VAR and STRESS TESTING

Measurement of Portfolio Risk

There are two potential ways to measure the overall risk of a firm's positions: (1) a statistically based approach called value-at-risk (VAR), (2) an approach based on economic insight rather than statistics, called stress-testing or scenario analysis. We will discuss the potential uses of such overall measures of firm risk in some detail, but first we will explore the methodology required to make such measurements. For the time being we will just point out 2 major advantages of overall firm risk measures relative to more traditional measures of risk such as value of a basis point or the greeks:

1. Traditional measures do not allow senior managers to form conclusions as to which are the most important risks currently facing the firm. It is not possible to meaningfully compare the value of a basis point size in two different currencies, since this comparison does not reflect the relative size of potential interest rate moves in the two currencies. Both VAR and stress-testing give a measure which combines the size of position and size of potential market move into a size of potential impact of firm P&L.

2. Traditional measures do not interact with one another. Should you add up the risks under different measures into some total risk? Clearly this would be wrong because it would ignore the effect of correlation between market factors. Both VAR and stress-testing account directly for correlation between market factors.

We will first discuss the methodology of statistical measures, VAR, and then discuss the methodology for non-statistical measures, stress-testing.

VAR Methodology

Since statistical overall risk measures first began to be calculated by financial firms, about 20 years ago, three methods have dominated:

1. Direct measurement of P&L distribution

2. Calculation of P&L distribution based on historical statistics representing the variance and covariance of market variables and the current size of position exposures to each of these market variables. So if $s_i$ represents the firm's exposure to each market variable, $\sigma_i$ represents the volatility of each market variable, and $\rho_{i,j}$ represents the correlation coefficient between each pair of market variables, the volatility of overall firm P&L is calculated as

$$\sqrt{\sum_{i,j} s_i s_j \sigma_i \sigma_j \rho_{i,j}}$$

The P&L distribution can now be calculated from this volatility.

3. Simulation of P&L distributions based on a selected set of possible moves of market variables and the current size of position exposure to each of those market variables. So if $s_i$ represents the firm's exposure to each market variable, $m_i$ represents the size of move of each market variable in each considered scenario, and $p_j$ represents the probability assigned to each scenario, with $\sum_j p_j = 1$.

Then the P&L movement in each scenario is calculated by

$$\sum_j s_i m_{i,j}$$

And the P&L distribution is calculated by multiplying each of these terms by its respective $p_j$.

The use of direct measurement of P&L distribution is still widely used, as can be seen from the frequent use of histograms of daily P&L distributions published in annual reports of financial firms. It has the advantage of simplicity of calculation, not having to make any use of models or statistical assumptions, and ability to capture effects of the trading culture (e.g., does management respond to periods of greater market volatility by reducing position size) that the competing methods do not. It is also the only method which is available for measuring risk when access to details of trading positions is not available (e.g., measurement
of a hedge fund's risk by one of its investors). But its inability to take into account the possibility that current position taking may be radically different than historical position taking renders it close to useless as a stand-alone risk measure, though it is still valuable as a complement to other measures.

The use of the variance-covariance method, which was popularized by J.P. Morgan under the brand name Risk Metrics, has now been virtually abandoned by sophisticated financial firms. The primary reason for this is that relative to the simulation method, the variance-covariance method provides very little flexibility in evaluating the contribution of non-linear positions, notably options positions, to P&L distributions. Secondary reasons are (1) the greater difficulty that the variance-co-variance method has in dealing with the fat-tailed distributions normally encountered in financial markets (the formula for combining distributions of individual variables assumes that variables are normally distributed, so fat-tails can only be accommodated by assuming and calculating mixtures of normal distributions or some generalization of the normal distribution — see Dowd, Chapter 3, section 4 for details); (2) the inability of variance-covariance to pick up the phenomenon, often observed in financial markets, that the largest changes in variables often cluster together (e.g., the 1987 stock crash) to a greater degree than will be indicated by correlation coefficients (i.e., the joint distribution is not binomial — see Shaw, "Beyond VAR and Stress Testing" for further discussion); and (3) the realization that almost all the benefits of simplicity and speed of computation claimed for variance-covariance relative to simulation were based on fallacious comparisons. As will be seen in our discussion of simulation methodology, the degree of simplicity and speed of computation is largely determined by the choice of the user. To achieve a level of accuracy similar to that obtained by variance-covariance, simulation is at least as simple and fast to compute as variance-covariance. Simulation offers the flexibility, which variance-covariance does not, of increasing accuracy as a trade-off against simplicity and computation time, but having more flexibility can surely not count as a disadvantage.

