Towards a Coherent and Self-Consistent Stress Testing Framework

Riccardo Rebonato

Global Head of Market Risk and of Quantitative Analytics, RBS

Visiting Lecturer, Mathematical Finance, OCIAM, Oxford University

Adjunct Professor, Tanaka Business School, Imperial College, London

Designing Coherent Aggregate Scenarios

This talk outlines a simple idea to aggregate in a coherent fashion the potential stress losses arising from a given set of trading-book positions.

- In the wake of recent events, both the industry and the regulators have keenly felt the need to complement traditional frequentist, percentile-based risk management tools – such as Value at Risk or Economic Capital – with subjective stress tests and scenario analyses.

- **One of the main unresolved questions is how to aggregate the results of these scenarios and stress tests**. This is of particular relevance if a capital charge is to be levied against the combined stress losses, as the recently introduced **Incremental Risk Charge (IRC)** requires.

- In the wake of recent market events the urgency of this capital charge has become evident, and the financial industry has been invited to submit proposals to address the problem.

- The main problem arises from the fact that simply adding the hypothetical losses from all the scenarios would result in a number that is of little meaning, and would certainly be prohibitively and unrealistically large.

- On the other hand, the regulators are correctly and understandably opposed to the endorsement of rules of thumb such as ‘square root of sum of squares’.

- The onus has been placed on the financial industry to come up with a realistic way to combine and aggregate the stand-alone stress losses in a coherent and justifiable manner.

- The magnitudes of the stand-alone losses associated with the individual stress events can either be estimated using traditional frequentist approaches, or it could be based on expert judgement and/or subjective probabilities.

- Given the regulators' examples of 'events' (eg, the depegging of a currency that has never depegged before), **we strongly prefer a subjective-probability, non-frequentist framework**. The risk manger, in our preferred approach, would start from an understanding of the portfolio under his watch, and determine severe but plausible moves of risk factors that would put under stress the largest positions in the portfolio.

- However, we stress that the approach we propose works just as well whichever meaning (frequentist or subjective) is given to the stand-alone (marginal) probability of the individual stress events.

- If the risk manager also believed that the data at his disposal were of sufficient quality and relevance to allow the estimation of the joint co-dependence of the stand-alone stress events (perhaps using a copula approach), then the procedure presented in this paper would be of no use.

- Our proposals are of interest in those situations where the available data are deemed to be either **insufficient**, or not sufficiently relevant to the problem at hand (eg, '**backward-**' rather than '**forward-looking**') or where the risk manager wants to **stress the 'objectively-determined' co-dependences**.

- In this latter sense the present paper can be seen as a generalization of the work by Rebonato and Jaeckel (1999). It is a generalization because we deal with the much richer concept of causation rather than correlation (see, eg, Williamson (2005)).

- The method we present requires the risk manager to provide **conditional probabilities** of rare events. While this is clearly not an easy task, it can be simpler than specifying the stand-alone probabilities of the individual rare events.

- Trivially, while it may be extremely difficult to estimate the probability of a person being involved in a car accident today, or of the same person being taken to hospital, it is relatively easy to give a plausible probability of the person being taken to hospital, conditional on a car accident having occurred.

- In practice, very often the risk manager will have at his disposal heuristics or imperfect models of the workings of sections of the financial markets in certain market conditions (eg, stylized facts such as: 'conditional on an equity market crash, volatilities are likely to increase and credit spreads to widen'). These partial models and rules of thumb fall well short of a coherent model of the market in its entirety.

- However, as former Chairman Greenspan recently pointed out in a related context, disregarding these heuristics because of their incompleteness would be a great mistake:

> "Policymakers often have to act [...] even though [they] may not fully understand the full range of possible outcomes, let alone each possible outcome's likelihood. As a result, [...] policymakers have needed to reach to broader, though less mathematically precise, hypotheses about how the world works..." (quoted in Frydman and Goldberg, 2007).

- Finally, we point out that the aggregate loss associated with these potential stress events can be directly linked to the event-risk charge (see, eg, BIS (2008a), (2008b)) that regulators have recently discussed. The question addressed in this section is therefore how to aggregate these losses and how to take into account possible offsetting gains.

1 Recent Developments

The Discussion Paper *Modelling Extreme Events** refers extensively to my work Rebonato *Plight of the Fortune Tellers*, PUP, (2007):

> 1.2.1 Any statistical estimate is a combination of data analysis, **prior beliefs** (see Rebonato (2007)) and an estimate of the uncertainty in the estimate
>
> [...]
>
> 1.2.2 Although in theory statistical estimates can be made without prior beliefs (a pure 'frequentist' approach), in practice there is rarely

**Modelling Extreme Market Events* – A Report of the Benchmarking Stochastic Moldes Working Party, Discussion Paper presented to the Institute of Actuaries, 3rd November 2008, and to the Faculty of Actuaries, 19th January 2009 sufficient data for the type of market investigation we are concerned with. [...] In the context of the FSA's ICAS regime, [...] the *role of prior beliefs is significant*.

