Principles of the skew

The ISMA Centre's Carol Alexander shows how principal component analysis should be used on the fixed strike deviations from at-the-money volatility

Many financial markets are characterised by a high degree of collinearity. It occurs when there are only a few important sources of information in the data, which are common to many variables. In this case, principal component analysis (PCA) is a standard method for extracting the most important uncorrelated sources of variation in a multivariate system.

PCA is not just about term structures of interest rates or futures, although most readers will be familiar with the method in this context. The standard interpretation of the first component as the trend, the second component as the tilt and the third component as the curvature holds for any highly correlated ordered system, not just a term structure. So when implied volatilities are ordered by strike or moneyness, a PCA application will reveal the standard trend-tilt-curvature interpretation of the first three principal components.

Several principal component models of volatility smiles and skews have been based on daily changes in implied volatilities, by strike and/or by moneyness. Derman and Kamal (1997) analysed S&P500 and Nikkei 225 index options, where the daily change in the volatility surface is specified by delta and maturity. Skiadopoulos, Hodges and Clewlow (1998) applied PCA to first differences of implied volatilities for fixed maturity buckets, across both strike and moneyness metrics. And Fengler et al (2000) employed a common PCA that allows options on equities in the Dax of different maturities to be analysed simultaneously.

There is an important difference between the research just cited and the approach taken in this paper. Instead of applying PCA to daily changes in implied volatilities, a PCA is applied to daily changes in the deviations of fixed strike volatilities from at-the-money (ATM) volatility. The advantages of this approach are both empirical and theoretical. On the empirical front, time series data on fixed strike or fixed delta volatilities often display much negative autocorrelation, possibly because markets over-react. But the daily variations in fixed strike deviations from ATM volatility are much less noisy than the daily changes in fixed strike (or fixed delta) volatilities. Consequently the application of PCA to fixed strike deviations from ATM volatility yields more robust and intuitive results.

The theoretical advantage of this approach is that the models of the skew in equity markets that were introduced by Derman (1999) can be expressed in a form where fixed strike volatility deviations from ATM volatility always have the same relationship with the underlying index. The particular market regime is determined only by a different behaviour in ATM volatility. Thus the stability of PCA on fixed strike deviations from ATM volatility is implied by Derman's models.

Derman (1999) asked: how should implied volatilities be changed as an equity index moves?. He developed three models that are described in box 1. In each model there will be a parallel shift in all volatilities as the index moves, where the size of this shift is determined by the current market regime. The model presented here extends Derman's models to allow non-parallel shifts in the skew as the index moves. It uses PCA to actually quantify the sensitivities of implied volatilities to changes in the underlying price.

The model has applications to all types of im-

Box 1. 'Sticky' models

Derman (1999) formulates three different regimes of price-volatility behaviour in equity index markets. A different linear parameterisation of the volatility skew applies in each regime, and each parameterisation implies a different type of 'stickiness' for the local volatility in a binomial tree. For this reason, the skew parameterisations are known as Derman's 'sticky' models.

Denote by $\sigma_K(t)$ the implied volatility of an option, with maturity t and strike K, $\sigma_{ATM}(t)$ the volatility of the t-maturity ATM option, S the current value of the index and σ_0 and S_0 the initial implied volatility and price used to calibrate the tree:

a) In a range bounded market, skews should be parameterised as

$$\sigma_K(t) = \sigma_0 - b(t)(K - S_0)$$

So fixed strike volatility $\sigma_K(t)$ is independent of the index level and σ_{ATM} will decrease as index increases:

$$\sigma_{ATM}(t) = \sigma_0 - b(t)(S - S_0)$$

b) In a stable trending market, skews should be parameterised as:

$$\sigma_K(t) = \sigma_0 - b(t)(K - S)$$

So fixed strike volatility $\sigma_K(t)$ will increase with the index level. But $\sigma_{ATM}(t)$ will be independent of the index, since

$$\sigma_{ATM}(t) = \sigma_0$$

c) In jumpy markets, skews should be parameterised as:

$$\sigma_K(t) = \sigma_0 - b(t)(K + S) + 2b(t)S_0$$

So fixed strike volatility $\sigma_K(t)$ will decrease when the index goes up, and increase when the index falls, since

