Crises and volatility

Allan Malz surveys the behaviour of forward-looking asset prices during market crises, focusing on historical and implied volatilities. The latter are shown to contain useful information for predicting market stress in the immediate future

Market risk measurement has come a long way since its value-at-risk origins in the 1980s. The most important developments come under the rubric 'stress testing' and address limitations of the classical joint-normal return model underpinning VAR. Stress testing measures how a trading book or portfolio would fare under a set of stress scenarios. Scenarios can be drawn from historical episodes or can be based on traders' judgement regarding likely or particularly adverse events. Scenarios can be parameterised to reflect current or anticipated correlations of the risk factors of primary focus with the remaining risk factors in the portfolio.[1]

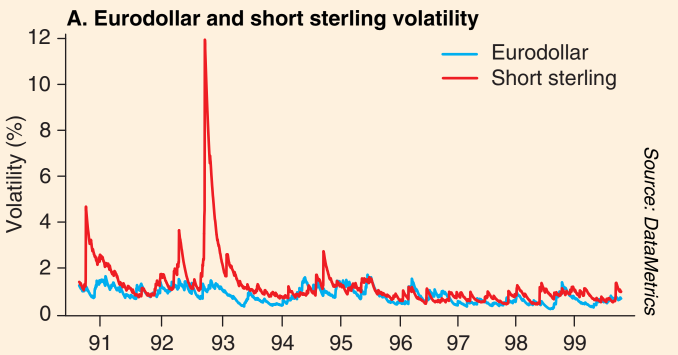
Stress testing is as much art as science. Since not all adverse scenarios can be tested all the time, choosing or designing stress scenarios involves judgement, as does the choice of how frequently to stress test. A frequent criticism of stress testing is that scenarios do not in themselves contain information on the probability of their realisation, and are thus difficult to evaluate and combine with other risk measures. We propose using market prices as signals of the likelihood of market stress, and hence as guidance on the design and timing of stress tests, and support this with anecdotal evidence and with time-series statistical tests that incorporate full-blown crises as well as less dramatic stress events. Market signals can be used to complement macroeconomic warning signals of economic stress, which have a much longer lead time, measured in months or years rather than days or weeks.[2]

Crises are marked not only by sharp asset price moves, but also by various additional phenomena, such as volatility spikes and correlation breakdown. Volatility spreads to markets unrelated to the source of the original disturbance via market sentiment, hedging and the search for liquidity. Central bankers speak of systemic risk, in which the price discovery, credit allocation and payment processing functions of the financial system, particularly as carried out by commercial banks, are compromised. Currency crises, which have had a special role in post-war instability, are closely associated with systemic risk because they involve the banking system in an essential way through short-term capital inflows and outflows.

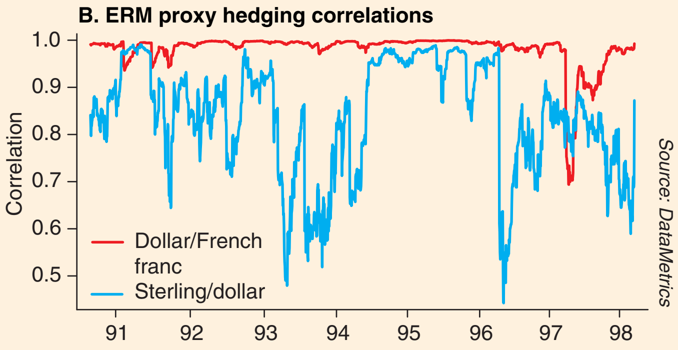
The rise in volatility can be quite dramatic. As panel A of figure 1 shows, short sterling volatility increased to more than 10 times its typical value during the 1992 Exchange Rate Mechanism (ERM) crisis, as the Bank of England attempted first to defend the ERM parities via penalty rates and then cut rates dramatically once the peg was broken and it was no longer obliged to follow Bundesbank tight-money policies.

