Incorporating Stress Tests into Market Risk Modeling

JOSE RAMON ARAGONÉS, CARLOS BLANCO, AND KEVIN DOWD

JOSE RAMON ARAGONÉS
is professor of finance at Complutense University of Madrid.

CARLOS BLANCO
is global support and educational services manager at Financial Engineering Associates, Inc. in Berkeley, CA.

KEVIN DOWD
is professor of financial risk management at Nottingham University Business School in Nottingham, U.K.

"The possibility that market participants are developing a degree of complacency or a feeling that (risk management) technology has inoculated them against market turbulence is admittedly somewhat disquieting. Such complacency is not justified. In estimating necessary levels of risk capital, the primary concern should be to address those disturbances that occasionally do stress institutional solvency—the negative tail of the loss distribution that is so central to modern risk management. As such, the incorporation of stress scenarios into formal risk modeling would seem to be of first-order importance. However, the incipient art of stress testing has yet to find formalization and uniformity across banks and securities dealers. At present, most banks pick a small number of ad hoc scenarios as their stress tests. And although the results of the stress tests may be given to management, they are, to my knowledge, never entered into the formal risk modeling process."

— Alan Greenspan [2000]

While VaR models have proven themselves to be very useful risk management tools, recent financial debacles have also highlighted their limitations—in particular, their excessive dependency on history or unrealistic statistical assumptions. The natural response to these limitations is for firms to resort to stress tests to complement the results of their VaR analyses. Stress tests are exercises to determine the losses that might occur under unlikely but plausible circumstances, and there has been a dramatic increase in the importance given to stress testing since the east Asian crisis and the LTCM affair. Indeed, many firms and regulators now regard stress tests as no less important than VaR methods for assessing a firm's risk exposure.

Stress tests can provide useful information about a firm's risk exposure that VaR methods can easily miss, particularly if VaR models focus on "normal" market risks rather than the risks associated with rare or extreme events. Such information can be fed into strategic planning, capital allocation, hedging, and other major decisions.

However, traditional stress testing is done on a stand-alone basis, and the results of stress tests are evaluated side-by-side with the results of traditional market risk (or VaR) models. This creates problems for risk managers, who then have to choose which set of risk exposures to "believe." There is also the related problem that risk managers often don't know whether to believe their stress test results, because the stress tests exercises give them no idea of how likely or unlikely stress-test scenarios might be.

Berkowitz [1999] suggests a solution to this problem by integrating stress testing into formal risk modeling by assigning probabilities to stress-test scenarios. The resulting risk estimates incorporate both traditional market risk estimates and the outcomes of stress tests, as well as the probabilities of each. They therefore give risk managers a single, integrated set of risk estimates to work with. Of course, these estimates are dependent on the judgmental factors that go into stress testing and into the evaluation of scenario probabilities, but in our view it is better to incorporate our judgments of stress-test events than to ignore them completely in risk modeling; it is better to be approximate and probably right in our risk assessments, than to be precise and probably wrong.

However, given the drawbacks of Value at Risk as a risk measure, it makes sense to complement scenario analyses by estimating the Expected Tail Loss (ETL) as well (see, e.g., Artzner et al. [1999]). Furthermore, since we are concerned about extreme events, it also makes sense to derive our risk estimates using Extreme Value Theory.

TRADITIONAL APPROACHES TO STRESS TESTING

Types of Stress Test

The essence of stress testing is the creation of user-defined scenarios, fed into a calculation engine to produce estimates of the profits or losses that can be expected under these scenarios. Stress tests fall into three main types, which differ in how the scenarios are constructed:

- The first type uses scenarios from recent history, such as the 1987 equity crash, the ERM crises of 1992-1993, the bond market crash of 1994, the 1994 peso crisis, and the 1997 east Asian crisis and its aftermath. In these scenarios, we simulate the effect on profit or loss of repeats of past historical events.
- The second type uses predefined or set-piece scenarios that have proven to be useful in practice. We simulate the impact on P/L of a fall in the stock index of x standard deviations, or of a change in an exchange rate of y%, or of a yield-curve shift of so–many basis points, and so on
- The third type are mechanical-search stress tests. These use automated routines to over prospective changes in risk factors, evaluate P/L under each set of risk-factor changes, and report the worst-case results (see, e.g., Dowd [1998], pp. 126-129).