[...]

Where relevant data is [*sic*] plentiful, the choice of prior belief is relatively unimportant. The estimation of rare percentiles, on the other hand, relies on only a handful of data points. As a result, **the choice of prior distribution is critical**, and, consequently, estimates vary wildly between [*sic*] market participants. Any estimate is **sub-tantially a reflection of prior views**, rather than objective data. [my emphasis]

- I could have written all of this myself. Up to this point I could not agree more strongly.

- From this premises, however, the discussion paper goes on to claim that

  there are overriding advantages in ‘standardization’ (comparability among firms and avoidance of the temptation to seek competitive advantage "by adopting weaker assumptions" –

as if different ‘objectiive’, data-driven models did not give a wide diversity of answers, minus the accountability from stating clearly the prior beliefs!)

2 Description of the Methodology

- Suppose that we have n positions, π_1, π_2, ..., π_n, which have been identified as ***vulnerable to losses*** by the desk risk managers. Let E_1, E_2, ..., E_n , $i = 1, 2, ..., 2n$, denote the **significant events that would give rise to large profits or losses arising from our positions** $\{\pi_i\}$ at a given point in time. Let P_i and L_i , $i = 1, 2, ...2n$, be the largest **plausible** profits and losses, respectively, associated with position π_i if event E_i occurs.

- So, to fix ideas, position 1, π_1, could be a large long equity position in the S&P. There are two events associated with position 1, E_1 and E_2: E_1 could be a 1987-like equity market crash; E_2 could be a large market rally. **The rise and fall need not be of the same magnitude**.

- Associated with event E_1 there is a profit, P_1 and a loss L_1. Both profit P_1 and loss L_1 are a function of the event E_1 (the equity market crash) and of the position π_1: $P_1 = P_1(E_1, \pi_1)$, $L_1 = L_1(E_1, \pi_1)$. Similarly, there are a profit, P_2 and a loss L_2 associated with the equity rally.

- It must therefore be stressed that these events do **not encompass only the large losses but also the large possible gains arising from the position $\{\pi_i\}$ which are most vulnerable to losses**. Of course, for a given position, the expected profit or loss need not be just the same number with the opposite sign. This can be either because the **moves in the underlying are not symmetric** (equity markets 'crash' in different ways than they spike, volatilities do not fall in the same way as they suddenly rise, etc); or **because the underlying position may be not linear** (eg, a short-gamma position).

- The notation used in the following assumes that, say, **a yield curve steepening and a yield curve flattening are two distinct events**. If our portfolio contains, say, a steepener position, then the ***loss*** for the steepener scenario would be zero, and the ***gain*** for flattening scenario would be zero. In short, given a large loss-vulnerable position, the associated events will, in general, be the largest plausible moves that give rise *to profit or losses*.

- We assume that the risk manager *can* express an informed opinion about the probability, $p_{j|i}$, of occurrence of event E_j conditional on event E_i having occurred. For instance, if event E_i were a major equity market crash, the event E_j associated with a *drop* in equity volatilities would have close-to-zero probability; however, the event E_k associated with widening in credit spreads could have a significant probability. In the following, **to avoid spurious precision**, these conditional probabilities can initially be grouped into buckets, say, 0.10, 0.3, 0.5, 0.7, 0.9.

- Given n large loss-vulnerable positions, there will therefore be in general n events, with a profit, P_i, and a loss, L_i, associated with each.

- Consider now the [2n x 1] matrix $\mathbf{y}$, defined by

$$\mathbf{y} = \mathbf{p} \cdot \mathbf{E} =$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ \cdots \\ \\ y_{2n-1} \\ y_{2n} \end{bmatrix} = \begin{bmatrix} 0 & 0 & p_{2|1} & p_{2|1} & \cdots & p_{2n|1} & p_{2n|1} \\ p_{1|2} & p_{1|2} & 0 & 0 & \cdots & p_{2n|2} & p_{2n|2} \\ \\ \\ \\ \\ & & & & & 0 & 0 \end{bmatrix} \begin{bmatrix} L_1 \\ P_1 \\ L_2 \\ P_2 \\ \\ L_n \\ P_n \end{bmatrix}$$