Correlations may change dramatically during crises. Risk managers added 'correlation breakdown' to their terminology in 1998, when typical correlations changed rapidly. When the durability of a currency peg is largely but not perfectly credible, interest rates for the anchor currency are lower than for the pegged currency. It is tempting for market participants with fixed-income obligations in the pegged currency to seek funding in the apparently near-perfectly positively correlated anchor currency. This variety of the 'carry trade' comes to an abrupt halt when a peg is broken (see panel B for the example of the ERM crisis).[3]

1. Volatility and correlation in the ERM crisis



Exponentially weighted moving average volatilities of daily returns, with decay factor 0.94, at an annual rate. Price volatilities for three-month constant maturity CME Eurodollar and Liffe short sterling futures



Exponentially weighted moving average correlations of daily returns, with decay factor 0.94. Deutschmark/French franc vis-à-vis dollar/French franc and sterling/Deutschmark vis-à-vis sterling/dollar

Expectations and market prices during crises

A growing number of market participants, central banks and international institutions have implemented asset price-based indicators of market sentiment alongside assessments based on macroeconomic data and forecasts, observation of transactions flows, and anecdotal evidence. Some of these indicators, such as the use of the slope of the term structure of interest rates or spreads between credit-risky and default-risk-free interest rates to forecast inflation or turning points in the business cycle, are well established within traditional market analysis. Others, such as forward, futures and option prices, are quite new.[4]

Forward and futures prices are closely related in theory to expected future spot prices, but are poor predictors in practice due to changing risk and liquidity premiums. Option implied volatility surfaces are also forward-looking asset prices, but react more strongly to changes in expectations.

To see why option implied volatility is potentially a more powerful stress predictor than forward rates, imagine a risk-neutral world, in which the

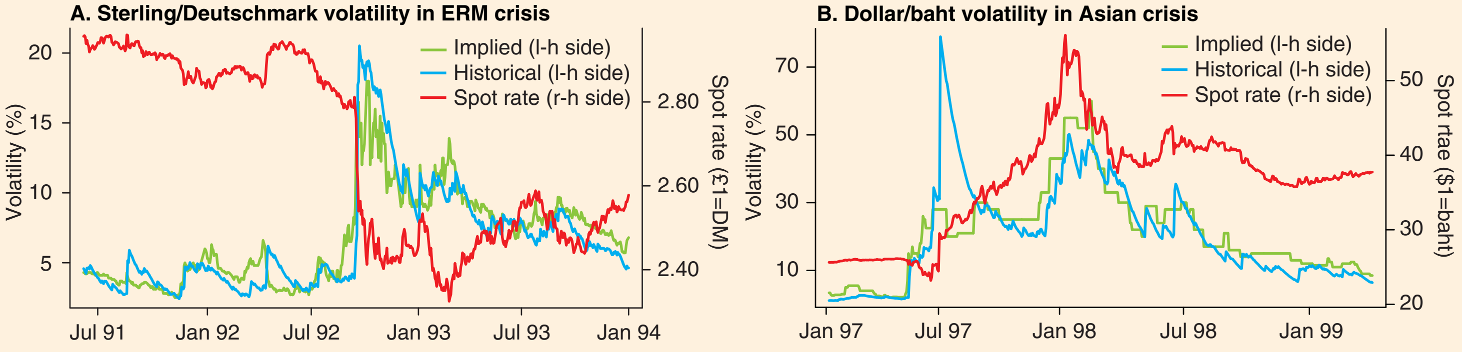
[1] *See Kupiec (1995) for a discussion of scenario choice and Mina & Xiao (2001), page 31 on, for a summary of current best practice*

[2] *See Kaminsky, Lizondo & Reinhart (1998) and the papers collected in the August 1999 issue of the* Journal of International Money and Finance *for examples of this approach*

[3] *Kim & Finger (2000) discuss the construction of stress tests when correlations are changing drastically*

[4] *A good example of the use of derivatives-based indicators are recent issues of the Bank of England's* Inflation Report and Financial Stability Review. *Market-based indicators are a potentially important response to the call for enhanced 'early warning capabilities' as part of improved crisis prevention in the International Monetary Fund communique of April 29, 2001*