Currently, the only users of variance-covariance would be smaller firms which do not hold significant options positions and who wish to outsource the market data component of their VAR computations. For such firms, variance-covariance does offer the distinct advantage that they only need to obtain volatilities and correlations rather than the day-by-day pricing histories required for simulation, a considerable savings in amount of data to be transferred as long as the number of different products a firm deals in does not grow to the point that the number of correlations needed get larger than the number of historical data points.

**Details of the Simulation Methodology for VAR**

Remember that the simulation approach consists of determining a number of possible scenarios, to be indexed by \( j \), determining the size of move of each market variable in each scenario \( m_{i,j} \), and then calculating \( \sum_i s_i m_{i,j} \) as the firm's total P&L movement in each scenario. The steps in a simulation of VAR consist of (1) determining a set of scenarios specified by the size of move in each of a set of underlying market variables and a probability to be assigned to each set, (2) translation from the size of move of underlying market variables to size of move for all market variables, and (3) calculation of the P&L distribution. There are 2 alternative approaches to the first step — historical simulation and Monte Carlo simulation. The decisions to be made for the second and third steps do not depend on the choice made for the first step. We will discuss each step in some detail.

**Step 1: Determine Underlying Market Volatilities**

The historical simulation approach is quite simple, a group of historical periods is chosen and the observed size of market move in each of these historical periods constitute the scenarios. So, for example, you could choose 300 scenarios consisting of all the most recent one business day changes in market variables — the changes in market variables from 6/7/99 to 6/8/99 would be one scenario, the change from 6/8/99 to 6/9/99 another scenario, and so forth. Or one could choose all the ten business day changes.

Scenario probabilities can be assigned based on perceived relevance to the current market situation — for example, greater probability weight could be assigned to more recent historical periods.

Historical simulation offers a large advantage in terms of simplicity — simplicity of implementation, simplicity of assumptions, simplicity of explanation. The advantage in terms of assumptions is that no
modeling assumption needs to be made beyond the assumption that the immediate future will resemble the past. There is no parameterization and no assumptions about distribution shape (e.g., normality). If fat tails or clustering of large moves between variables are present in the historical data, they will be reflected in the simulation.

The advantage in terms of explanation is that any questions raised by traders or managers concerning a VAR which seems too high can be easily traced to a subset of specific historical dates which would show large losses against the current firm holdings. Disagreement can be quickly focused on accuracy of data for a few specific dates or on arguments about the probabilities to be assigned to repetition of particular historical events. By contrast, the variance-covariance approach and the Monte Carlo simulation approach make it far more difficult to resolve such questions.

This advantage of simplicity of historical simulation also underlies its primary disadvantage — the VAR produced is dominated by market moves on a few specific historical days. If a particular combination of market events did not occur in the historical period being considered, it cannot contribute to VAR. You cannot overcome this problem by just expanding the historical period you are considering. Data availability tends to get sparse once you go back more than a few years, because of failure to retain data, because data becomes more difficult to "clean" the further back you go in time, and because some currently traded instruments may not have histories which go back that far.

This disadvantage of generating scenarios utilizing the historical method is the primary argument in favor of the Monte Carlo method. The Monte Carlo method starts with a specification of the underlying market variables which is similar to that of the variance-covariance approach, but may have a richer specification of each single variable than just a volatility — a multi-parameter specification allows the generation of distributions which are fat-tailed; see, for example, the article by Shaw, "Beyond VAR and Stress Testing." Monte Carlo techniques are then used to generate a set of scenarios which fit the desired statistical specifications. Usually, users of Monte Carlo simulation want to take advantage of the flexibility it offers to generate many more scenarios than can be practically generated with historical simulation. This has led to the incorrect assertion that Monte Carlo simulation requires more scenarios than historical simulation. In fact, there are strong reasons to believe that Monte Carlo simulation will deliver a greater degree of accuracy than historical simulation for the same number of scenarios employed, in addition to which Monte Carlo simulation offers the flexibility of achieving even greater accuracy if the greater expense of running more scenarios is justified by the increase in accuracy. Standard computerized techniques for improving the tradeoff between accuracy and speed for Monte Carlo can be employed (e.g., stratified sampling, low-discrepancies sequences, importance sampling, Cholesky decomposition — see Hull, 16.7).