Problems with Stress Tests

However, stress tests have a number of problems. Perhaps the most obvious is that stress tests are inevitably subjective because they depend on scenarios chosen by the stress tester. As a result, the value of stress testing depends critically on the choice of scenarios and therefore on the skill of the modeler. This subjectivity also makes it difficult to assess the value of a stress test objectively, and poses obvious problems for senior management, regulators, and other interested parties trying to assess a firm's stress testing procedures.

A related problem is that the results of stress tests are difficult to interpret because they give us no idea of the probabilities of the events concerned, and in the absence of such information we often don't know what to do with them. Suppose for instance that stress testing reveals that our firm will go bust under a particular scenario. Should we act on this information? The only answer is that we can't say. If the scenario is very likely, we would be very unwise not to act on it. But if the scenario was extremely *unlikely*, then it becomes almost irrelevant, because we would not usually expect management to take expensive precautions against events that may be too improbable to worry about. So the extent to which our results matter or not depends on unknown probabilities. As Berkowitz [1999] nicely puts it, this absence of probabilities puts "stress testing in a statistical purgatory. We have some loss numbers but who is to say whether we should be concerned about them?"

This problem leads to another: it implies that we have two sets of separate risk estimates—probabilistic estimates (e.g., such as VaR), and the loss estimates produced by stress tests—and no way of combining them. How can we combine a probabilistic risk estimate with an estimate that such-and-such a loss will occur if such-and-such happens? The answer, of course, is that we can't. We therefore have to work with these estimates more or less independently of each other, and the best we can do is use one set of estimates to check for prospective losses that the other might have underrated or missed.

Another problem is that the methodology of stress testing is still in its infancy, and the approaches commonly used are often open to question. For example, a very common stress-testing procedure is to "shock" certain prices or returns to particular values, assume that other prices take their usual values, and derive the firm's risk exposure accordingly. This approach is conceptually straightforward and easy to carry out, but is open to objection because it ignores correlations between the stressed prices and other prices, and recent empirical evidence suggests that the failure to allow for these correlations can make a big difference to results (Kupiec [1998], p. 8).

There is also the problem that stress-test procedures are very difficult to backtest. As Schachter [1998] points out,

> Neither the completeness nor the reliability of the information provided [by stress tests] can be scientifically assessed. Also hypothetical stress scenarios cannot be "validated" based on actual market events. That is, even when the events specified in a hypothetical scenario actually occur, there is usually no way to apply what was "right" and "wrong" in the scenario to other hypothetical scenarios to improve them. These limitations are not shared by value at risk models, which are statistically based, and where it is possible to construct statistical confidence intervals around the VaR estimate, and to conduct meaningful "backtests" of the model's predictions.
>
> — Schachter [1998], p. 5F-8

The net result is that we cannot rely on established backtesting procedures to tell us the difference between a good approach to stress testing and a bad one.

INCORPORATING STRESS TESTS INTO MARKET RISK MODELING

An answer to most of these problems is to bring stress testing into the market risk modeling process used by the firm—to unify stress testing and probabilistic risk estimation—and we can do so by putting probabilities to the scenarios used in stress testing. Once our scenarios are put in probabilistic form, we will have an idea of whether to take a particular scenario seriously, we will have one unified and coherent risk measurement system rather than two incompatible ones, and we will be able to apply backtesting procedures to impose some (albeit limited) check on our scenarios. Inevitably, the choice of scenarios will remain subjective, but even there, the need to assign probabilities to scenarios will impose some discipline on risk managers and put pressure on them to distinguish between those scenarios that matter and those that don't.