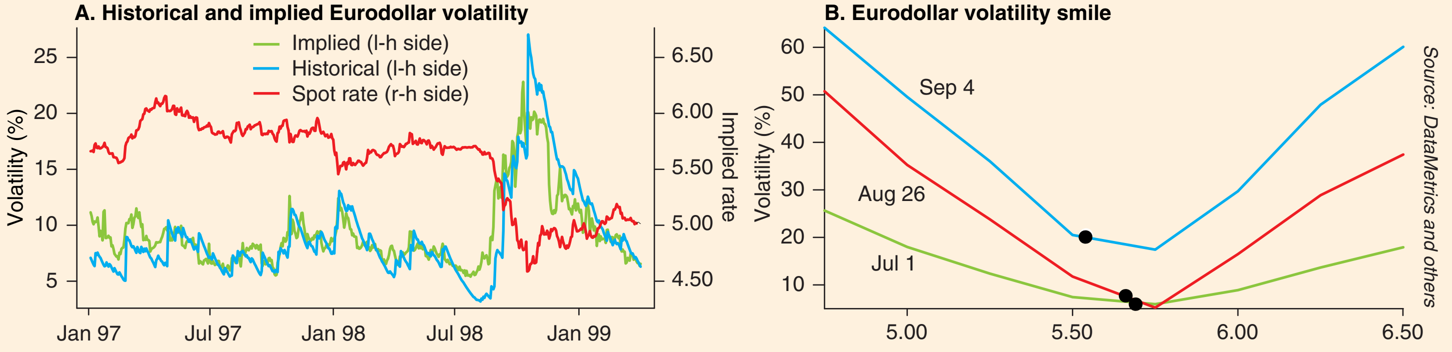
2. Implied and historical volatility in the ERM and Asian currency crises



Historical: exponentially weighted moving average volatilities of daily returns, with decay factor 0.94, at an annual rate. Implied: one-month at-the-money forward implied volatilities

Source: DataMetrics and others

3. CME Eurodollar volatilities during the LTCM crisis



Panel A: Historical: exponentially weighted moving average volatilities of daily returns, with decay factor 0.94, at an annual rate. Implied: three-month constant maturity at-the-money implied volatilities for CME Eurodollar options on futures. Panel B: implied volatilities of options on CME Eurodollar front contract. Heavy dot indicates futures settlement price

forward price is equal to the mean future price and implied volatility is equal to the return standard deviation. Market participants agree on the set of possible outcomes:

☐ A ('random walk'): there are two outcomes, in which the future asset price is either \$0.90 or \$1.10, with a probability of 50% each. The forward price will be \$1.00 and the implied volatility 10%.

☐ B ('jump-diffusion'): there are three outcomes, in which the future asset price is \$0.90 or \$1.10, with a probability 45% each, or \$0.50, with a probability of 10%. The forward price will be \$0.95 and the implied volatility 17.75%.

If the outcomes contemplated by the market change from scenario A to B, with the emergence of a small probability of a large price decline, the forward price drops by only 5%, but implied volatility nearly doubles. Implied volatilities are therefore a more sensitive indicator of anticipated stress than forwards and futures. While they are commingled with risk and liquidity premiums, these premiums are less apt to overwhelm the indicator properties of implied volatility than those of forwards and futures, because the reactivity of implied volatility to changes in market expectations is so much greater.