The advantages of Monte Carlo simulation over historical simulation, which lead to the inference that it will be more accurate for the same amount of computing power employed are:

1. Ability to select the most suitable technique to estimate each parameter. Volatilities and correlations can be forecast using statistical techniques such as GARCH. Where implied volatilities are available they can be substituted for or blended with statistical measures. The choice can be separately made for each variable, though you do need to be careful not to generate impossible or implausible combinations of correlation coefficients.

2. Ability to select the most relevant data set for estimating each parameter. You might have 10 years of good historical data for one variable and only 2 years for another. Historical simulation would force you to use only 2 years worth of data for both. Monte Carlo simulation lets you choose the data set individually for each variable and can also choose the weighting of data individually.

3. Greater flexibility in handling missing data. Data for individual dates can be missing because a particular market was closed for a holiday or because of errors in data gathering. In fact, all sources of market data, whether data vendors, brokers, or data bases internal to the firm, are notoriously poor in quality and require major data scrubbing efforts. But some data will not have sufficient duplication of sources to scrub successfully and must be regarded as unavailable. Monte Carlo simulation can exclude periods for which a particular data series is missing from the calculation of each individual variable without excluding this period from the calculation of other variables for which the data is available. Historical simulation lacks this flexibility — it must either completely include or completely exclude a particular time period.
Greater flexibility in handling asynchronous data. Correlations observed between variables which are sampled at different times of the day can be highly misleading and lead to significant misstatements of risk. Monte Carlo simulation has the flexibility to measure correlation for each individual pair of variables based on quotations from the best time of day to represent that particular pair, or by basing the correlation on a multi-day time interval which will tend to smooth out asynchronous effects.

Ability to combine histories. Consider a corporate bond held in the firm's portfolio. By historical experience, one knows that some of these bonds may suffer a ratings downgrade and subsequent large fall in price. But it may be that none of the bonds currently held has suffered such a downgrade since the firm avoids holding such bonds. Historical simulation would show no ratings downgrade events for these bonds. But Monte Carlo simulation could be used to combine ratings downgrade possibilities based on the history of a large pool of bonds with specific pricing history of actual bonds held.

Given all these advantages to Monte Carlo simulation in its flexibility to handle data and estimation issues, it is preferable, and sometimes even unavoidable, to still employ some Monte Carlo simulation techniques when you have chosen historical simulation as your primary methodology. Consider two examples:

1. A certain stock held in your portfolio has only recently been issued. To develop a past history for the price of this stock for use in historical simulation, you may represent it by some formula based on a selected stock index. But if you are long this stock and short this index, you would measure your position as having no risk during the period when it is represented by the index. To avoid this you need to introduce a random element into your generation of the stock's back price history, basing the size of the random element on observed changes during the period since the stock began trading. But this is precisely the Monte Carlo approach.

2. If two stocks have begun trading in very tightly related fashion since a merger announcement, you would not want to reflect their previous more volatile arrangement as part of the history which determines VAR. So you must generate the price of one stock as a function of the other. If you are to avoid treating a merger arbitrage position as having zero risk, you must introduce a random element as in the case above.

Similarly, Monte Carlo simulation techniques can be used to fill in missing data in historical simulations.

One issue which is difficult for Monte Carlo simulation to handle is generation of the right degree of clustering of largest changes in variables. Shaw's paper gives some very interesting ideas on how to approach this within a Monte Carlo simulation framework. In particular he recommends an algorithm of Stein's which allows the generation of scenarios which combine a Monte Carlo generation of individual variables with a joint distribution pattern which takes into account not just the correlation coefficient but the actual observed distribution of rank orders (e.g., if the 3rd highest move in variables actually occurred at the same time as the highest move in variable 2, this pattern would tend to be reproduced by the Monte Carlo simulation).

Step 2: Determine All Market Variables

For spot positions, the translation from underlying market variables to the full set of market variables which you want to multiply by the firm's positions is quite direct. Spot positions such as spot FX or the holding of an individual stock or stock index or spot gold or spot oil is just directly multiplied by the generated price change from Step 1.