$$\sigma_{ATM}(t) = \sigma_0 - 2b(t)(S - S_0)$$

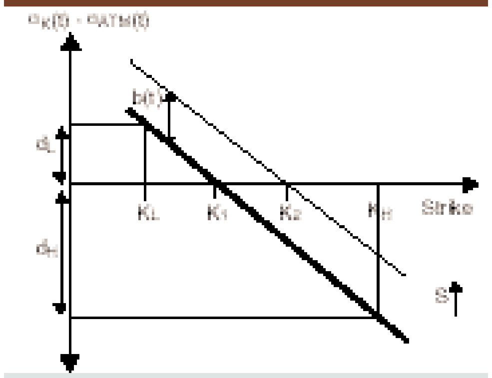
the ATM volatility will also decrease as the index goes up and increase as the index falls, twice as fast as the fixed strike volatilities.

The range-bounded model (a) is called the 'sticky strike' model because local volatilities will be constant with respect to strikes. That is, each option has its own binomial tree, with a constant volatility that is determined by the strike of the option. For a given change in the index, the root of the tree is moved to the new level of the index. The same tree is still used to price the option.

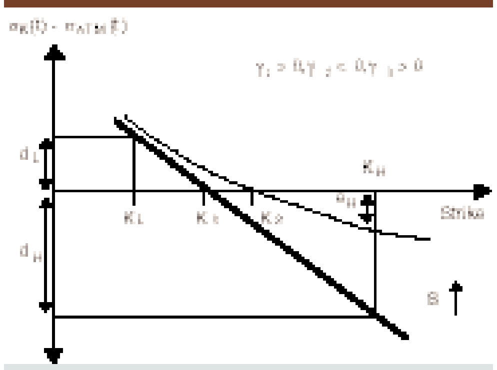
The trending markets model (b) is called the 'sticky delta' model because local volatilities are constant with respect to the moneyness (or equivalently the delta) of the option. That is, it is the moneyness of the option that determines the (still constant) local volatility in the tree. As the index changes, the delta of the option also changes and we consequently move to a different tree, the one corresponding to the current option delta.

In the 'sticky tree' model (c), the local volatilities are no longer constant. There is, however one unique tree that can be used to price all options determined by the current skew. This is the implied tree described in Derman and Kani (1994). ■

1a. Parallel shift in skew deviations as price moves up



1b. Non-Parallel shift in skew deviations as price moves up



plied volatility surfaces, including currency option smiles and swaption skews. But the present paper focuses on its application to the skew in the FTSE 100 between January 4 1998 and March 31 1999. It is found that the sensitivity of a fixed strike volatility to movements in the index can change according to market conditions. In particular, for short term volatility the range of the skew (the difference between low strike volatility and high strike volatility) will normally fluctuate over time.

Fixed strike deviations from ATM volatility

The relationship between fixed strike deviations from ATM volatility and the underlying price is the same in all of Derman's 'sticky' models. In fact, for any maturity t there will be a linear relationship between the deviation of a fixed strike volatility from ATM volatility and the underlying price given by:

$$\sigma_K(t) - \sigma_{ATM}(t) = -b(t)(K - S) \quad (1)$$

For any given maturity, the deviations of all fixed strike volatilities from at-the-money volatility will change by the same amount b(t) as the index level changes, as shown in figure 1a. Four strikes are marked on this figure: a low strike K_L, an initial ATM strike K_1, a new ATM strike after the index level moves up K_2 and a high strike K_H. In each of the three market regimes, the range of the skew between K_L and K_H, that is $\sigma_L - \sigma_H$, will be the same after the index move. Thus as the underlying price moves, the fixed strike volatilities will shift in parallel, and the range of the skew will remain constant. The direction of the movement in fixed strike volatilities depends on the relationship between the original ATM volatility σ_1 and the new ATM volatility σ_2:

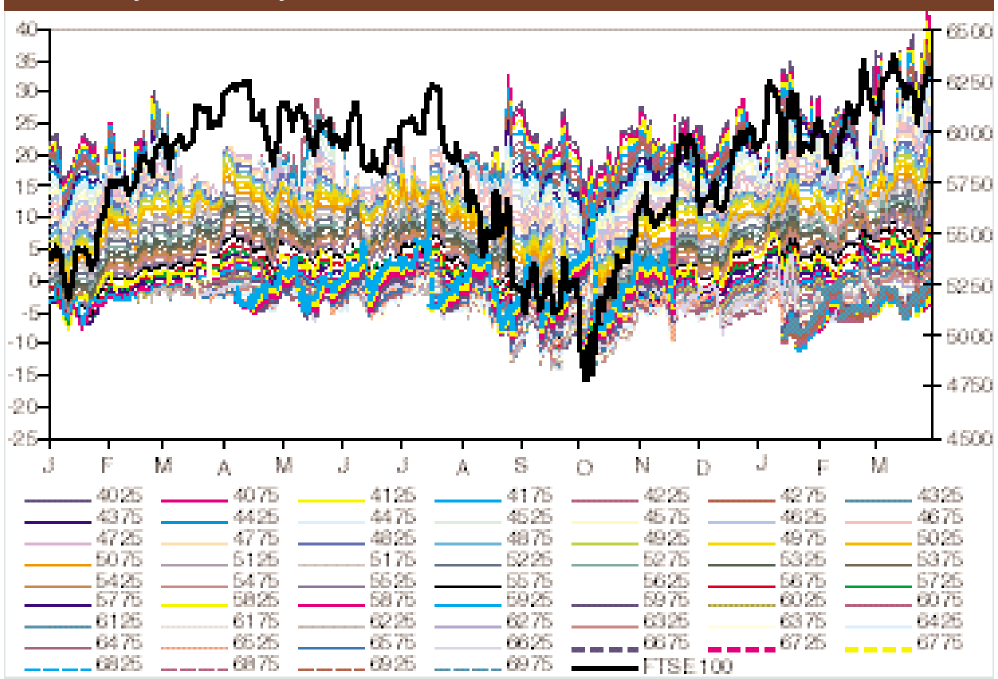
☐ In a range-bounded market, $\sigma_2 = \sigma_1 - b(t)$, but fixed strike volatilities have all increased by the same amount b(t), so a static scenario for the skew by strike should be applied;

☐ When the market is stable and trending, $\sigma_2 = \sigma_1$ and there is an upward shift of b(t) in all fixed strike volatilities;

☐ In a jumpy market, $\sigma_2 = \sigma_1 - 2b(t)$, so a parallel shift downward of b(t) in the skew by strike should be applied.

These observations suggest that a method for testing whether Derman s sticky models are supported by empirical evidence is to perform a PCA of the daily change in fixed strike volatility deviations from ATM volatility, denoted $\Delta(\sigma_K - \sigma_{ATM})$. Equation (1) implies that only the first principal component should be significant, but if it is found that the second or higher principal components are significant factors for determining movements in $\Delta(\sigma_K - \sigma_{ATM})$, then the parallel shifts in the skew that are implied by the sticky models will not apply.

2. Deviations of one-month fixed strike volatility from ATM volatility, January 1998-March 1999



PCA of skew deviations

The model has been developed on one-month, two-month and three-month implied volatilities for European options of all strikes on the FTSE 100 index for the period stipulated.[1] These implied volatilities form a correlated, ordered system that is similar to a term structure. It is therefore natural to consider using PCA to identify the key uncorrelated sources of information, and there will only be a few. However, a PCA of daily changes in the fixed strike volatilities themselves may not give very good results as the data will be rather noisy. But figure 2 shows that the fixed strike deviations display less negative autocorrelation and are even more highly correlated and ordered than the fixed strike volatilities. A strong positive correlation with the index is evident during the whole period.

The PCA of fixed strike deviations is based on the model:

$$\Delta\big(\sigma_K(t) - \sigma_{ATM}(t)\big) = \omega_{K,1}(t)\, P_1(t) + \omega_{k,2}(t)\, P_2(t) + \omega_{k,3}(t)\, P_3(t) \quad (2)$$

where t is the volatility maturity and K is the strike of the volatility, both fixed. Daily data on $\Delta(\sigma_K(t) - \sigma_{ATM}(t))$ is used to estimate the principal components $P_1(t)$, $P_2(t)$ and $P_3(t)$, and the factor weights $\omega_{K,1}(t)$, $\omega_{K,2}(t)$ and $\omega_{K,3}(t)$.