To see how sensitive implied volatilities are to expectations or anxieties about large market moves, let us focus again on the ERM crisis and the devaluation of the Thai baht. Figure 2 shows spot exchange rate levels and historical and implied volatilities for the sterling-Deutschmark and dollar-baht exchange rates during the crises. Both historical and implied volatility are typically mean-reverting, and generally not far apart.[5] Implied volatility anticipates the crises, in the case of sterling-Deutschmark rate by a few weeks, and in that of dollar-baht rate by a few days. Historical volatility reacts to crises, lagging implied slightly. Historical volatility spikes much higher than implied at first, but settles down rapidly after the massive asset level changes of the devaluation itself. Implied may continue to rise, and decays slowly, as the devaluations of recent decades have been more severe in the event than the market had anticipated. The markets expected a modest change in central parity for sterling, not an outright exit from the ERM followed by a 15% depreciation in a month. Nor did the markets expect an 85% depreciation of the baht within six months.[6]

Anticipation of stress also influences the volatility smile, the relationship between equally out-of-the-money puts and calls on the same underlying asset and with the same maturity. The degree of curvature of the volatility smile indicates a market perception of kurtosis in the return distribution of the underlying asset, while the skewness indicates a perception of skewness in the return distribution.

Tracing these changes over time for Chicago Mercantile Exchange (CME) Eurodollar futures shows how perceptions of the size and timing of rate cuts by the Fed evolved during the Russia-LTCM crisis. The prior change in key rates (a tightening) had been in March 1997, 16 months earlier, and the target Fed funds rate had been between 5.25% and 5.75% for about three years. Just before the onset of the crisis, therefore, implied and historical volatilities had been subdued (panel A of figure 3), while the smile had a modest degree of curvature and no visible skewness (panel B). As the crisis developed, the market anticipated a cut in the Fed funds rate as part of the classic central bank monetary easing response to financial crisis. Eurodollar implied volatility began to rise, well ahead of historical volatility, the curvature of the smile became more pronounced, and the

[5] *See Malz (2000/2001) on the statistical properties of implied volatility*
[6] *A detailed description of the cash market and option data is provided in Malz (2000)*

smile became skewed. By early September, options with positive payouts in the event of a 25-basis-point rise in the three-month forward rate had far lower implied volatilities than options with positive payouts in the event of a 25-basis-point decline in the forward rate.

A compact way to capture the information contained in forwards and implied volatilities is via estimates of risk-neutral probability distribution functions.[7] Figure 4 compares estimates of the probability of a 10% depreciation of the dollar-Mexican peso exchange rate over the subsequent month, estimated in three different ways:

- ☐ From the forward premium, using a simple 'peso problem' approach in which the market expects a 10% depreciation with a probability π and no change in the exchange rate with a probability $1 - \pi$.
- ☐ Using RiskMetrics historical volatilities with a decay factor of 0.97 and a drift rate of zero.[8]
- ☐ From the option-implied probability distribution, using a set of at-the-money and out-of-the-money option prices.

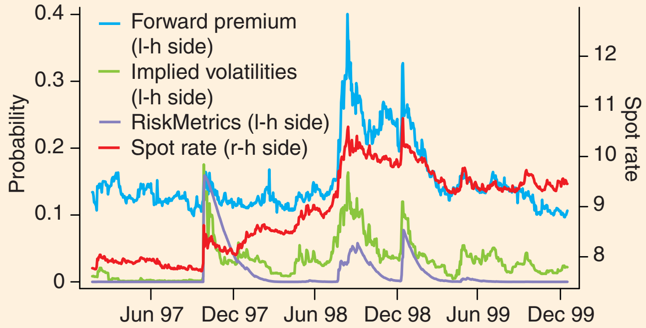
The forward-based estimates, always well over 10% even in the peso's quietest trading periods, are much higher than using the other two techniques. The RiskMetrics and the option-based estimates are not generally far from one another, but there are important differences that point to greater accuracy of the latter for tracking market expectations of extreme moves. The implied volatility-based forecasts always show a small positive probability, typically lead the RiskMetrics forecasts and generally display a more plausible variability, while the RiskMetrics forecasts fall to zero during quiet trading periods. These results underpin the importance of stress testing as a supplement to risk measures based on historical volatility such as VAR (as in Mina & Xiao, 2001).