Issues are less straight-forward for forward positions. If you are currently holding a Treasury bill maturing 1 month from now, you don't want to apply to it the price move you observed for that Treasury bill on a date 6 months ago, since at that point the Treasury bill had 7 months to maturity, and you expect 7 month instruments to demonstrate much larger price changes than 1 month instruments. So you want to utilize yield curve parameters as underlying market variables and then multiply those yield curve parameters by the appropriate value of a basis point measures of forward position. This has the important added advantage of not having to separately price each interest rate instrument but instead working with a summary description of the entire position.

Issues are most complex for option positions (in which we include any non-linear payoff positions). The conceptually simplest and most accurate approach would be to value each individual option separately based on the changes in the underlying market variables of forward price and implied volatility. Even such
a simple approach has complications, since it is necessary to decide which point in the implied volatility surface is the right one to apply. If you are repricing an option with 1 year to expiry, a strike of 125, and current underlying price of 80, which implied volatility shift do you use when sampling from a period 6 months ago when the underlying price was 100. Most practitioners would opt for looking at the shift in options with a 1 year expiry and a strike of 125, since that would give the same "moneyness", i.e., a strike 25% above current spot. But this is clearly open to interpretation and a variety of theories on what drives options pricing (see the article “Regimes of Volatility” by Derman, RISK, April, 1999). Very similar considerations apply to option-adjusted spreads on mortgage and mortgage-backed securities, which should be related to the security which had a comparable relationship to the prevailing new mortgage rate. The reasoning is similar, since option-adjusted spreads represent the market pricing of uncertainty in option exercise by homeowners.

While the simplest approach is the most accurate, it is clearly also the most costly and the heavy expense of doing full individual revaluation of each option position is what was primarily responsible for incorrect claims that simulation methodology for VAR was inherently expensive to perform. In fact, one simulation methodology can achieve better accuracy than variance-covariance at no greater cost by the easy trick of representing option portfolios by summary statistics of deltas, gammas, and vegas and multiplying these by the appropriate price change, half the square of change in price, and change in implied volatility, respectively. So it is a matter of tradeoff in desired accuracy vs. cost to be determined for each options position. There are also intermediate approaches. One which can provide quite accurate approximations is to interpolate results based on a spot-vol matrix representations of the options portfolio. If a reasonably detailed spot-vol matrix is already being calculated as part of the trading desk’s own risk reporting, this is a good way of taking advantage of a large number of full revaluation runs which are already being made (since each bucket of the matrix requires all options in the portfolio to receive a full revaluation) without needless duplication of effort. As we noted in discussion of the spot-vol matrix, it can potentially capture all higher order terms in the Taylor series of both the underlying price and the volatility, as well as cross-terms between them. It will not capture impacts such as non-parallel shifts in volatility surface, so the sensitivities will need to be separately accounted for. Whatever approximations are used should be occasionally tested against a full revaluation by individual option to see if a finer degree of detail is needed. The scenarios involving the very largest shifts should probably always be evaluated by full revaluation by individual option.

Finally, we will note that some of the determinants of exotic derivative prices are not market variables whose price history can be observed and so are not suitable for inclusion in a VAR analysis. Consider an option on a basket of stocks. The impact of changes in the prices of the stocks and in the implied volatilities of each stock in the basket can be computed and included in the VAR. But there will probably be no liquid market quotations for the implied correlations impacting this option. Analysts are occasionally tempted to substitute changes in historical correlation for unobservable changes in implied correlation. I would argue that this is an error. If the basket option has 3 years remaining, you should presumably look at the change from one business day to the next of a change in the 3 year historical correlation. But since these two 3 year periods will share all but one day at the beginning and end in common, the change in correlation must be tiny. We know from experience that implied volatility can change far more rapidly than a similarly computed change in historical volatility, and I do not know of any reason why correlations should behave differently. If, on the other hand, you decided to choose a much shorter period for computing the historical correlation in order to increase the potential size of the change from day to day, how would the choice of period be justified? I believe it is better to acknowledge that such non-market observables cannot be included in VAR analyses and that their risks should be accounted for separately, through reserves and separate allocations against capital.

