is far higher than that of the underlying markets: the latter is typically about 0.0 for Swiss franc/US dollar and −0.2 for S&P 500 to one decimal place. This shows that the skewness comes entirely from the momentum strategy. The Gram-Charlier formula for the probability of exceeding zero is modified to $\Phi(\tau) - \kappa_3/(6\sqrt{2\pi})$ in the presence of non-zero expected return, where $\tau = \kappa_1/\kappa_2^{1/2}$ is the Sharpe ratio and Φ is the normal cumulative distribution function. For horizons M in the range of 100–200 days the Sharpe ratio of each is roughly +0.2 and the skewness is around 2, so this gives the probability of exceeding zero as about 0.45, which corresponds well with the empirical value – note that it is less than one-half.

Option-like nature of trend-following

As pointed out in Till & Eagleeye (2011), trend-following systems are often thought to have a long-option-type payout because of the positive skewness. For linear systems, this can be formalised as follows. The M-period trading return is:

$$Y_n^{(M)} = \mathbf{u}'\boldsymbol{\Gamma}\mathbf{u}, \quad \mathbf{u} = \begin{bmatrix} U_{n+M} & U_{n+M-1} & \cdots \end{bmatrix}'$$

where the symmetric matrix $\boldsymbol{\Gamma}$ is given by:

$$\boldsymbol{\Gamma} = \frac{1}{2}\begin{bmatrix} 0 & a_0 & a_1 & a_2 & \cdots & & \\ a_0 & \ddots & \ddots & \ddots & \ddots & \ddots & \\ a_1 & \ddots & 0 & a_0 & a_1 & a_2 & \cdots \\ a_2 & \ddots & a_0 & 0 & 0 & 0 & \cdots \\ \vdots & \ddots & a_1 & 0 & 0 & 0 & \cdots \\ & \ddots & a_2 & 0 & 0 & 0 & \cdots \\ & & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{9}$$

where the dots indicate that there are M rows of the same form as the top one. The moments of $Y_n^{(M)}$ relate to the spectrum of $\boldsymbol{\Gamma}$, and direct calculation reveals:

$$\left\langle Y_n^{(M)} \right\rangle = tr(\boldsymbol{\Gamma}) = 0, \quad \left\langle \left(Y_n^{(M)}\right)^2 \right\rangle = 2tr\left(\boldsymbol{\Gamma}^2\right), \quad \left\langle \left(Y_n^{(M)}\right)^3 \right\rangle = 8tr\left(\boldsymbol{\Gamma}^3\right)$$

Writing $\boldsymbol{\Gamma}$ in terms of its eigenvalues γ_j and normalised eigenvectors $\mathbf{e}_j$, we have an expression that is a weighted sum of 'orthogonal quadratic bets', that is, squared linear combinations of returns, which are like straddle payouts but have constant convexity:

$$Y_n^{(M)} = \sum_j \gamma_j \left(\mathbf{e}_j \cdot \mathbf{u}\right)^2 \tag{10}$$

Note the weights sum to zero as $tr(\boldsymbol{\Gamma}) = 0$. Now $tr(\boldsymbol{\Gamma}^r) = \Sigma_j \gamma_j^r$, so the moments of $Y_n^{(M)}$ relate to the moments of the eigenvalue distribution. It is easy to see the rank of $\boldsymbol{\Gamma}$ is $\leq 2M$ (and is $M + 1$ in the EMA1 case as then the rows after the Mth are linear multiples of each other), which limits the number of non-zero eigenvalues to $2M$. The interpretation of all this is that a positively skewed system has a small number of large positive eigenvalues and a larger number of smaller negative ones. This generates a small number of large positive-convexity bets and a larger number of smaller negative-convexity bets, which is where the positive skewness comes from.

Hybrid linear systems

Suppose that a system has trend-following and counter-trending characteristics, as would happen if its weights were obtained from a linear combination of EMA2 systems, with opposite signs. It may be desirable to ensure that the long-term skewness remains positive, as this is associated with the longevity of the strategy. There are two situations in which this arises. In what is basically a trend-following system, it is desired:

- to make small bets on short-term reversion through a 'fast' component to the strategy without this upsetting the behaviour if a longer-term trend occurs; and
- to make a small bet against very long-term trends on the supposition that what goes up must eventually come down (or vice versa) through a 'slow' component, provided this bet is not too large.

In the first case, the weights on most recent returns will be negative; in the second, it is the weights on the distant past that will be negative. The idea is to make sure that they are not too negative, in a sense to be made precise.

We have a system of the form:

$$A(z)=\frac{\lambda_F\left(\alpha_F-\beta_F\right)}{\left(1-\alpha_F z^{-1}\right)\left(1-\beta_F z^{-1}\right)}+\frac{\lambda_S\left(\alpha_S-\beta_S\right)}{\left(1-\alpha_S z^{-1}\right)\left(1-\beta_S z^{-1}\right)}$$

where λ_F and λ_S are the multipliers on the fast and slow components. Positive asymptotic skewness is ensured by (4):

$$\sum_j \rho_j A\left(\alpha_j^{-1}\right)^2>0 \qquad (11)$$

where there are now four poles, $\alpha_1 = \alpha_F$, $\alpha_2 = \beta_F$, $\alpha_3 = \alpha_S$ and $\alpha_4 = \beta_S$. So we have:

$$\rho_1=\alpha_F^2\lambda_F,$$

$$A\left(\alpha_1^{-1}\right)=\frac{\lambda_F\left(\alpha_F-\beta_F\right)}{\left(1-\alpha_F^2\right)\left(1-\beta_F\alpha_F\right)}+\frac{\lambda_S\left(\alpha_S-\beta_S\right)}{\left(1-\alpha_S\alpha_F\right)\left(1-\beta_S\alpha_F\right)}$$