Implied volatility warning signals: statistical tests

In a full-blown crisis, the early-warning properties of implied volatility are so clear they can be seen graphically, as in figures 2 and 3. We can also detect these predictive properties statistically in less extreme situations. We carried out statistical tests of the predictive ability of implied volatility for stress events for a range of assets, including major and emerging-market currency pairs, money markets, bonds, equity indexes and commodities.[9] We used squared returns as a metric for market stress, focusing on the kurtosis of the return distribution. If squared returns are plotted, a handful of large realisations is visible.

One test of predictive ability is the Granger causality test. A time series y is said to cause another time series x if past values of y help to predict the current value of x, that is, the conditional forecast variance of x can be reduced by including information on past y in the conditioning set along with past values of x. A standard test for causality is to set a lag length k, carry out the ordinary least squares regression:

4. Alternative measures of the probability of a 10% depreciation in the value of the Mexican peso against the dollar



See text for explanation — *Source: DataMetrics*

$$x_t = \sum_{i=1}^{k} \alpha_i x_{t-i} + \sum_{i=1}^{k} \beta_i y_{t-i} + u_t, \qquad t = 1, \ldots, T \tag{1}$$

and test the exclusion restrictions $\beta_1 = \beta_2 = \ldots = \beta_k = 0$, that is y fails to Granger cause x. In our application, squared returns are the dependent x variable, while implied volatilities and historical volatilities are the potentially causative y variables.

We find that for all assets, the null hypothesis that implied volatility Granger-causes – that is, systematically precedes – large-magnitude returns cannot be rejected, even at high confidence levels (columns one and two of table A). Interestingly, we also find confirmation of the predictive ability of RiskMetrics exponentially weighted moving average volatilities for large-magnitude returns (columns three and four), while in contrast, the predictive ability of conventionally calculated equally weighted historical volatilities was readily rejected (columns five and six).

These results suggest that much of the information about future squared returns in RiskMetrics and implied volatilities is common to both, raising the question as to what incremental information content implied volatility contains over and above that also contained in RiskMetrics volatilities. Columns seven and eight of table A display results of a causality test of ex-

[7] *See Jackwerth (1999) and Bliss & Panigirtzoglou (2000) for recent surveys of these techniques. An extension of this approach is the extraction of implied correlations between dollar-based exchange rates from implied volatilities of cross-currency options. See, for example, Campa & Chang (1998) and Lopez & Walter (2000)*

[8] *See Mina & Xiao (2001) for details on RiskMetrics volatilities*

[9] *This section summarises results presented in more detail in Malz (2000)*

A. Granger causality tests for squared returns, implied, historical and RiskMetrics volatility, and test of independence of an implied volatility signal and large returns

	Implied		Historical		RiskMetrics		Implied given RiskMetrics		Returns 99th pctile			Returns 99.9th pctile		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	χ^2 test			χ^2 test		
Asset	p-value	$\bar{R}^2$	p-value	$\bar{R}^2$	p-value	$\bar{R}^2$	p-value	$\bar{R}^2$	Q	p-value	HGD	Q	p-value	HGD
CME S&P 500	0.000	0.081	0.669	0.076	0.000	0.091	0.023	0.093	0.92	0.339	0.096	4.26	0.039	0.013
CME Eurodollar	0.000	0.042	0.004	0.022	0.000	0.029	0.000	0.043	4.78	0.029	0.012	0.71	0.401	0.071
Liffe short sterling	0.000	0.055	0.310	0.005	0.000	0.048	0.000	0 081	34.21	0.000	0.000	27.82	0.000	0.000
CBOT 30-year bond	0.000	0.059	0.000	0.035	0.000	0.041	0.000	0 066	11.37	0.001	0.001	11.03	0.001	0.002
Nymex crude light oil	0.000	0.054	0.015	0.034	0.008	0.035	0.000	0 057	6.24	0.012	0.006	10.92	0.001	0.000
Gold	0.080	0.162	0.044	0.163	0.000	0.238	0.152	0.241	6.27	0.012	0.006	0.13	0.717	0.152
Dollar/yen	0.000	0.079	0.213	0.061	0.000	0.083	0.009	0.088	0.38	0.536	0.309	0.13	0.707	0.128
Dollar/euro	0.000	0.055	0.209	0.032	0.000	0.046	0.000	0.056	2.08	0.149	0.037	8.26	0.004	0.003
Sterling/euro	0.000	0.110	0.012	0.078	0.000	0.091	0.000	0.112	3.21	0.073	0.013	0.11	0.735	0.104
Dollar/Mexican peso	0.000	0.255	0.863	0.133	0.000	0.205	0.000	0.299	3.45	0.063	0.017	10.16	0.001	0.001
Dollar/baht	0.000	0.023	0.280	-0.000	0.000	0.020	0.241	0.022	9.87	0.002	0.001	18.90	0.000	0.000