The key is therefore to assign probabilities to defined stress events. Suppose that we have m stress scenarios. Following Berkowitz [1999], assume that scenario i will occur with probability α. We can then presume that our returns, x, come from a combined distribution that takes the form $f_{stress,1}(.)$ with probability $\alpha_{1,\ fstress,2}(.)$ with probability $\alpha 2$, and so forth, and f(.) with probability $1-\sum_{i=1}^{m}\alpha_i$. $f_{stress,i}(.)$ is the probability of scenario i occurring, and f(.) is the probability that none of the specified scenarios occur. $f_{stress,i}(.)$ and f(.) are n-variate distributions of the n underlying risk factors. We could think of the latter as the distribution generated by a conventional historical simulation analysis or Monte Carlo simulation, and the $f_{stress,i}(.)$ could be distributions that capture stress scenarios not covered by the conventional distribution (e.g., correlation breakdown, etc.).

Berkowitz went on to suggest that we could model the realization of the $f_{stress,1}(.)$ or f(.) distribution by means of a dummy uniform random variable that takes a value between 0 and 1. For example, if this variable takes a value from 0 to α_1, the distribution takes the form fstress,1(.) , if the variable takes the value bigger than α_1 up to α_2, it takes the form $f_{stress2}(.)$, and so on. If we wish, we can also modify the α-values, the probabilities of scenarios, to allow them to depend on other factors (e.g., history, market conditions, etc.).

STAGES OF THE RISK MODELING PROCESS

There are therefore four key stages in our risk modeling process.

Stage One: We go through our stress testing in the traditional way, and the outputs of this process will be a set a of realized profits/losses associated with each scenario.

Stage Two: Once we have gone through our scenarios and established their P/L outcomes, we go through a second, very judgmental, process and assign probabilities to each of our scenarios.

Stage Three: We then go through a formal risk modeling process of the traditional kind, and model our risks using appropriate risk measurement techniques. We can think of the outcome of this process as a set of P/L figures and their associated probabilities.

Stage Four: We now have all the information we need, so we bring together our two sets of P/L figures and two sets of associated probabilities, and carry out an integrated risk estimation.

We now present the integrated risk estimation approach to VaR and Expected Tail Loss, and give a concrete example.

INTEGRATED RISK ESTIMATION

Another problem with traditional approaches to risk modeling is that it does not take proper account of the more extreme observations in our data set. This shortcoming can be rectified by using Extreme Value (EV) theory, which provides a tailor-made approach to the estimation of low frequency events with limited data. EVT is very useful because it guides us in the selection of distribution to use when modeling our risks. This can be very useful, because we often don't know the distributions from which our P/L observations are drawn. If we have P/L observations x from an unknown distribution, then EVT tells us that subject to certain innocuous conditions, the distribution of excess returns x beyond a threshold u, converges asymptotically to a Generalized Pareto distribution:

$$G_{\xi,\beta}(x) = \begin{Bmatrix} 1-(1+\xi x/\beta)^{-1/\xi} \\ 1-\exp(-x/\beta) \end{Bmatrix} \text{ if } \begin{cases} \xi \neq 0 \\ \xi = 0 \end{cases}$$

where $\beta > 0$, and the support is $x > 0$ when $\xi \geq 0$, and $0 \leq x < \beta/\xi$ when $\xi < 0$. The parameter β corresponds to the standard deviation, and the parameter ξ gives an indication of the heaviness of the tails: the bigger ξ, the heavier the tail. This parameter is known as the tail index, and the case of most interest in finance is where $\xi > 0$, which corresponds to the fat tails commonly founded in financial return data.

EVT tells us that the limiting distribution of extreme excess returns always has the same form—whatever the distribution of the parent returns from which our extreme returns are drawn. It is important because it allows us to estimate extreme probabilities and extreme quantiles, including VaRs and ETLs, without having to make strong assumptions about the full shape of the unknown parent distribution.

The size of the tail u can be determined judgmentally or by other methods (e.g., such as the historical-simulation (HS)–EV approach suggested by Aragonés, Blanco, and Dowd [2000a, b]. Once the size of the tail has been determined we can then use nonlinear regression analysis to obtain a best fit estimate for β and ξ.