- Consider, say, the second entry of the matrix $\mathbf{y}$. Given the definition above it is given by

$$y_2 = (L_1 + P_1)\, p_{1|2} + 0 + (L_3 + P_3)\, p_{3|2} + \ldots (L_n + P_n)\, p_{2n|2} \quad (1)$$

- This can be interpreted as the (conditional) expectation of the **profits and losses incurred by our portfolio** *due to positions other than position 2* **if the second scenario event materializes**. The total stress loss, SL_2, if the second *event* scenario materializes is therefore given by the sum of y_2 and the *loss* associated with the second scenario:

$$SL_2 = L_2 + y_2 \quad (2)$$

- As Equation (1) shows, the matrix equation above can therefore be written more concisely as

$$\mathbf{y} = \mathbf{p} \cdot \mathbf{E} =$$

$$
\begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ \cdots \\ \\ y_{n-1} \\ y_n \end{bmatrix}
=
\begin{bmatrix}
0 & p_{2|1} & p_{3|1} & \cdots & \cdots & p_{n-1|1} & p_{n|1} \\
p_{1|2} & 0 & p_{3|2} & \cdots & \cdots & p_{n-1|2} & p_{n|2} \\
 & & 0 & & & & \\
 & & & 0 & & & \\
 & & & & 0 & & \\
p_{1|n-1} & & & & & 0 & p_{n|n-1} \\
p_{1|n} & p_{2|n} & & & & p_{n-1|n} & 0
\end{bmatrix}
\begin{bmatrix} L_1 + P_1 \\ L_2 + P_2 \\ \\ \\ \\ \\ L_n + P_n \end{bmatrix}
$$

3 A worked out example

- A simple example can clarify the notation and the ideas behind the approach.

- Let event 1 be an **equity market crash** (S&P). (Event 2 can be an equity market rally, but we do not consider this in this example.)

- Let event 3 be a **flattening of the US$ yield curve** and

- Let event 4 be a **steepening of the US$ yield curve**.

- Let's assume that **we are long the S&P and we have a yield curve steepener on**. Given an equity market crash we have a conditional probability of a curve steepening of the US$ curve, $p_{4|1}$, of 80%, and a conditional probability, $p_{3|1}$, of a flattening of 20% (the two probabilities only happen to add up to 1 by chance.)

Since we have a long equity position and a yield curve steepener we then have

- a gain, P_4, and a loss, L_4, of, say, \$100m and \$0m, respectively, associated with event 4 (the steepening of the curve) – L_4 is 0 because there is no loss from our position (a steepener) if event 4 happens, ie, if the curve indeed steepens;

- a gain, P_3, and a loss, L_3, of, say, \$0m and \$100m, respectively, associated with event 3 (the flattening of the curve) – it is P_3 that is now 0 because there is no gain from our position (a steepener) if event 3 happens, ie, if the curve flattens;

- a stand-alone loss, L_1, of, say, \$200m on the S&P position if the equity market crash occurs.

- **The conditional losses associated with event 1** (equity market crash) are given by:

$$y_1 = p_{3|1}(L_3) + p_{3|1}(P_3) + p_{4|1}(L_4) + p_{4|1}(P_4) =$$

$$p_{3|1}(L_3) + p_{3|1}(0) + p_{4|1}(0) + p_{4|1}(P_4)$$

- **The stress loss 1 associated with the equity market crash is then given by**:

$$SL_1 = L_1 + (P_3 + L_3)\,p_{3|1} + (P_4 + L_4)\,p_{4|1} =$$

$$L_1 + (0 + L_3)\,p_{3|1} + (P_4 + 0)\,p_{4|1} =$$

$$-\$200m + \$100m * 0.8 - \$100m * 0.2 = -\$140m$$

- It must be stressed again that the profits and losses L_2, P_2 and L_3 and P_3 do not depend on event 1 (the equity market crash), and are unconditional (marginal) quantities. The link with the equity market crash only comes via the conditional probabilities, $p_{2|1}$ and $p_{3|1}$.

- As our intuition suggests, in the example above **the steepening position mitigates the stand-alone loss** in the naked equity position, reducing it from \$200m to \$140m.

4 Some Observations

The following obvious points are worth noting:

- In general

$$\sum_{j=1,n} p_{i|j} \neq 1 \tag{3}$$

- In general

$$p_{i|j} \neq p_{j|i} \tag{4}$$

  ie, the matrix $\mathbf{p}$ is not symmetric: the probability of, say, equity volatility increasing given a market crash is not the same as the probability of a market crash given a an increase in equity volatility.

- In the worked-out example above, a (linear) position in a steepener gave a profit and zero loss if the curve did steepen, (and vice versa for the flattener.) This need not be case: if we had a short gamma position on the level of a yield curve, for instance, we could record a loss both on yield curve moving up and down.