Note: p-values for test of exclusion restrictions; $\bar{R}^2$ is that of the unrestricted regression; HGD is 1 minus the CDF of the hypergeometric distribution

B. Two-way contingency table

	$N(B_{S\&P\ 500})$	$N(\sim B_{S\&P\ 500})$	Total
$N(A_{S\&P\ 500})$	2	61	63
$N(\sim A_{S\&P\ 500})$	5	742	747
Total	7	803	810

clusion restrictions on implied volatility in a regression also including lags of squared returns and RiskMetrics volatility. In almost all cases, implied volatility is shown to contain information regarding future squared returns, even after the information also in RiskMetrics volatility is accounted for.[10]

Another test of the predictive ability of implied volatility for stress events is to formulate a warning signal summarising whether implied volatility is high relative to recent levels and rising fast, and examine whether that signal helps to predict the event that cash asset returns are unusually high or low. As a measure of how high implied volatility is currently, we calculate for each date the mean and standard deviation of implied volatility over the past year (250 trading days) and define as high any implied volatility higher than one standard deviation above its one-year mean. As a measure of how fast implied volatility is rising, we make use of the mean root square or volatility of daily logarithmic changes in implied volatility, the 'volatility of volatility'. High returns are defined as those exceeding 2.33 standard deviations (99th percentile) or 3.09 standard deviations (99.9th percentile) in magnitude.

Let $T = \{t_1, t_2, \ldots, t_N\}$ be the set of dates in the sample for an asset, using weekly data. On each observation date $t_\tau \in T$, if implied volatility is high and has risen more than 0.675 volatilities of volatility (75th percentile) over the past week, the implied volatility is deemed to have sent a warning signal that high returns are likelier to occur over the next week. The signal is thus an event $A = \{t_\tau \in T\text{: IV high and rising}\}$ in T, as is the event $B = \{t_\tau \in T\text{: high returns}\}$. The intersection of the events $A \cap B$ is the set of dates on which the signal has been accurate.

Consider the S&P 500 index, for which we have 810 weekly observations. On 63 occasions, implied volatility was high (more than one standard deviation above its recent mean) and rising (by 0.675 volatilities of volatility over the previous week). The signal lights up frequently enough to be useful, and not so frequently as to be meaningless. On 22 occasions, returns over the subsequent week were more than 2.33 standard deviations in magnitude, of which three coincided with high and rising implied volatility. On seven occasions, returns were more than 3.09 standard deviations, of which two coincided with high and rising implied volatility. For 99.9th percentile returns, we can summarise these results in a two-way contingency table (see table B). The null hypothesis of independence between events A and B:

$$\mathcal{H}_0 : P\left\{t_\tau \in B \middle| t_\tau \in A\right\} = P\left\{t_\tau \in B\right\}$$

that is, the hypothesis that implied volatility has no signalling value for large-magnitude returns, can be tested two ways. One is a chi-square test using Pearson's Q. Under $\mathcal{H}_0$, the maximum likelihood estimate of the probability of $A \cap B$ is 63/810 7/810. Under the alternative hypothesis:

$$\mathcal{H}_1 : P\left\{t_\tau \in B \middle| t_\tau \in A\right\} \neq P\left\{t_\tau \in B\right\}$$

the estimate of the probability of $A \cap B$ is 2/810. Pearson's Q is the sum, over the four cells in the contingency table, of the differences between these estimates.