and similarly for the other three. The left-hand side of (11) is a homogeneous cubic in λ_F, λ_S, which will factorise as:

$$\mathcal{P}\left(\lambda_F,\lambda_S\right)=\left(\lambda_F-\zeta_1\lambda_S\right)\left(\lambda_F-\zeta_2\lambda_S\right)\left(\lambda_F-\zeta_3\lambda_S\right)$$

where the ζ's are functions of the four poles. It is possible to identify the coefficients of λ_F^3, $\lambda_F^2\lambda_S$, $\lambda_F\lambda_S^2$, λ_S^3 as functions of the poles, then evaluate them and factorise the cubic by the Cardano-Tartaglia formula. However, for practical purposes one might just as well write a numerical routine for (11) and find the roots ζ_i numerically. One root ζ_1 has to be real, and the other two are likely to be complex because we expect $\mathcal{P}$ to be strictly increasing in λ_F and in λ_S. Raising either weight should enhance the trending behaviour and hence the asymptotic skewness.

As a particular example, let $N_{\alpha_F} = 5$, $N_{\beta_F} = 10$, $N_{\alpha_S} = 20$, $N_{\beta_S} = 40$. Then $\zeta_1 \approx -1.476$, and the condition for positive asymptotic skewness is simply:

$$\lambda_F+1.476\lambda_S>0$$

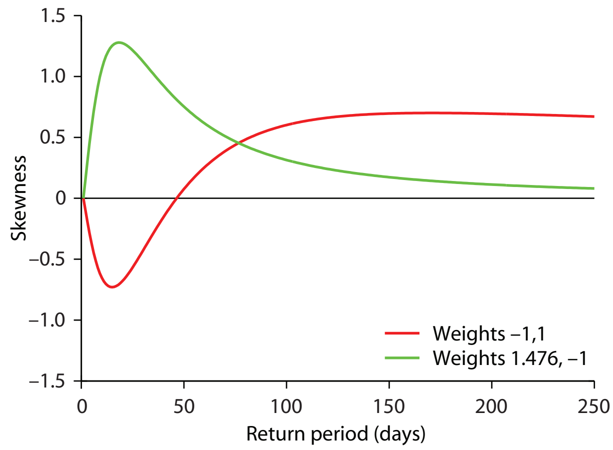
This being so, it is easily incorporated into an optimisation as a 'style constraint'. Figure 3 shows the results for two examples: (i) $\lambda_F = -1$, $\lambda_S = 1$, so the short-term behaviour is counter-trending and generates negative skewness; and (ii) the critical case $\lambda_F = 1.476$, $\lambda_S = -1$, where now just enough long-term counter-trending behaviour is added to make the asymptotic skewness zero at leading order. These exemplify the cases (i) and (ii) discussed above. The results were obtained using (5) again, which is not laborious despite there being four poles (so that the double summation has 16 terms). It is preferable to Monte Carlo simulation, which even with a few hundred thousand simulations generates noticeable uncertainty.

Conclusion

We have shown how to analyse the behaviour of a variety of trend-following models by particular reference to the skewness of the distribution of trading returns. To do this, we have needed only the first three moments of the market returns, keeping the modelling quite general. The most important of the formulas we have derived are (3) and (5), giving the second and third moments of the trading returns in an elegant application of residue calculus. Pure momentum systems generate positive skewness even though the market returns might be totally symmetrical, but the skewness depends on the return period and has a characteristic term structure that we have derived, illustrated and verified with real data. Hybrid strategies, with trending and counter-trending behaviour, may exhibit a more complex term structure of skewness, and we have shown how to analyse a general linear system.

A related question is the full distribution of trading returns, for which we need to make an assumption about the corresponding full distribution of the market returns: considerably stronger assumptions must be made. In some cases, notably normal market returns, this can be calculated with the aid of the moment-generating function, $\langle \exp(sY^{(M)})\rangle$. When φ is linear in the U_n, as we have considered here, the moment-generating function is $[\det(\mathbf{I} - 2s\mathbf{\Gamma})]^{-1/2}$ with $\mathbf{\Gamma}$ as in (9). From (10), we see that $Y_n^{(M)}$ is a weighted sum of independent χ_1^2-distributed random variables. An accurate approximation to the distribution and related risk measures may be easily obtained by saddlepoint methods (see Feuerverger & Wong, 2000, and Martin, 2011). ■

3 Skewness of trading returns, as a function of period, for hybrid model with both trending and counter-trending behaviour, in two cases



Source: Bloomberg

Richard Martin is chief quantitative investment officer at Longwood Credit Partners in London. David Zou is a quantitative analyst at AHL in London. A longer version is available upon request. This paper was written when both authors were at AHL. Email: martinr@longwoodcp.com

References

Björk T, 1998
Arbitrage theory in continuous time
Oxford University Press

Feuerverger A and A Wong, 2000
Computation of value-at-risk for nonlinear portfolios
Journal of Risk 3(1), pages 37–55

Haykin S, 1989
Modern filters
Macmillan

Herbst A, 1992
Analyzing and forecasting futures prices
Wiley

Martin R, 2011
Saddlepoint methods in portfolio theory
In The Oxford Handbook of Credit Derivatives, chapter 15, edited by A Lipton and A Rennie, Oxford University Press. Also at *arxiv.org/pdf/1201.0106*

Potters M and J-P Bouchaud, 2005
Trend followers lose more often than they gain
Wilmott Magazine, Nov/Dec, also at *arXiv:physics0508104v1.pdf*

Till H and J Eagleeye, 2011
A hedge fund investor's guide to understanding managed futures
Technical report, EDHEC-Risk Institute