Style-based value-at-risk for UK equities

What is the best risk measure for active fund managers? Stephen Rees argues that tradtional tracking error measures should be replaced by a form of value-at-risk. Here he constructs and tests a parametric VAR model in the context of UK equity 'style' investment

In my paper 'VAR for fund managers' (Rees, 2000), I argued that tracking error, long the preferred risk measure of portfolio managers, would be better replaced by a relative value-at-risk measure suitably tailored for the longer-term time horizons encountered in the investment field. The basis of my argument was that tracking error takes into account only the volatility of active (ie, benchmark-relative) returns, not any systematic trend to them. Any such trends could become significant over time-scales of months or years, and have a bearing on downside performance that tracking error would fail to capture. I suggested that systematic portfolio active returns might be modelled in terms of so-called 'investment styles', essentially the characteristics of particular types of security such as 'value' or 'growth' stocks. I gave some examples of the trending nature (ie, non-zero means) of the returns associated with these styles and suggested how this property might be used to estimate a 'drift' term, which could be added to tracking error to give a relative value-at-risk number.

Here I outline the development and backtesting of such a model for the UK equity market. Quantitative evidence is presented for the effectiveness of styles as a set of factors to model common residual risk. That the returns to the style factors trend over time, and that this trending is statistically significant is then shown. A prescription for building a relative VAR model using the styles is laid out and the model backtested by means of Monte-carlo simulations over a 14-year period. By this means, the model's efficacy, versus tracking error, in measuring downside active returns is demonstrated.

Why VAR is a more apt risk measure than tracking error

Most conventional fund managers have their performance measured with respect to a benchmark. Typically, benchmarks are published indexes representing the underlying markets. Examples are the S&P 500 index for US equities or the Salomon World Government Bond Index for global bonds. In quantifying their risk, fund managers look for an estimate of their portfolios' possible underperformance of these indexes (downside active return). For this, most of them tend to use 'tracking error' σ. This is the annualised standard deviation of active returns:

$$\sigma = \sqrt{\mathrm{Avg}\left(\left[\left(R_p(t) - R_M(t)\right) - \mathrm{Avg}\left(R_p(t) - R_M(t)\right)\right]^2\right)} \quad (1)$$

where $R_P(t)$ and $R_M(t)$ are the annualised portfolio and benchmark returns, respectively, at times t.

If the long-term expected mean of the active return is zero, σ is a perfectly good measure for likely downside. At least in a world of normally distributed returns, the portfolio would not deviate from its benchmark by more than σ in two out of any three given years, and would not underperform by more than σ in 8.4 out of every 10 years. However, the expected active return of an active portfolio is anything but zero. Fund managers are paid to outperform their benchmark, so 'expect' a significant positive active return. The only time this is not so is in the case of index tracking funds. It is interesting to note that 'tracking error' as a term and a measure, came from index fund management, and has crossed the border, unchallenged, into the active management field.

Any non-zero expected active return is referred to as 'drift'.

$$\mathrm{Drift} = \mathrm{Avg}(R_p(t)\text{-}R_M(t)) <> 0 \quad (2)$$

If this drift term were negative then in 8.4 out of every 10 years, the portfolio's underperformance would not exceed its magnitude plus the tracking error. It is this number which is the portfolio's relative VAR[1]:

$$\mathrm{VAR} = \sigma + \mathrm{Abs(drift)} \quad (3)$$

Of course, the actual underperformance and the two components of VAR, above, can always be measured after the fact, or *ex-post.* Modern portfolio theory provides us with the framework to estimate, or model VAR *ex-ante.* This is Ross's Arbitrage Pricing Theory (APT), which decomposes returns into a security-specific portion and a number of systematic components (Ross, 1976):

From APT, we deduce that, to first order, the active return R^i_A of a security is given by:

$$R^i_A = R^i - R_M = \Sigma^N_{j=1} \ \beta^i_j . F_j + \alpha^i \quad (4)$$

Where R_i is the total return of security i and R_M is the market return. F_j are the factor returns of the factors $j=1$, N. β^i_j is the sensitivity of stock i to factor j and α^i is the stock's alpha

Ross' theory in its most general form is:

$$R^i - r_f = \beta^i_M.(R_M - r_f) + \Sigma^N_{j=1} \ \beta^i_j . F_j + \alpha^i$$

r_f is the risk-free interest rate, with R^i -r_f referred to as the excess return. β_M is the *CAPM beta* (as in the Capital Asset Pricing Model) and $\beta_M.(R_M - r_f)$ is commonly referred to as the systematic return and is that portion of the stock's return explained by the market's return, R_M, alone. The components, $\beta^i_j . F_j$, when taken together comprise the common residual return. α^i is the specific residual return or more commonly alpha.