A PCA of $\Delta(\sigma_K(t) - \sigma_{ATM}(t))$ for a fixed maturity t gives excellent results (box 2). Alexander (2000) shows that for fixed maturity volatility skews in the FTSE 100 index option market during most of 1998, 80-90% of the total variation in skew deviations can be explained by just three key risk factors: parallel shifts, tilts and curvature changes. The parallel shift component accounted for around 65-80% of the variation, the tilt component explained a further 5 to 15% of the variation and the curvature component another 5% or so. The precise figures depend on the maturity of the volatility (one-month, two-month or three-month) and the exact period in time that the principal components were measured.

The immediate conclusion must be that the limitation of movements in volatility surfaces to paral-

[1] The fixed maturity implied volatility data used in this section have been obtained by linear interpolation between the two adjacent maturity option implied volatilities. However this presents a problem for the one-month volatility series. During the last few working days before expiry data on the near maturity option volatilities are totally unreliable. So the one-month series rolls over to the next maturity, until the expiry date of the near-term option, and thereafter continues to be interpolated linearly between the two option volatilities of less than and greater than one month

Box 2. PCA for a fixed maturity

A PCA for three-month implied volatility skew deviations gives the output in the tables shown. It is clear from table A that the first principal component only explains 74% of the movement in the volatility surface, and that the second principal component is rather important as it explains an additional 12% of the variation over the period. It is interesting that the factor weights shown in table B indicate the standard interpretation of the first three principal components in a term structure, as parallel shift, tilt and convexity components. Note that sparse trading in very out-of-the money options implies that the extreme low strike volatilities show less correlation with the rest of the system, and this is reflected by lower factor weights on the first component. ■

A. Eigenvalues of correlation matrix

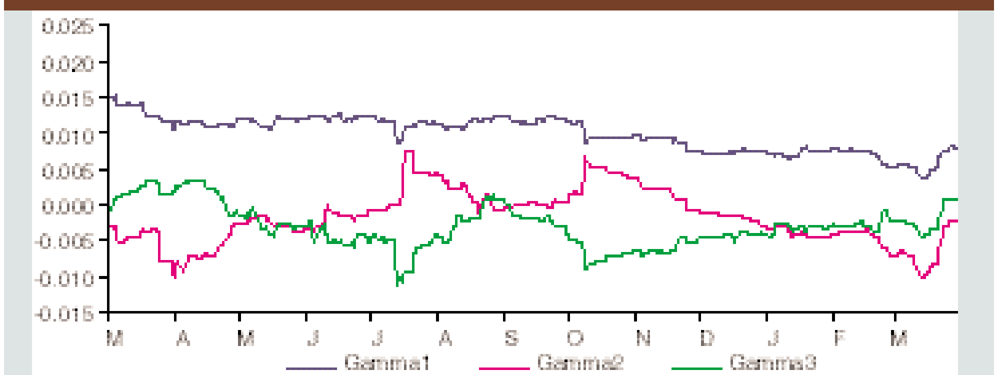
Component	Eigenvalue	Cumulative R^2
P1	13.3574	0.742078
P2	2.257596	0.8675
P3	0.691317	0.905906

B. Eigenvectors of correlation matrix

Factor Weights

	P1	P2	P3
4225	0.53906	0.74624	0.26712
4325	0.6436	0.7037	0.1862
4425	0.67858	0.58105	0.035155
4525	0.8194	0.48822	-0.03331
4625	0.84751	0.34675	-0.19671
4725	0.86724	0.1287	-0.41161
4825	0.86634	0.017412	-0.43254
4925	0.80957	-0.01649	-0.28777
5025	0.9408	-0.18548	0.068028
5125	0.92639	-0.22766	0.13049
5225	0.92764	-0.21065	0.12154
5325	0.93927	-0.22396	0.14343
5425	0.93046	-0.25167	0.16246
5525	0.90232	-0.20613	0.017523
5625	0.94478	-0.2214	0.073863
5725	0.94202	-0.22928	0.073997
5825	0.93583	-0.22818	0.074602
5925	0.90699	-0.22788	0.068758

3. Gamma estimates for one-month volatilities



lel shifts alone is an over simplification of what is actually happening in the data. The next section develops a model that encompasses changes in the tilt and curvature of the volatility skew, as well as the parallel shift. So the range of the skew can widen or narrow as the underlying price moves up or down, and change its curvature.