Following McNeil [1999], the formula used to obtain VaR for a given probability q is:

$$\hat{VaR}_q = u + \frac{\hat{\beta}}{\hat{\xi}}\left(\left(\frac{n}{N_u}(1-q)\right)^{-\hat{\xi}} - 1\right)$$

where n represents the number of observations and N_u is the number of observations in the tail beyond u. The ETL is:

$$ETL = VaR + E\left(X - VaR \middle| X > VaR\right)$$

which we can estimate by plugging estimates of the relevant parameters into:

$$ETL = [VaR + (\beta - \xi u)]/(1-\xi)$$

In effect, we use first estimate of the parameters of our EV distribution, and then project the tail out beyond our sample, thereby allowing us to estimate extreme risk measures and the probabilities associated with them.

MODELING ENERGY PORTFOLIO RISK WITH AND WITHOUT USER-DEFINED SCENARIOS

To illustrate the usefulness of the integrated stress-VaR approach, we calculated VaR and ETL for an energy derivatives portfolio using Historical Simulation, and then calculated the same parameters adding two hypothetical high-loss scenarios to the historical set of P/Ls.

In an HS context, an integrated stress-VaR approach involves adding the scenario to our historical data set, assigning a probability to that scenario so that the probability of all scenarios adds to one. Our portfolio contains energy swaps, futures, and option contracts. The risk factors are the forward curves with the 12 nearby contracts for the West Texas Intermediate Crude prices, Natural Gas Prices, and Heating Oil Prices from the New York Mercantile Exchange. We use 300 daily returns in our analysis.

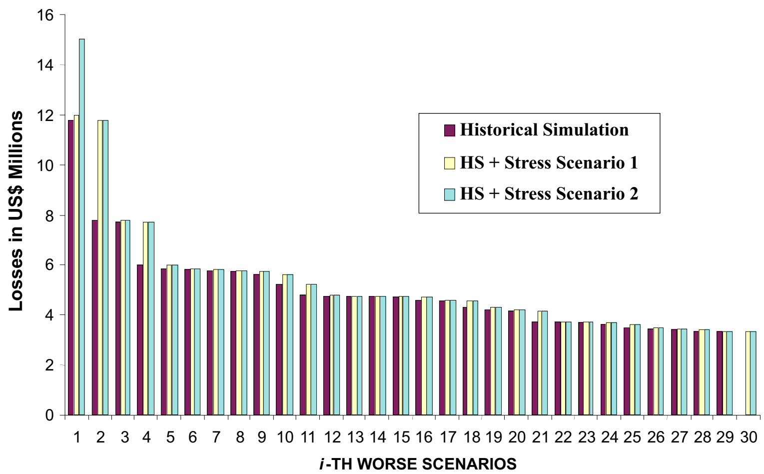
Stage One: Stress Testing. We conduct stress tests based on two scenarios.

Scenario #	Description	Profit/Loss
Scenario 1	OPEC decision to reduce production	Loss US$ 12 million
Scenario 2	Breakdown of Middle East peace negotiations	Loss US$ 15 million

In order to calculate expected P/L for each scenario, we would need to determine the impact in prices, volatilities, and correlations of each of those scenarios, and revalue our portfolio under each set of possible futures prices. The scenarios chosen would result in losses greater than the largest historical loss in the sample.

Stage Two: Assign probabilities to each scenario. For simplicity, we have given that scenario the same weight as the other 300 historical scenarios in the sample (0.33%). This step of the analysis is crucial, and insti-

Exhibit 1
Losses for Historical Simulation VaR and Integrated Stress-VaR



tutions should allocate resources to try to get the best estimates of occurrence of catastrophic scenarios.

Scenario 1	OPEC decision to reduce production	0.33%
Scenario 2	Breakdown of Middle East peace negotiations	0.33%

Stage Three: Perform Historical Simulation Analysis to derive a set of P/L figures and their associated probabilities.

Stage Four: Once we have the historical and scenario P/L figures and their respective probabilities, we can carry out an integrated risk estimation. We calculate EV-VaR and ETL for a 99% confidence level (see Exhibit 1).

Exhibit 2 shows portfolio losses using Historical Simulation alone, HS plus scenario 1 loss, and HS plus scenario 2 loss. The results clearly show the added value of the unified VaR—Stress Test approach over traditional VaR analysis.

Exhibit 2
Stress-VaR Results

	Historical Simulation	HS + Scenario 1	HS + Scenario 2
VaR	7.85	8.68	9.17
ETL	11.35	13.59	16.05
Pareto ξ	0.26	0.33	0.42
Pareto β	1.42	1.55	1.49

Note: Estimated with a 1-day horizon, 99% VaR quantile, and u corresponding to a 10% tail.