Clearly, then, armed with a complete set of factor returns and exposures (and assumptions on the distribution of α^i), equation (4) allows us to estimate both σ, in equation (1) and the drift term in equation (2), and hence the relative VAR in equation (3).

'Styles' as appropriate risk factors for the UK equity market

Equation (4) tells us nothing about what the risk factor set { j } actually are, and they have been represented in a number of different ways including

[1]. VAR in its most general form is defined with respect to a particular time period and level of confidence (again, see ref.1 for full details). The above is the special case for 1 year and 83.5% confidence - the natural extension of tracking error. It should be noted that the VAR described in this portfolio is a special type of relative VAR appropriate for long term VAR estimates of equity portfolios. It is also of the 'parametric' type. For a more general discussion on VAR, including stochastic and historical approaches, please refer to ref.1]

as macro-economic influences and membership of different industry groups (Grinold & Kahn, 1994). In fact, virtually any set of factors can be used as long as they explain a high proportion of the active return. An arbitrarily chosen set of factors will normally give rise to a high degree of correlation between the factor returns F_j so that cross terms need to be added to equation 4. Factors where the cross correlations are low would be a more 'natural' set. A further plus point would be parsimony: ie, the smaller the number of factors required to explain the most active return, the better. In short, an ideal set of risk factors would be complete, orthogonal and parsimonious.

A number of commentators have suggested that investment styles can sometimes fit this prescription (eg, ref. 4). Styles are factors representing different types of investment management. The most popular styles are: value, where the managers buy stocks which are cheap according to such criteria as book/price or price/earnings ratios; growth, where growth stocks, as defined by measures like return on equity and earnings growth, are bought; size, where managers buy securities with either large or small market capitalisation; and momentum, corresponding to trend-following behaviour.

Given suitable, quantifiable definitions of the styles, any portfolio's exposure to each of them can be measured. To gauge their usefulness as a set of risk factors for the UK equity markets, a style analysis was carried out on the constituents of the FTSE 350 index. This index is a broad one covering most of the large and mid-cap stocks likely to be held in a UK active equity portfolio. Book/price was used to represent value, return on equity for growth. Market capitalisation was used for size and rolling three-month *ex-ante* return was used for momentum. On the last day of each calendar month from December 1986 to December 2000, exposures of each stock, i, in the index to each style, labelled j, β^i_j ,were calculated as the Z-score of the style :

$$\beta^i_j(t) = \left(S^i_j(t) - \sum_i w^i(t) S^i_j(t) \big/ \sigma S^i_j(t)\right) \quad (5)$$

$W^i(t)$ is the weight of stock i in the index at time t so that $\Sigma_i\, w^i(t)S^i_j(t)$ is the cap-weighted cross-sectional average of style $S^i_j(t)$ of index stocks at time t. Thus if j represented book/price, the summation would be the cap weighted average book/price of the FTSE-350 index. $\sigma(S^i_j(t))$ is the cross-sectional standard deviation of the style calculated about this mean. Thus, a stock with a style number equal to that of the index (eg, book/price same as the index) would have a sensitivity to that style of zero. The cross-sectional distribution of sensitivities for each style will have a standard deviation of 1.