Fixed strike volatilities in different market regimes

It follows from equation (2) that the movement in fixed strike volatilities as the underlying moves will be determined by the movement in the principal components. Each component is assumed to have a linear relationship with daily changes ΔS in the underlying:

$$P_i(t) \approx \gamma_i(t)\Delta S \qquad (3)$$

Thus the movement in fixed strike volatility deviations in response to movements in the underlying will be determined by the factor weights in the principal component representation equation (2) and the gamma coefficients in equation (3).

Note that γ_1 represents the trend component and this is always assumed to be positive, an assumption that is justified by the empirical analysis below. The coefficient γ_2 determines the tilt of the fixed strike deviations and γ_3 determines the convexity of the skew. There are four stylised movements in the skew deviations, according to the signs γ_2 and γ_3.

Figure 1b depicts the movement in skew deviations as the index price moves up when $\gamma_1 > 0$, $\gamma_2 < 0$ and $\gamma_3 > 0$. The deviation at the high strike K_H is denoted d_H before the move and e_H after the move, and similarly d_L and e_L denote the before and after deviations at the low strike volatility K_L. The relation between d_H and e_H and the relation between d_L and e_L will depend on the values of γ_1, γ_2 and γ_3.[2] Normally there will be a change in the range of the skew as the index moves. When γ_2 is negative the range will narrow as the index moves up and most of the movement will be coming from low strike volatilities. But when γ_2 is positive, the range will widen as the index moves up and there will be more movement in high strike volatilities.

Exponentially weighted moving average estimates (with $\lambda = 0.94$) of γ_1, γ_2 and γ_3 for each of the one-month, two-month and three-month maturities have been calculated. The gamma estimates for the one-month volatilities are shown in figure 3. For reasons of space the two-month and three-month gamma estimates are not shown, but are given in Alexander (2000). In all cases, the estimate of γ_1 is strongly positive throughout. Since γ_1 captures the parallel shift component of movements in the skew, we can deduce that most of the movement in the skew at all maturities can be attributed to a parallel shift up when the index falls.

For the two-month and three-month maturities, the index seems to have little effect on the second and third principal components – the estimates of γ_2 and γ_3 are close to zero for almost all the sample period. There are a couple of negative γ_2 periods during the springs of 1998 and 1999, when the range of the skew will have narrowed as the index moved up and widened as it moved down. But this effect is not as pronounced as in the one-month skew. The rest of the time, and particularly during the crash period, it would be reasonable to apply parallel shift scenarios for fixed strike volatilities of two-month and three-month skews.[3]

A slightly different picture emerges for the movement of the one-month skew (figure 3). The estimate of γ_2 is often negative, particularly during the spring of 1998 and the spring of 1999. At these times the range of the skew was clearly decreasing when the index rose and increasing when the index fell, an effect which is very evident in figure 2. But there are two notable periods, just before the beginning of the crash and during the market recovery, when the estimate of γ_2 was strongly positive and γ_3 was strongly negative. On July 14 1998, several days before the FTSE 100 price started to plummet, there was a dramatic increase in γ_2 and decrease in γ_3 so that $\gamma_2 > 0$ and $\gamma_3 < 0$. During this period, the range of the one-month skew narrowed as the index fell. Then between October 8-12 1998, the FTSE 100 jumped up 8% in 2 days' trading, from 4803 to 5190. At the same time, γ_2 jumped up and γ_3 jumped down, so that again $\gamma_2 > 0$ and $\gamma_3 < 0$, and the range of the one-month skew widened as the index moved up. The narrowing of the range of the skew as the index fell, and the consequent widening again as the market recovered, was driven by unexpectedly big movements in high strike volatilities.