Step 3: Calculation of the P&L Distribution

Let us contrast two extremes of approach before suggesting a compromise. One extreme is to calculate any desired statistic of the P&L distribution as a population statistic rather than a sample statistic. So if you want to know the 99th percentile of possible P&L losses and you have simulated 300 possible P&L shifts with equal probability weight, just sort the P&L results, pick the 3rd worst, and report that as your 99th percentile. This approach makes no parametric or modeling assumptions and will pick up fatness of tails, but can produce very unstable results due to small sample bias, compounded by a great sensitivity to errors in data (just one bad data point out of 300, and the 99th percentile you are reporting is actually the 99.33rd
percentile). It has the added disadvantage, which can make explanation of results to senior management quite difficult, that apparently negative diversification effects could arise. Consider the following example:

<table>
<thead>
<tr>
<th>Case</th>
<th>Portfolio A</th>
<th>Portfolio B</th>
<th>Combined Portfolio A &amp; B</th>
</tr>
</thead>
<tbody>
<tr>
<td>3rd worst case for A</td>
<td>– 20MM</td>
<td>+ 10MM</td>
<td>– 10MM</td>
</tr>
<tr>
<td>2nd worst case for A</td>
<td>– 25MM</td>
<td>– 17MM</td>
<td>– 42MM</td>
</tr>
<tr>
<td>1st worst case for A</td>
<td>– 30MM</td>
<td>– 10MM</td>
<td>– 40MM</td>
</tr>
<tr>
<td>3rd worst case for B</td>
<td>– 7MM</td>
<td>– 20MM</td>
<td>– 27MM</td>
</tr>
<tr>
<td>2nd worst case for B</td>
<td>– 10MM</td>
<td>– 40MM</td>
<td>– 50MM</td>
</tr>
<tr>
<td>1st worst case for B</td>
<td>+ 5MM</td>
<td>– 60MM</td>
<td>– 55MM</td>
</tr>
<tr>
<td>99th percentile (3rd worst case)</td>
<td>– 20MM</td>
<td>– 20MM</td>
<td>– 42MM</td>
</tr>
</tbody>
</table>

Negative portfolio effects are undesirable both from the standpoint of clarity of exposition, when explaining risk measures to managers, and from the standpoint of control structure — even if all units of the firm are within allocated VAR risk limits, the firm itself may be outside its risk limits. To avoid negative portfolio effects, it is a sufficient condition that you work with a risk measure which is capable as being represented as the worst case among a number of cases of different weighted sums across scenarios with different probability weights being used to construct each case. (This result is part of a broader study of developing "coherent" risk measures, see "Thinking Coherently" by Artzner, Delbaen, Eber, and Heath). A measure based on the worst of 300 cases would meet this criteria. So would a measure of the 99th percentile based on a weighted average of the worst cases, provided that the weights are assigned so that in going from a case to the next-worse case, the weights are non-increasing since this can be represented as the worst case of all such weightings across all cases. So would a measure of the expected loss beyond the 97th percentile which consisted of a straight average of the nine worst cases, since this can be represented as the worst case of all possible equal weightings of nine cases.

The second extreme is to calculate any desired statistic based on parameters computed from the distribution. You could for example, make the assumption that the final P&L is normally distributed, calculate its volatility from the generated P&Ls, and then derive percentiles from the volatility. This approach is obviously very model dependent and can easily miss fat tails effects (and would miss them in the example given). But results are far more stable and less impacted by data error, and negative diversification effects are not possible.

The suggested compromise is to make separate estimates of each desired statistic using some parametric assumptions. For a brief discussion of this issue along with further references, see Dowd, Chapter 4, section 2.4 and Box 4.2 and particularly the accompanying footnotes.

If you want to use simulation results to project possible extreme results (i.e., many standard deviations), then either you must use a Monte Carlo simulation with a truly large number of runs or you must make heavy use of parametric assumptions. For a brief discussion of the use of extreme value theory to generate such results along with further references, see Dowd's appendix to Chapter 6.