The first column shows the HS-VaR at the 99% level, and the second and third columns show the VaRs and ETLs obtained after adding one or more additional scenarios that would result in losses of US$12 million and US$15 million, respectively.

CONCLUSIONS

If VaR is used for capital allocation, limits setting, and performance measurement, it is important that the scenarios used are transparent to risk managers, senior management, and traders, whose actions may be constrained by those limits. Avoiding the "black box syndrome" will facilitate the acceptance of risk models for active risk management. Incorporating stress scenarios into the VaR model also helps fill in "risk holes" and reduce the ability of traders to "game" the risk model.

One of the main advantages of the suggested framework is that scenarios become a direct input into the risk model, instead of being generated internally. That adds transparency to the risk measurement process, as the different participants are fully aware of the scenarios being used in the VaR calculation and the probability assigned to each of them.

The new approach suggested here has two big attractions—it is theoretically correct, in the sense that it makes use of EVT in a theoretically sound way, and it incorporates user-defined scenarios. However, to make good use of this approach, we must also bear in mind that we need a decent run of data, typically several years or more of suitable daily historical observations. This condition is easily met in many markets, but can be a problem in immature markets (e.g., energy, IPOs, etc.), or markets that experience structural changes (e.g., in exchange-rate regimes).

The introduction of subjective scenarios with assigned probabilities allows us to create a more comprehensive picture of risk including all available information to the risk managers. The choice of those scenarios and the probabilities assigned to them will determine the quality of the risk analysis. However, it is important to remember that it is better to have a model that it is approximately right than one that it is accurate from a mathematic point of view, but wrong.

Finally, this approach helps to alleviate the main limitation of all HS-based approaches to VaR—that they only pay attention to real historical events. The introduction of stress tests into the formal risk model gives us a convenient and plausible way to include events that could plausibly occur, but did not actually occur, in our data set.

ENDNOTE

The authors thank the FEA Support and Financial Engineering teams, particularly Angelo Barbieri and Robert Erickson, for helpful comments and software assistance.

REFERENCES

Aragonés, J.R., and Carlos Blanco. "Complementing VaR with Stress Tests." *Derivatives Week,* August 9, 1999, pp. 5-6.

Aragonés, J.R., Carlos Blanco, and Kevin Dowd. "Extreme Value VaR." *Derivatives Week*, March 6, 2000a, pp. 7-8.

——. "Extreme Value VaR (2)." *Derivatives Week*, March 20, 2000b, pp. 8-9.

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath. "Coherent Measures of Risk." *Mathematical Finance,* 9 (1999), pp. 203-228.

Bank for International Settlements. *Stress Testing by Large Financial Institutions: Current Practices and Aggregation Issues.* Basel, Switzerland: Committee on the Global Financial System. April 2000.

Berkowitz, Jeremy. "A Coherent Framework for Stress-Testing." Manuscript, Board of Governors of the Federal Reserve, July 1999.

Dowd, Kevin. *Beyond Value at Risk: The New Science of Risk Management.* Chichester and New York: John Wiley & Sons, 1998, pp. 126-129.

Greenspan, Alan. "Remarks [on Banking evolution]." 36th Annual Conference on Bank Structure and Competition of the Federal Reserve Bank of Chicago. Federal Reserve Bank of Chicago, Chicago, Illinois, May 4, 2000.

Kupiec, Paul."Stress Testing in a Value at Risk Framework." *The Journal of Derivatives,* 6 (1998), pp. 7-24.

McNeil, A.J. "Extreme Value Theory for Risk Managers." *Internal Modeling and CAD II, RISK Books,* 93-113. http://www.math.ethz.ch/~mcneil/ftp/cad.pdf.

Schachter, Barry. "The Value of Stress Testing in Market Risk Management." Derivatives Risk Management Service, March 1998, 5F-1-5F-11.

Financial Engineering Associates. *VaRworks 3.0 Technical Document.* Berkeley, CA: Financial Engineering Associates, 2000.