Factor returns, $F_j(t)$ were then estimated cross-sectionally for each style (j=1,4) from equation (5) for each month, t. This was done using OLS regression. $R^i(t)$ and $R_M(t)$ are the one-month *ex-post* returns of stock i and the FTSE-350 index respectively.[2]

For each stock, at each time, the residual active return, $\varepsilon^i(t)$ – ie, that portion of active return not explained by the four styles – was calculated as:

$$\varepsilon^i(t) = R^i_A - \sum_{j=1}^{4} \beta^i_j(t) F_j(t) \quad (6)$$

The 'explanatory power', R^2, of the styles was estimated as the percentage of total variance of the FTSE 350 universe they accounted for according to the following formula:

$$R^2 = 1 - VAR(\varepsilon^i) / VAR(R^i_A) \quad (7)$$

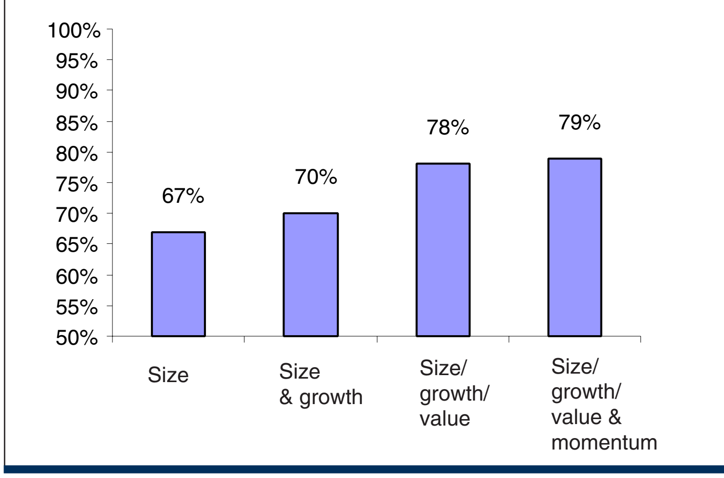
Where $VAR(R^i_A)$ is the sum of the elements in the covariance matrix of active returns, and $VAR(\varepsilon^i)$ is the same for the covariance matrix of residual active returns. The summations were performed in two different ways – equal-weighted and cap-weighted. The latter is probably a more appropriate estimate of explanatory power as it would correspond more closely to that for 'real portfolios', which tend to have a cap distribution more like the index. The covariance matrices were calculated using the entire 14 years of return history.

It turns out that the explanatory power of the styles is high. Size has the highest, accounting alone for 66% of the variance, cap-weighted (and almost 80% equal-weighted). All four together explain 79%, cap-weighted and a very high 91% when equal weighted. The incremental explanatory power (for the cap-weighted case is illustrated in figure 1.

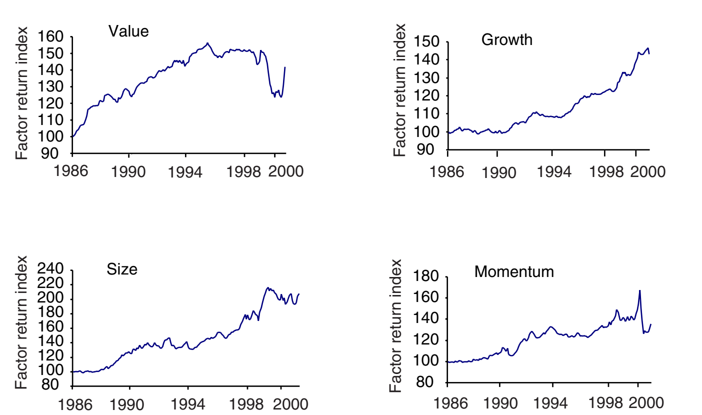
Thus, the styles appear to satisfy the first criteria of a good set of risk factors in that they explain a high proportion of active returns. They also come close to satisfying the second condition – that of orthogonality. The correlation matrix of style factor returns (table A) shows this, with generally low correlations. It is really only the -0.35 number between growth and value which might stretch the argument a bit.

With only 4 styles, the third condition of parsimony is also clearly satisfied.

1. % variance explained by styles (FTSE 350)



2. Trending behaviour of UK style-factor return



A. UK style-factor return correlations

	Value	Growth	Size
Growth	-0.35		
Size	0.23	-0.03	
Momentum	0.07	0.07	0.06

So the styles comprise a good set of risk factors. Using other factors, such as industry membership or macro sensitivities instead of – or in addition to – styles might allow us to explain even more variance.[3]

However, what makes styles so appropriate for VAR estimation is their 'trending' behaviour – which provides a possible means to model the drift term.