5 The Link To the Event Risk Charge

On the basis of this interpretation it is reasonable to equate the Event Risk Charge, ERC, to

$$\text{ERC} = \max\{SL_i\} \tag{5}$$

This measure has some nice properties.

- **If we add a new scenario, its marginal contribution on ERC is not purely additive**, reflecting the intuition that all the **stress losses will not occur at the same time**.

- If we add a new scenario, say, the $n+1$th scenario, it can produce the largest entry in the ERC calcualtion even if the stand-alone loss L_{n+1} is not the largest, provided that the conditional losses for the other events ($i \neq n+1$) add up to the largest SL. **This loss diversifes poorly**.

- If we add a new scenario, **its marginal contribution on ERC can be negative**, provided that the marginal conditional profits for events $i \neq n+1$ reduce the largest SL.

- As one position becomes larger and larger, it will tend to dominate the ERC number, both via its stand-alone loss, and via the conditional profits and losses. In other words, **the measure ERC knows both about concentration and diversification**.

6 Internal Consistency of the Conditional Probabilities

- In the presentation above the conditional probabilities are supposed to be exogenously provided by the risk managers. However, as these conditional probabilities ultimately reflect dependence among events (and actually can be linked to causal relationships in a way that the correlation coefficient cannot), it is not surprising that they cannot be arbitrarily assigned.

- When correlation matrices are exogenously assigned, the consistency requirements are that the real symmetric matrix should be positive (semi)-definite. **We derive similar conditions for the matrix of conditional probabilities**.

- We note that an arbitrarily-chosen set of conditional probabilities are extremely unlikely to be consistent, the more so when some conditional probabilities are ‘large’ and some ’small’ – as will typically be the case with the conditional probability matrix of interest in our application.

- We stress that **the constraints are actually of great help in guiding the intuition of the risk manager when, as we advocate, the conditional probabilities are subjective**. As we shall see, the **space of admissible solutions can be remarkably narrow**. This implies that, given one broad type of assumed joint co-dependences, the possible solutions cannot differ too much and remain admissible.

- The degree of uncertainty and arbitrariness in assigning the conditional probability matrix is therefore greatly reduced. "Islands of aceptability."

6.1 Preliminary Analysis and Notation

- We present a **Linear Programming approach** to find an admissible solution 'close', in some sense to be discussed, to a subjective input conditional probability matrix.

- It is useful to apply some 'pre-processing' of this input matrix in order to ensure that the risk manager's intuition is fully reflected in the matrix that is input into the more opaque Linear Programming algorithm. This 'pre-processing' takes the form of a few preliminary tests that should be run to eliminate obvious inconsistencies. **Highlighting the inconsistencies at an early stage forces the risk manager to re-analyze his input matrix**.

- We have events x_i for $i = 1, 2, ..., n$. In this section and for the rest of the paper, in order to lighten notation we will use (i) to represent the (unknown) unconditional probability of x_i and $[i|j]$ to represent the conditional probability x_i of given x_j. The following identity is used extensively in the following:

$$\frac{(i)}{(j)} = \frac{[i|j]}{[j|i]} \tag{6}$$

- Obviously, in specifying $[i|j]$ and $[i|j]$ the risk manager is already making a statement about the relative likelihood of (i) and (j) – **a useful first 'sanity check' can be carried out here** – for instance, both i and j should be interpreted as 'rare' events of not-too-dissimilar likelihhood. If this is the case, the ratio should be close to 1. A ratio of 2, or even 10 is still acceptable. But a ratio of, say, 1,000 would suggest that we re mixing 'apples and oranges' – **events of very different (subjective) likelihood of occurrence**. This can be justifiable, but must be questioned.

- Our first constraint (**Constraint Triplets**) refers to triplets, and requires that

$$[i|j] = [j|i]\,\frac{(i)}{(j)} = [j|i]\,\frac{(i)}{(k)}\frac{(k)}{(j)} \Longrightarrow 1 \geq [j|i]\,\frac{[i|k]}{[k|i]}\frac{[k|j]}{[j|k]} \geq 0 \text{ for } i \neq j \neq k \tag{7}$$

- We have 6 combinations of $[i|j]$ for $i \neq j$, **Constraint Triplets** shows that, given the conditional probabilities of 5 combinations, the 6^{th} ($[i|j]$) is uniquely determined (and, of course, must be smaller than or equal to 1 and non-negative).

- A simple routine can easily be written that returns all the inconsistent triplets, **and allows the risk manger to query the inputs and make changes accordingly**.