Scenario quantification for implied volatility surfaces

We assume for a fixed maturity t of 1, 2 or 3 months that:

$$\Delta\sigma_{ATM}(t) \approx \beta(t)\Delta S \qquad (4)$$

To capture the dependence of β on the current market regime, we estimate β with an expo-

[2] *When γ_2 is negative it is clear that e_H will be less than d_H and that e_L is normally a little greater than d_L, unless γ_2 is very large and negative. If γ_3 were extremely large and negative then e_H would be less than d_H but this never occurs empirically. On the other hand if γ_2 is positive then $e_L > d_L$ but now the sign of $e_H - d_H$ will be ambiguous. But normally e_H will be a little less than d_H unless γ_2 is very large indeed*

[3] *If the volatility by strike has a parallel shift then the volatility by moneyness or delta is static. Equivalently if volatility by delta has a parallel shift, volatility by strike is static*

[4] *Of course, if one uses equally weighted averages to estimate beta then it may appear to jump, but this would only be because of the 'ghost features' that are an artefact of pushing a constant parameter model into a time-varying framework*

nentially weighted moving average (again with $\lambda = 0.94$). These estimates are shown in figure 4. As expected, the sensitivity of ATM volatility to changes in the FTSE is greater in one-month options than in two-month options, which in turn have greater sensitivity than three-month options. There is a striking pattern in figure 4 and it is very clear indeed that the sensitivity of ATM volatility moves with the level of the index. It does not jump unless the index jumps.[4]

Combining equations (2), (3) and (4) yields:

$$\Delta\sigma_K(t) \approx \beta_K(t)\Delta S \qquad (5)$$

where the sensitivity of the fixed strike volatility $\sigma_K(t)$ to the index is given by:

$$\beta_K(t) = \beta(t) + \Sigma\, \omega_{K,i}(t)\gamma_i(t) \qquad (6)$$

This completes the model for changes in all fixed-strike volatilities as the index moves.

Figure 5 shows the estimates of β_K for strikes K between 4675 and 5875 and volatilities of one-month maturity that are obtained from equation (6). These lowest and highest strikes are picked out in red and green. The index sensitivity of all fixed strike volatilities are negative, so they move up as the index falls, but by different amounts.

During the crash period the sensitivities of all volatilities are greater and the change in the 5875 strike volatility sensitivity is very pronounced. Before the crash, it ranged between -0.005 and -0.01, indicating an increase of between 0.5 and 1 basis points for every FTSE point decrease. At the beginning of the crash the 5875 sensitivity increased to about 1.5bp, and since the FTSE fell by 1500 points during the crash, that corresponds to a 22.5% increase in 5875 volatility. Then at the height of the crash, between October 1-9, the 5875 sensitivity became increasing large and negative as the FTSE index reached a low of 4786 on October 5. On October 9, the 5875 sensitivity was an impressive -0.028, indicating a further 2.8bp increase in 5875 volatility would have occurred for every point off the FTSE at that time.

What is interesting about the 4675 volatility sensitivity is that it is often far greater (in absolute terms) than the high strike volatility sensitivities. So most of the movement will be coming from the low strikes as the range of the skew narrows when the index rises and widens when the index falls. The 4675 volatility approximately gains about 1 or 2bp for every point fall in the FTSE index during the period, although the sensitivity varies considerably. At the end of the data period it is extraordinarily large because the range narrowing of the skew was very considerable.

Conclusion

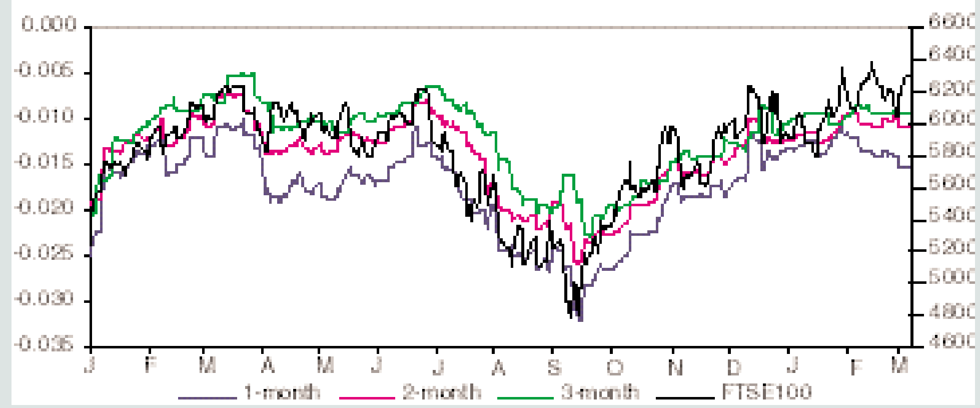
This paper has presented a new principal component model of fixed strike volatility deviations from ATM volatility. It is based on the observation that fixed strike volatility deviations from ATM volatility are independent of the market regime, and that the regime is determined only by different behaviour in ATM volatility. The model presented here actually quantifies the change that should be made to any given fixed strike volatility per unit change in the underlying, at any point in time. Empirical application of the model to the FTSE 100 index options has shown that two-month and three-month skews should normally be shifted parallel as the index moves, as predicted by Derman's models. In the 'range-bounded' markets in the spring of 1998 and 1999, there was also some narrowing in the range of the skew as the index moved up (and widening as the index moved down). But the range narrowing of the skew in a range bounded market scenario is much more pronounced in the one-month skew.