As with any model, a VAR model needs to have its predictions tested against real results to see if it is sufficiently accurate. This process is sometimes known as "backtesting", since you are looking back to see how the model would have performed in the recent past. It has been particularly emphasized for VAR models, owing to insistence by regulators that if firms are to be allowed to use internally built models for calculation of regulatory capital, they must be able to demonstrate that the models fit real results. The suggested regulatory backtest is a straightforward comparison between the 99th percentile produced by a VAR model on each day during a specified period (since it is this percentile which determines regulatory capital) and the actual P&L on each day. The model is considered satisfactory (or at least erring acceptably on the side of too much capital) if the number of days on which P&L exceeds the predicted 99th percentile is not statistically significantly greater than 1%. While this approach has the virtue of simplicity, it is
statistically quite a blunt instrument. Much more information can be extracted by comparing VAR projections to actual results at many different percentiles.

A methodological question is whether to backtest against actual reported P&L or against P&L which has been adjusted for components which the VAR cannot reasonably be expected to pick up. Such components are revenue from newly booked transactions, revenue from intra-day or (when running VAR for periods longer than a day) intra-period trading, and gains or losses due to operational error (e.g., trades incorrectly booked). The argument in favor or using unadjusted P&L in the comparison, besides simplicity of computation, is that these are all real components of P&L which can be quite difficult to identify, so it is better to be aware of the extent to which your model is underpredicting actual reported loss events. In favor of making at least the largest adjustments is that without getting the target data to line up with the forecasting process, you are working with a suboptimal diagnostic tool.

**Stress-Testing**

Stress testing involves using economic insight rather than strict reliance on statistics to generate scenarios against which to measure firm risk. From a computational standpoint, it is simply another variant of simulation — it just uses a different method to generate the scenarios of underlying market variables. But after that, the other two steps in simulation analysis, translation to all market variables and calculation of firm P&L can be carried out exactly as per simulation VAR — indeed, the exact same system can be used for both.

The advantage of using stress-testing as a supplement to VAR is that it can pick up possible extreme events which can cause large losses to the firm’s positions which may be missed by a purely statistical approach. The disadvantage is that once we leave the realm of statistics, we must substitute a standard of plausibility for one of probability, and plausibility is a very subjective notion. However subjective, plausibility must still be insisted upon. Without such a standard, stress-testing becomes equivalent to the child's (and childish) game, "who can name the largest number?" No one ever wins, because one can always be added to the last number. And you can always specify a stress-test which is one shade more extreme than the last one specified.

One question to consider is why bother departing from statistics? Couldn't we just rely on extreme value analysis applied to VAR to generate highly unlikely but still plausible scenarios? There is not a consensus in the industry on this question — for one thing, the application of extreme value analysis to VAR is still relatively new, so we don't have a lot of experience with the results yet. I will give my reasons for personally placing my money on investing in more and better stress tests. Consider the Fall 1998 crisis in credit spreads triggered by the collapse of the Russian economy. The following table shows the degree to which credit spreads blew out in both high-yield corporate bonds and government debt of key emerging market countries.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Chase Securities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Yield Index</td>
<td>780 to 1,060</td>
<td>382 to 433</td>
<td>313 to 361</td>
<td>390 to 716</td>
</tr>
<tr>
<td>Mexican Pars</td>
<td>516 to 1,698</td>
<td>324 to 599</td>
<td>538 to 1,187</td>
<td></td>
</tr>
<tr>
<td>Brazil C Bonds</td>
<td>770 to 1,334</td>
<td>375 to 811</td>
<td>632 to 1,165</td>
<td></td>
</tr>
<tr>
<td>Argentine FRBs</td>
<td>823 to 1,836</td>
<td>238 to 637</td>
<td>428 to 1,284</td>
<td></td>
</tr>
</tbody>
</table>

The size of this rise in credit spread could have easily been anticipated by either a statistical extreme value approach or an economically-based stress-test approach, on the basis of prior experience, as also seen in the table. The key question is should one have anticipated that both events would occur simultaneously. There was no prior experience in which both had simultaneously shown such large moves. You can't posit a rule to assume that all prior worst cases occur simultaneously, or you will wind up with scenarios in which 3 year Treasury rates go up the maximum amount you've ever seen while 4-year Treasury rates go down the maximum amount you've ever seen, which defies my concept of plausibility and would lead to the
conclusion that if you have on a position in which you are long 3 year bonds hedged by a short 4-year bond position, you can reduce your risk by taking off the hedge.