- We note in passing that setting $[i|j] = [j|i]$ ensures that **Constraint Triplets** is automatically satisfied for any triplet, but this is in general not a good place to start – the probability of a general equity market crash given a widening in the credit spread of firm X can be very different from the probability of widening in the credit spread of firm X given a general equity market crash.

- **If two scenario are mutually incompatible**, there are severe constraints on the other conditional probabilities. We make use of the identity

$$[i \cup j|k] = [i|k] + [j|k] - [i \cap j|k] \tag{8}$$

(where $\cup$ and $\cap$ indicate union and intersection, respectively). If $[i\,|j|] = [j|i] = 0$ for all other scenarios k it must hold that

$$[i \cap j|k] = 0 \tag{9}$$

- Therefore (**Constraint Zeros**)

$$[i|k] + [j|k] = [i \cup j|k] \leq 1 \tag{10}$$

- A simple routine can be written to identify violations of this inequality for any zero conditional probability.

6.2 Systematic Solution

- The above analyses can be used to remove the more obvious inconsistencies, but there are still likely to be subtler inconsistencies.

- The general methodology consists in checking the proposed conditional probabilities for consistency, and until consistency is proven **we allow the conditional probabilities to be moved within a widening range of values**.

- **At each iteration the algorithm suggests a consistent solution**, and at any time **this solution can be accepted (even if it is outside the current range)**.

- The formulation is as follows.

- With each of the scenarios we associate an **indicator variable**, $I_i, i = 1, 2, ..., N$, where $I_i = 1$ if scenario i occurs, and $I_i = 0$ otherwise.

- There are 2^N mutually **exclusive and exhaustive joint events** corresponding to the distinct combinations of the set of I_is.

- We signify **any given combination of I_is by the vector $\boldsymbol{I}$**, and **the set of all $\boldsymbol{I}$s by $\mathcal{I}$**.

- To each vector $\boldsymbol{I}$ we associate a probability $P[\boldsymbol{I}]$.

- A conditional probability $[i|j]$ is linked to the vector probabilities $P[\boldsymbol{I}]$ by

$$[i|j] = \frac{\sum_{\boldsymbol{I}:\{I_i=1, I_j=1\}} P[\boldsymbol{I}]}{\sum_{\boldsymbol{I}:\{I_j=1\}} P[\boldsymbol{I}]} \tag{11}$$

- In order for any set of proposed conditional probabilities $\{[i|j]\}$ to be consistent there must exist (at least) one set of $P[\boldsymbol{I}]$ such that

$$[i|j] = \frac{\sum_{\boldsymbol{I}:\{I_i=1,I_j=1\}} P[\boldsymbol{I}]}{\sum_{\boldsymbol{I}:\{I_j=1\}} P[\boldsymbol{I}]} \tag{12}$$

$$\sum_{\boldsymbol{I}\in\mathcal{I}} P[\boldsymbol{I}] = 1 \tag{13}$$

$$P[\boldsymbol{I}] \geq 0 \text{ for any } \boldsymbol{I} \in \mathcal{I} \tag{14}$$

(note that the sum is over all the possible vectors $\boldsymbol{I}$s **in the set** $\mathcal{I}$ – this includes the vector $\mathbf{0}$, with associated $P[\mathbf{0}]$, ie, **the probability than none of the specified events will happen**. Note that this probability should be 'rather high'.

A short detour about event correlation

- Risk managers tend to think mor readily in terms of correlation rather than conditional probabilities

- Similarly, the copula models popular for pricing credit derivatives are expressed in terms of default correlation

- I briefly explore the links between the two concepts.

- From Equation (12) we can easily work out the correlation between Event i and j.

- A natural definition for event correlation is given by

$$\rho_{ij} = \frac{E\left[(I_i - E[I_i])\left(I_j - E\left[I_j\right]\right)\right]}{\sqrt{var(I_i)\,var\left(I_j\right)}} \tag{15}$$

where the expectation is taken over the 2^N combinations of scenarios.

- Now,

$$E[I_i] = \sum_{\boldsymbol{I}:\{I_i=1\}} P[\boldsymbol{I}] = \sum_{k}^{1,\ldots,2^N} I_i^k P[\boldsymbol{I}_k] = p(i) \tag{16}$$

where the sum is over the 2^N possible vectors (scenario combinations) $\boldsymbol{I}_k$.

p1)=	0.455
p(2)=	0.640
p(3)=	0.295

E(3)^2	E(2)^2	E(2)E(3)	E(1)	E(2)	E(3)		P[E(i)]
1	1	**1**	1	1	1	**E(1)**	0.10
0	1	**0**	1	1	0	**E(2)**	0.30
1	0	**0**	1	0	1	**E(3)**	0.02
0	0	**0**	1	0	0	**E(4)**	0.04
1	1	**1**	0	1	1	**E(5)**	0.02
0	1	**0**	0	1	0	**E(6)**	0.22
1	0	**0**	0	0	1	**E(7)**	0.16
0	0	**0**	0	0	0	**E(8)**	0.14