For very short term volatility in equity markets, the empirical analysis has revealed two distinct regimes. In stable markets, the range of the one-month skew narrows as the index moves up and widens as the index moves down. Most of the movement is in low strike volatilities and high strike volatilities remain relatively static as the underlying moves. But there is a second regime in short-term volatilities that applies to the jumpy markets during a market crash and recovery period. In this regime, the high strike volatilities move much more than usual, and in the recovery period after the 1998 crash the one-month skew range actusally widened as the index moved up.

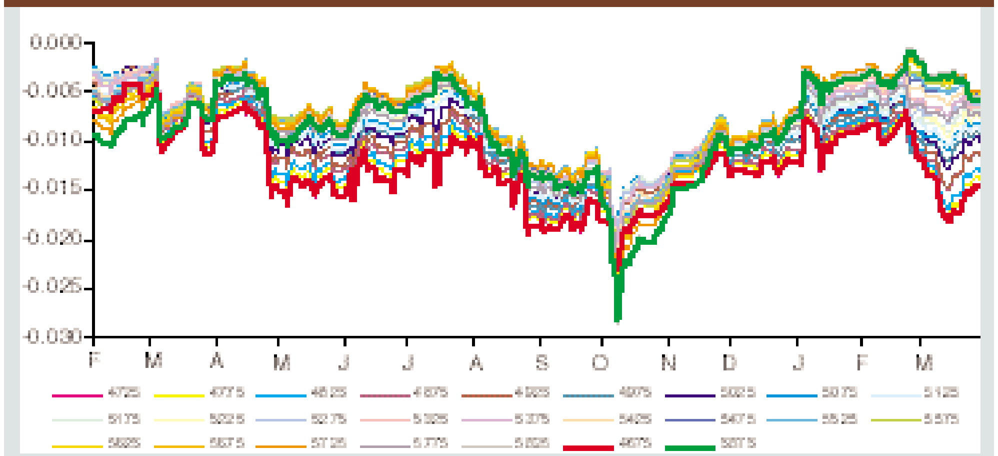
The model presented has very general applications because it admits non-linear movements in the volatility smile as the underlying moves. PCA is shown to be a powerful analytical tool for the construction of an implied volatility smile surface in different market regimes. ■

Professor Carol Alexander is chair of risk management at the ISMA Centre, University of Reading, UK. My sincere thanks to Emanuel Derman of Goldman Sachs, New York and to Jacques Pezier of CAL-FP Bank, London for many clarifying and useful discussions

4. ATM volatility sensitivity and the FTSE 100, January 1998-March 1999



5. Change in one-month fixed strike volatility per unit increase in index, February 1998-March 1999



REFERENCES

Alexander C, 2000
Principal component analysis of implied volatility smiles and skews
ISMA Centre, discussion papers in finance, 2000-10

Derman E, 1999
Volatility regimes
Risk 12:4 pages 55-59

Derman E and M Kamal, 1997
The patterns of change in implied index volatilities
Quantitative Strategies Research Notes, Goldman Sachs

Derman E and I Kani, 1994
Riding on a Smile
Risk 7:2 pages 32-39

Fengler M, W Hardle and C Villa, 2000
The dynamics of implied volatilities: a common principal component approach
Available from fengler@wiwi.hu-berlin.de, September

Skiadopoulos G, S Hodges and L Clewlow, 1998
The dynamics of implied volatility surfaces
Financial Options Research Centre Preprint 1998/86
Warwick Business School, University of Warwick, UK