Another test of the null is based on the hypergeometric distribution (HGD). Under $\mathcal{H}_0$, the probability of the event $A \cap B$ follows the hypergeometric distribution. Returning to the example of the S&P 500 index, the probability of drawing three or more high-return observations in a sample of 63 from a population of 810 containing 22 high-return observations is 8.4%, if high-return dates are as frequent in the entire population as in the high-and-rising implied volatility sub-population.

Table A displays Pearson's Q and it's p-value, as well as one minus the cumulative distribution function of the hypergeometric distribution evaluated at $N(A \cap B)$, which can be treated as the p-value of a test of $\mathcal{H}_0$. Apart from the dollar/euro exchange rate, the null hypothesis is rejected at high confidence levels for all assets for either returns at the 99th percentile or at the 99.9th percentile, or both. If the null is rejected for returns at the 99.9th percentile but not the 99th percentile, implied volatility is likely to rise in anticipation of extremely large price moves, but less likely to move in anticipation of more routine large moves. Both the chi-squared test and the test based on the hypergeometric distribution indicate that the implied volatilities provide a useful warning signal for large-magnitude returns.[11]

Conclusion

One of the most fundamental changes in risk measurement in recent years has been the emergence of stress testing as a crucial supplement to classical loss measures such as VAR. Stress testing has retained something of an *ad hoc* character, as the design and timing of stress scenarios, as well as how stress test results should be combined with one another and with classical measures, remains a matter of judgement. The use of implied volatility as a stress indicator can increase the role of measurement over judgement in this process. In particular, it permits a market-based assessment of the likelihood of market stress. Implied volatility-based indicators and warning signals should be part of a wider toolkit of asset price-based stress indicators derived from forwards and futures, interest rate term structures and spreads, and option prices. ■

Allan Malz is head of research at the RiskMetrics Group in New York

Comments on this article may be posted on the technical discussion forum on the Risk website at http://www.risk.net

[10] The exceptions are gold and dollar/baht. In the case of gold, the results are dominated by the 15% price rise following the surprise announcement on September 26, 1999 of limits on gold sales by European central banks. For the Thai baht, although figure 2 makes it clear that implied did in fact anticipate historical volatility, the observation period is comparatively short and contains only one long run-up and one long wind-down of implied and historical volatilities

[11] Changing the signal (for example, changing the volatility-of-volatility threshold for rising implied volatility to the 90th rather than the 75th percentile) can change the results by increasing the number of false positives or reducing the number of successfully anticipated large returns. The thresholds used in this article provided good performance across all asset classes, but could in further work be tailored to individual assets

REFERENCES

Bliss R and N Panigirtzoglou, 2000
Testing the stability of implied probability density functions
Working paper 114, Bank of England

Campa J and P Chang, 1998
The forecasting ability of correlations implied in foreign exchange options
Journal of International Money and Finance 17(6), pages 855–880

Jackwerth J, 1999
Option-implied risk-neutral distributions and implied binomial trees: a literature review
Journal of Derivatives 7(2), pages 66–82

Kaminsky G, S Lizondo and C Reinhart, 1998
Leading indicators of currency crises
IMF Staff Papers 45(1), pages 1–48

Kim J and C Finger, 2000
A stress test to incorporate correlation breakdown
Journal of Risk 2(3), pages 5–19

Kupiec P, 1995
Stress testing in a value at risk framework
Journal of Derivatives 6(1), pages 73–84

Lopez J and C Walter, 2000
Is implied correlation worth calculating? evidence from foreign exchange options
Journal of Derivatives 7(3), pages 65–82

Malz A, 2000
Do implied volatilities provide early warning of market stress?
RiskMetrics Journal 1(1), pages 41–60

Malz A, 2000/2001
Vega risk and the smile
Journal of Risk 3(2), pages 41–63

Mina J and J Xiao, 2001
Return to RiskMetrics: the evolution of a standard
RiskMetrics Group