Exp[E(2)E(3)]=	0.120
p(2)=	0.640
p(3)=	0.295
Exp[E(2)^2]=	0.640
Exp[E(3)^2]=	0.295
Var[E(2)]	0.2304
Var[E(3)]	0.207975

p(2|3)=	**Numerator**	0.12	**0.40678**
	Denominator	0.30	
p(3|2)=	**Numerator**	0.12	**0.1875**
	Denominator	0.64	
Event Correlation	**Numerator**	-0.0688	**-31.43%**
	Denominator	0.2189005	

Check	***Check***
2.169492	2.169492

Figure 1:

- Remember: $\boxed{P[\boldsymbol{I}_k]}$ is the probability of the kth **scenario combination**, and $\boxed{p(i)}$ is the probability of the ith **scenario** – the value of the indicator of the ith scenario averaged over the scenario combinations.

- As for the denominator,

$$var\left(I_i\right) = E\left[I_i^2\right] - \left(E\left[I_i\right]\right)^2 \tag{17}$$

- But (see the table above)

$$E\left[I_i^2\right] = \sum_{k}^{1,\dots,2^N} \left(I_i^k\right)^2 P[\boldsymbol{I}_k] = \sum_{k}^{1,\dots,2^N} I_i^k P[\boldsymbol{I}_k] = E\left[I_i\right] = p(i) \tag{18}$$

- Therefore

$$var\left(I_i\right) = p(i) - p(i)^2 \tag{19}$$

(which, by the way, nicely ensures that in all non-degenerate situation the variance will be strictly positive.)

- Therefore Equation (15) can be re-written as:

$$\rho_{ij} = \frac{E\left[\left(I_i - p(i)\right)\left(I_j - p(j)\right)\right]}{\sqrt{\left(p(i) - p(i)^2\right)\left(p(j) - p(j)^2\right)}} \tag{20}$$

- In the discrete case above, this is equal to

$$\rho_{ij} = \frac{E\left[\left(I_i - p(i)\right)\left(I_j - p(j)\right)\right]}{\sqrt{\left(p(i) - p(i)^2\right)\left(p(j) - p(j)^2\right)}} =$$

$$\frac{\sum_{k}^{1,\ldots,2^N} \left(I_i^k - p(i)\right)\left(I_j^k - p(j)\right) P[\boldsymbol{I}_k]}{\sqrt{\left(p(i) - p(i)^2\right)\left(p(j) - p(j)^2\right)}} \tag{21}$$

- How can we interpret this expression?

- The the term $\left(I_i^k - p(i)\right)$ in the **numerator gives the deviation in the kth scenario combination of value of the indicator for event i from its average value** (which is just $p(i)$); ditto for event j. The **product** $\left(I_i^k - p(i)\right)\left(I_j^k - p(j)\right)$ therefore **gives an indication of how the deviations (positive or negative) of the values of events i and j are related in the kth scenario combination** (positive deviation times positive deviation, positive deviation times negative deviation, etc).

- All these products of deviations are then weighted by the probability of occurrence of scenario combination k, and normalized by the variances.

- The table above was constructed from the **assumed knowledge** of the scenario probabilities $P[\boldsymbol{I}_k]$.

- Note, however, that **in general we do not have this information**. That is why we did not provide this link between the conditional probabilities and the correlation in the sanity checks presented above. Equation (15) still **provides a useful link between the various building blocks of the problem**.

Let's go back to the optimization problem

- Since any proposed set of conditional probabilities is very unlikely to be internally consistent, we allow the conditional probabilities to lie anywhere between **lower and upper limits**, $\sigma_{i|j}$ and $\vartheta_{i|j}$, respectively, for $[i|j]$. **These upper and lower limits can be specified by the risk manager**.

- Hence we can impose for each i and j, the two linear constraints:

$$\sigma_{i|j} \sum_{\boldsymbol{I}:\{I_j=1\}} P[\boldsymbol{I}] - \sum_{\boldsymbol{I}:\{I_i=1,I_j=1\}} P[\boldsymbol{I}] \leq 0 \tag{22}$$

and

$$\vartheta_{i|j} \sum_{\boldsymbol{I}:\{I_j=1\}} P[\boldsymbol{I}] - \sum_{\boldsymbol{I}:\{I_i=1,I_j=1\}} P[\boldsymbol{I}] \geq 0 \tag{23}$$