[2] The return variable used, here and throughout, was natural logarithm of wealth ratio as this is more normally distributed than ordinary returns and so more appropriate for OLS regression. To first order, this equals simple return. A process of winsorisation was also employed where stocks with bij greater than 3 standard deviations, were removed from the analysis and Si wi(t).Sij(t) recalculated to remove the effect of these 'outliers'

[3] The assumption here, of course, is that the explanatory power of an arbitrary portfolio's active variance would be similar to that of the index. No exhaustive testing has been done to prove this is the case and it is perfectly possible that portfolios with, for example, industry biases might not be catered for so well by the styles

B. Long-term average style-factor returns

	Value	Growth	Size	Momentum
Mean factor return (per annum)	3.20%	2.60%	5.90%	2.40%
t-stat	2.7	3.1	4.7	1.9

C. Downside active return over/under-prediction (% per annum)

		Ex-post drift	Ex-ante drift
Ex-post	d(σ)	-1.2%	-1.2%
tracking error	d(VAR)	-0.2%	0.3%
Ex-ante	d(σ)	-1.6%	-1.6%
tracking error	d(VAR)	-0.8%	0.07%

D. Downside active return over/under-prediction (% per annum): 1998–2000

		Ex-post drift	Ex-ante drift
Ex-post	d(σ)	-0.6%	-0.6%
tracking error	d(VAR)	0.3%	2.3%
Ex-ante	d(σ)	-2.2%	-2.2%
tracking error	d(VAR)	–1.5%	0.7%

Figure 2 shows the cumulative factor return indexes for each style, based to 100 at December 1986. Despite some short-term volatility, the long-term trending of the factor returns can clearly be seen (the implication for active managers is long-term outperformance from purchasing cheap, large stocks with good growth and momentum). The significance of the drift, ie, mean factor return significantly different from zero, is shown in table B.

From these non-zero style means $<F_j(t)>$, the drift term for stock i can be estimated as:

$$Drift^i = \Sigma^4_{j=1} \beta^i_j(t). <F_j(t)> \quad (8)$$

Thus, from equation 3, the relative VAR of a portfolio with stock weights w^i can be calculated as:

$$\begin{aligned} VAR &= \sigma + Abs\left(\sum_i W^i Drift^i\right) \\ &= \sigma + Abs\left(\sum_i W^i \sum_{j=1}^{4} \beta^i_j(t) < F_j(t) >\right) \end{aligned} \quad (9)$$

Testing the efficacy of UK style-based VAR using Monte-Carlo simulations

The acid test of the style-based VAR model is whether or not it does a better job of modelling downside active returns. To this end, a series of Monte-Carlo simulations were carried out in which random UK equity portfolios' returns were compared with the two metrics. In these simulations, one hundred 50-stock portfolios were drawn from the FTSE 350 at the end of each calendar month from December 1986 to January 2000-a total of 14,600 portfolios overall. 50 stock portfolios were chosen because this is fairly typical of the number of names that would be held in an active UK equity fund.

The FTSE 350 index is cap-weighted, so is skewed towards the larger stocks. At any time, only about 20% of the stocks (70 names) are 'large' companies, having a capitalisation greater than that of the index mean. The remaining 80% (280 names) are 'smaller companies'. Completely random portfolios would thus invariably have a small-cap tilt. To prevent this systematic bias, and further make the random portfolios look more like 'real' ones, their split between 'large' and 'small' cap stocks was made to accord with the index. 10 stocks (20%) were drawn from the 70 large cap names and 40 (80%) from the 280 small caps for each portfolio. The stocks were then market-cap weighted.

Each portfolio's active return was measured over the following 12 months and compared with both tracking error and VAR.

A couple of different measures were used for tracking error. We could, of course, have used the style factors to model the tracking error as well as the drift. This is not done here, both for simplicity sake when conducting the simulations and because there are a variety of perfectly good tracking error forecasting models available, which could be used in practice and over which a style-based model is unlikely to offer any improvement. Here, we use an *ex-post* measure calculated from perfect foresight of standard deviation of monthly active returns over the actual 12-month simulation periods and an *ex-ante* measure calculated from the previous 12 months' active returns. The first measure is the actual measured tracking error of the portfolio, and any forecasting model could clearly not give a better tracking error estimate than this. The second measure is meant to be a proxy for the tracking error forecast by some model.