I think that only economic reasoning, which can speculate about underlying causes and innovative factors which may not have been present before (in this case the large amount of money being invested by the same proprietary position traders in both high yield and emerging markets), will be able to disentangle the plausible from the implausible.

There are two fundamental type of stress-tests. The easy kind are just complete replays of a previous stressful event, like the 1987 stock market crash. All you need to do is select the proper start and end dates (we'll say more about that in a minute), make sure you've stored or have researched the historical values of the market variables, and do some artful creation of variables which don't have historical values (e.g., there was no significant emerging market in 1987, so you have to create values based on how emerging market debt fared in subsequent large stock market downturns). The hard kind are the hypothetical scenarios which require economic insight. A few rules of thumb regarding the creation of scenarios:

- Given the difficulty of developing hypothetical scenarios, it is unreasonable to think that more than a handful (say between 5 and 20) can be active at any one time. Given all the potential combinations of events in markets, it is important to focus on those possibilities which are most significant to the types of positions your firm generally holds.
- Anchoring the assumptions for the move of a particular variable to the largest previously move observed historically in either that variable or one which has a close economic relationship is a good preventative against playing the "who can name the highest number?" game and overcoming some of the inherent subjectivity.
- The most important choices are always about which variables can plausibly move together, not about the size of moves. History can be some guide, particularly experience in prior large moves — history of statistical correlations is virtually worthless. It is important to consider linkages which are caused by investors as well as linkages caused by economics.
- Large moves in variables are closely associated with market illiquidity. The size of variable moves chosen must correspond to moves that occur from the time a liquidity crisis begins to the time it ends — prices recorded in between these times often have little meaning, since you can't really do any significant size of business at them. This rule should govern how you choose start and end dates for historical scenarios as well as how you choose start and end dates in determining the largest historical move you've seen for a given variable.

One point of contention between traders on one side and risk managers and regulators on the other side is the assumption that no delta rehedging of options positions will take place during the unfolding of a stress scenario (there is a parallel contention about the same assumption when used for the largest moves seen in VAR simulation). Traders rightly point out that they often have firm rules and limits which would require them to perform a delta rehedge when underlying prices move sufficiently. However, the reason risk managers and regulators insist on assuming no rehedging is the fear that leak of market liquidity in a crisis will prevent rehedging from being executed successfully.

**Uses of Overall Measures of Firm Position Risk**

For a detailed discussion of these points, read sections 3.2.1 and 3.2.2 of Wilson's article on "Value at Risk." He emphasizes the use of VAR as a way of making comparisons between the risks taken in different positions and by different businesses. Since VAR is a statistical measure it can be used to develop a relatively objective comparison of different types of risk by looking at specified probability of loss. Since VAR is expressed as a potential dollar loss amount, it can also be compared to similar measures of other types of risk, such as credit risk and operational risk.
Wilson argues that these features which make VAR a measure of comparison across many different types of position, business, and risk make it a good candidate for determining the amount of capital needed to support a firm's trading risk and to use in internal performance measures comparing the risk capital use of a business line with its P&L. He also argues that VAR is not a good control to prevent firms from having embarrassing, catastrophic losses, since these come from a combination of hidden positions, incorrect pricing, and major economic events which will not be picked up by VAR. I would argue that prevention of catastrophic losses by measuring the potential impact of major economic events is a role stress-testing can play, and because of this stress-testing should also play a role in determining the capital assigned to business lines, even though this does introduce a subjective element which will lead to accusations that different businesses are being judged by a different standard. The operational risk of hidden positions and the market risk of incorrect pricing are beyond the scope of stress-testing as well as VAR. To repeat what I said at the beginning of the term, no system for measuring aggregate risk can be any better than the models and systems which feed it, which is why I have placed so much emphasis on measurement of individual risks.

Wilson is also skeptical of the role VAR can play in control of risks through limits on trading desks. I will discuss this in the overall context of limit controls.

Portfolio risk measures, in addition to being valuable as indicators of overall firm risk, are also useful as guides to which product lines, trading desks, and risk components are the largest contributors to risk. There are varying approaches to representing the composition of risk by component:

1. Each component can be represented by the scenario risk measure it would have as a stand-alone portfolio. This is the easiest approach to implement and certainly gives a good indicator of relative risk, but fails to capture any correlation effects with other risk components which contribute to overall firm risk.