- We then define non-negative 'slack' variables, $s_{i|j}$ and $t_{i|j}$:

$$s_{i|j} = -\sigma_{i|j} \sum_{\boldsymbol{I}:\{I_j=1\}} P[\boldsymbol{I}] + \sum_{\boldsymbol{I}:\{I_i=1, I_j=1\}} P[\boldsymbol{I}] \tag{24}$$

$$t_{i|j} = \vartheta_{i|j} \sum_{\boldsymbol{I}:\{I_j=1\}} P[\boldsymbol{I}] - \sum_{\boldsymbol{I}:\{I_i=1, I_j=1\}} P[\boldsymbol{I}] \tag{25}$$

where

$$s_{i|j} \geq 0 \quad \text{for any } i, j \tag{26}$$

$$t_{i|j} \geq 0 \quad \text{for any } i, j \tag{27}$$

- We refer to any set of probabilities that obey Equations (13), (14) a '**feasible**' solution, and any set that in addition obey Equations (22) and (23) a '**coherent**' solution. **If we can find a coherent solution then the given constraints are consistent, otherwise they are not**.

- In order to find a coherent solution we use Phase 1 of the Revised Simplex Method (see, eg, Press et al (1996)). This involves **defining non-negative artificial variables** to give an initial 'basic' solution[†], and **minimize the sum of the artificial variables, say, z, subject to the given constraints**.

- **If the minimum value of z, z_{min} say, is greater than zero, the constraints do not allow a coherent solution (ie, a solution consistent with the lower and upper limits assigned by the risk manager)**.

- However, in this case our optimal solution will be, in some sense, as 'close' as we can get to a coherent solution. The risk manager can then either

[†]A 'basic' solution is one in which all but M of the varibales are set to zero (where M is the number of constraints – excluding the non-negativity constraints). At each iteration each of the basic variables and the objective function (z in our case) are expressed as affine functions of the non-basic variables. See the worked-out example below.

accept the resulting 'close' solution or **widen the limits,** $\sigma_{i|j}$ **and** $\vartheta_{i|j}$**, within which the conditional probabilities are allowed to lie**.

- It is worth noting that any feasible solution gives rise to an infinite number of related solutions. Representing the given solution by $P_{opt}[\boldsymbol{I}]$, any solution such that:

$$P[\mathbf{0}] = \frac{1+\alpha}{1+\alpha P_{opt}[\mathbf{0}]} P_{opt}[\mathbf{0}] \tag{28}$$

and

$$P[\boldsymbol{I}] = \frac{1}{1+\alpha P_{opt}[\mathbf{0}]} P_{opt}[\boldsymbol{I}] \quad \text{for} \quad \boldsymbol{I} \neq \mathbf{0} \tag{29}$$

is also feasible for $-1 \leq \alpha$.

- In the expressions above, $\mathbf{0}$ is the vector $\boldsymbol{I}$ whose components are all zero, and $P[\mathbf{0}]$ is the associated probability that no scenario occurs.

- Scaling all the individual probabilities (except $P[\mathbf{0}]$) by the same proportion clearly does not affect the conditional probabilities. This means that the solution can be associated with unconditional probabilities that are arbitrarily small or large (provided, of course, that none of them is greater than 1).

- Note also that, given any proposed solution with $P[\boldsymbol{I}] \geq 0$ for any $\boldsymbol{I} \in \mathcal{I}$, we can always find a *feasible* solution (i.e. the constraint $\sum_{\boldsymbol{I} \in \mathcal{I}} P[\boldsymbol{I}] = 1$ can always be satisfied) simply by dividing each $P[\boldsymbol{I}]$ by the sum of the $P[\boldsymbol{I}]$s, and this will not affect the conditional probabilities.

7 Implementation

- We prototyped the simplex solution in VBA; for 12 scenarios this was taking of the order of 5-10 minutes to run. Coding it in C++ makes the routines runs 10 to 20 times faster.

- Preliminary analysis suggests that each additional scenario will increase the run time by a factor of about three; considering that the analysis is likely to be required infrequently (say, on a weekly basis), **the computation time should not be a problem with up to, say, 15-20 scenarios**. In fact we are more likely to run into memory-space problems first

- The ability of filling the matrix in a meaningful and thoughtful manner is more constraining.

8 A Worked-Out Example

- In order to illustrate the procedure in practice we present in this section the results of a fictitious case study.

- We have four scenarios, A, B, C, and D, and we suppose that the risk-manager supplied the following conditional probabilities:

$$\begin{bmatrix} & \textbf{A} & \textbf{B} & \textbf{C} & \textbf{D} \\ \textbf{A} & 1 & 0.60 & \mathbf{0.50} & \mathbf{0.60} \\ \textbf{B} & 0.80 & 1 & 0.22 & 0.20 \\ \textbf{C} & 0.50 & 0.20 & 1 & \mathbf{0.00} \\ \textbf{D} & 0.40 & 0.40 & \mathbf{0.00} & 1 \end{bmatrix} \quad (30)$$

where, for example, the figure 0.5 in the first row represents $[C|A]$.