Ex-post and *ex-ante* versions of the drift term (equation 8) were also used. In the *ex-post* version the long-term mean style-factor returns shown in table B were used for $<F_j(t)>$. In the *ex-ante* case $<F_j(t)>$ were estimated as rolling average style-factor returns calculated over the previous three months. Thus, two versions of tracking error and four versions of VAR were considered. For each version, the figures were examined to see which gave the best approximation for downside active return, R_A for each portfolio in the simulation. The overall efficacy of the models was gauged by measuring the average difference of the measures, d(σ) and d(VAR) in equation 10 and 11, across all the underperforming portfolios in the simulation, to see which was closer to zero:

$$d(\sigma) = Avg\left(\sigma - Abs(R_A)\right) \quad (10)$$

$$d(VAR) = Avg\left(VAR - Abs(R_A)\right) \quad (11)$$

The results are shown in table C.

It appears from this table that the VAR model does indeed do a better job than tracking error. On average, the *ex-post* tracking error under-represents downside by 1.2% per annum. VAR gets closer, underestimating downside by just 0.2% when long-term style drifts are used, and overestimating by just 0.3% with short-term *ex-ante* style drifts. Similarly, the *ex-ante* tracking error measure underestimates downside by 1.6%, improving to 0.8% when VAR with long-term mean drifts is used and a very small overestimate, just 0.07%, when short-term *ex-ante* drifts are used. Note that this last case uses data, all of which would have been available prior to the start of each simulation. These VAR estimates, then, could actually have been used as downside risk predictions. Given the large number of simulations used to reach these figures, it would appear that these differences are statistically significant. For example, a t-test between the means $d(\sigma) = -1.2\%$ and $d(VAR) = -0.2\%$ gives a figure of 17.6 standard errors.

It is in recent years that, anecdotally at least, people have increasingly been reporting the failure of tracking error to adequately measure downside performance. Markets have certainly been more volatile recently and, as can be seen from figure 2, this has been reflected in UK style-factor return volatility, especially apparent in the years 1998–2000. Mostly, this has been due to the boom and bust in technology and media (TMT) companies. How would VAR have fared versus tracking error during this period? Table D gives the answer.

Interestingly, *ex-post* tracking error does a better job during this period than it did over the long history, underestimating downside by just 0.6% (cf 1.2%). This is not really surprising, as it is simply a reflection of the increased volatility actually seen in the market. Indeed, adding the *ex-ante* drift term produces a substantial overestimate of downside (+2.3%). Again, this is not surprising, as very high mean-style returns were seen over this period. Note, though, that *ex-ante* tracking error underpredicts downside quite substantially (-2.2%) while VAR with an *ex-ante* drift added to the same tracking error does a good deal better, getting to within 0.7% of the actual downside.

Conclusion

The investment styles: value, growth, size and momentum together comprise a good set of risk factors for the UK equity market. The trending behaviour of the factor returns to these styles (ie, significant non-zero means) enables us to estimate a 'drift' term, as well as a tracking error, and so derive a full-blown relative VAR model. The Monte-Carlo simulations presented here indicate that this VAR model does a better job of predicting downside active returns than does tracking error alone.

The differences between the accuracy of the VAR models and tracking error measures may not seem especially large from tables B and C (given that typical tracking errors of the portfolios were around 6%), so the results here admittedly only weakly support the assertion that 'VAR is better than tracking error'. However, because the portfolios in these simulations are randomly selected, we would mostly not expect their style exposures to be very large. Many fund managers would be likely to introduce stronger biases by virtue of their own styles of management (value, growth etc). Stronger biases would result in larger drift terms and thus bigger differences between the VAR and tracking error numbers. The question of whether the improvement of VAR over tracking error, in accounting for underperformance, would be more substantial under these circumstances provides scope for further interesting research. □

Stephen Rees is director, head of quantitative research, Baring Asset Management

References

Rees S, 2000
VAR for fund managers
Risk, June

Ross S, 1976
The arbitrage theory of capital asset pricing
Journal of Economic Theory 13

Grinold R & Kahn R, 1994
Multiple factor models in portfolio risk
A Practitioner Guide to Factor Models, AIMR

Schwob R, 2000
Style & style analysis from a practitioner's perspective
Journal of Asset Management 1