2. Each component can be represented by the impact on total firm risk the full elimination of that risk component would have. This captures correlation effects, but may be unrealistic in that full elimination of a business line may not be a feasible alternative.

3. Each component can be represented by its marginal impact on total firm risk. This captures correlation effects and gives a good measure of the immediate impact on firm risk of adding to or offsetting some of a component's risk, but it is very dependent on the current mixture of risk components. A very risky business line may get represented as having a small contribution to risk just because it has low correlation with the current mix of risk for the firm. It may be best to use a stand-alone risk measure in conjunction with a marginal impact measure to make sure that components which can potentially make large contributions to risk receive timely management focus.

The marginal impact measure has a nice side benefit — when you take the weighted sum of marginal impact, weighted by current positions, you get the total risk measure for the firm. This makes the marginal impact a convenient tool for exercises such as allocation to business line of firm capital where you need the sum of the parts to equal the whole. In order to have this property, a risk measure need only satisfy the condition that it scales directly with position size; i.e., a position with the same composition but \( k \) times as large has a risk measure \( k \) times as large as the original position. This homogeneity condition is clearly met by both VAR and stress-testing measures.

To see that the weighted by position sum of marginal impacts equals total risk, first write the total portfolio as \( \Sigma x_i \) where each \( x_i \) is a component of the portfolio, and let \( R \) be the risk measure. By hypothesis, \( R(\Sigma k x_i) = k R(\Sigma x_i) \). Taking the derivative of both sides with respect to \( k \),

\[
\frac{dR(\Sigma k x_i)}{dk} = \sum \frac{\partial R(\Sigma k x_i)}{\partial k x_i} dk x_i = \sum \frac{\partial R(\Sigma x_i)}{\partial x_i} x_i
\]

\[
\frac{dk R(\Sigma x_i)}{dk} = R(\Sigma x_i)
\]
Hence,

\[ R(\Sigma x_i) = \sum \frac{\partial R(\Sigma x_i)}{\partial x_j} x_j \]

and this partial derivative is just the marginal impact risk measure.

**Trading Desk Limits**

There are 3 fundamental types of trading desk limits: (1) stop loss limits which halt a desk's trading when they have exceeded some predetermined level of losses over a specified time period, (2) limits on aggregate risk taking in the form of VAR or stress scenario loss limits, and (3) limits on specific risk positions, such as limits on rate or volatility exposure across the curve or in a given time bracket.

Virtually everyone would agree to the necessity of stop loss limits. Once losses exceed a certain level, a more senior level of trading management needs to be alerted and involved in the decision as to whether to continue with a trading strategy. Persistent loss can be an indication of risks which are not understood or markets which have changed their character and require fresh thinking. As we discussed in the simulation model for spot risk, it is necessary to have planned for an adequate stop loss level in relation to the revenue level targeted.

Stop loss limits cannot be adequate by themselves, since by time a trader reaches a stop loss level he may have already committed the firm to a size of position which it will take some time to work out of. The traditional way firms have dealt with this risk is to place specific limits on the type of trades each desk can do and the size of specific positions which they can put on, with limit sizes chosen in relation to the degree of liquidity in a given market and a particular desk's expertise in dealing with that market.

Aggregate risk limits were originally proposed as supplements to these specific limits for two reasons: (1) fear that aggregate positions could be built up across a number of markets that would be inside each individual limit but would constitute a dangerously large position in total, and (2) recognition that individual limits changes often lagged behind market changes in price volatility which a VAR limit would immediately adjust to. Some trading desks began to see these aggregate risk limits as giving them a new flexibility — they could expand risk in one area as long as they decreased it in other areas and kept within their aggregate risk limit. They then argued for the elimination of more specific limits to give them room to take advantage of this flexibility.

Generally, risk managers believe that specific risk limits are needed in addition to aggregate limits for 3 reasons: (1) an aversion to putting too much faith in historical experience, (2) the need to restrict specific position taking based on lack of market liquidity and lack of trader experience — a trader should not be given the same degree of freedom in every possible market, and (3) the need for senior trading managers to create risk diversification effects in order to stay in the efficient frontier of risk/return tradeoff — this requires that not all desks be free to pursue the same positions or the same trading styles; individual limits can be used to impose diversity of risk.