SANITY CHECKS

- We first observe that $[C|A]+[D|A] = 1.1$. But, from the matrix, C and D and are mutually exclusive ($[C|D] = [D|C] = 0$). Therefore we need to reduce either or both of $[C|A]$ and $[D|A]$ so that their sum is ≤ 1. In fact we would make the sum **strictly less than 1** to avoid making the overly strong statement that if A occurs then either C or D must occur.

- Suppose that the risk manager therefore sets $[C|A] = 0.4$ and $[D|A] = 0.5$. If we now look at the triplet $\{A, B, D\}$, we should have:

$$[A|D] = [D|A]\,\frac{[A|B]}{[B|A]}\frac{[B|D]}{[D|B]} = 0.5\frac{0.80}{0.60}\frac{0.40}{0.20} = 1.333 \qquad (31)$$

So the 5 probabilities on the right hand side are clearly inconsistent.

- To fix the problem, the risk manager could set $[D|B] = 0.3$, for example, which gives $[A|D] = 0.89$. Our revised matrix therefore becomes:

$$\begin{bmatrix} & \textbf{A} & \textbf{B} & \textbf{C} & \textbf{D} \\ \textbf{A} & 1 & 0.60 & 0.40 & 0.50 \\ \textbf{B} & 0.80 & 1 & 0.22 & \mathbf{0.30} \\ \textbf{C} & 0.50 & 0.20 & 1 & 0.00 \\ \textbf{D} & 0.89 & 0.40 & 0.00 & 1 \end{bmatrix} \tag{32}$$

  which shows no obvious inconsistencies.

- Note, however, that, to 'fix the problem', **the risk manager has had to change substantially his original estimate of** $[A|D]$ (from 0.40 to 0.89). **The plausibility of this change should be questioned**.

  1) Was the original suggestion ill-thought-out?

2) Or is the ‘fix’ forcing assumptions on the risk manager she is not comfortable with?

- **We view this as an important and useful part of the process (the ‘auditable decision process’)**.

- We now move to the next phase, to ensure consistency of the plausible and ‘cleansed’ matrix. Suppose that the risk manager allows each of the conditional probabilities[‡] to lie within a range

$$[i|j]_1\,(1-\delta) \leq [i|j]_2 \leq [i|j]_1 + \delta\,(1-[i|j]_1) \tag{33}$$

where the $[i|j]_1$s are taken from the revised matrix above (the subscripts 1 and 2 indicate the first or second iteration of the procedure, resepctively)

[‡]Except, of course, those that are 1 or 0.

- Note that we are not necessarily dealing yet with *bona fide*, internally consistent conditional probabilities.)

- Taking $\delta = 0.01$ this gives lower and upper bounds for the conditional probabilities:

$$\left[\sigma_{i|j}\right] = \begin{bmatrix} & \textbf{A} & \textbf{B} & \textbf{C} & \textbf{D} \\ \textbf{A} & 1 & 0.59 & 0.40 & \mathbf{0.50} \\ \textbf{B} & 0.79 & 1 & 0.22 & 0.30 \\ \textbf{C} & 0.50 & 0.20 & 1 & 0.00 \\ \textbf{D} & 0.88 & 0.40 & 0.00 & 1 \end{bmatrix} \tag{34}$$

and

$$\left[\vartheta_{i|j}\right] = \begin{bmatrix} & \textbf{A} & \textbf{B} & \textbf{C} & \textbf{D} \\ \textbf{A} & 1 & 0.60 & 0.41 & 0.51 \\ \textbf{B} & 0.80 & 1 & 0.23 & 0.31 \\ \textbf{C} & 0.51 & 0.20 & 1 & 0.00 \\ \textbf{D} & 0.89 & 0.41 & 0.00 & 1 \end{bmatrix} \tag{35}$$

respectively.

- We now define the 16 indicator vectors and their corresponding probabilities:

$$\begin{bmatrix} \text{Indicator} & \text{A} & \text{B} & \text{C} & \text{D} & \text{Prob} \\ \boldsymbol{I}_0 & 0 & 0 & 0 & 0 & P(0) \\ \boldsymbol{I}_1 & 0 & 0 & 0 & 1 & P(1) \\ \boldsymbol{I}_2 & 0 & 0 & 1 & 0 & P(2) \\ \boldsymbol{I}_3 & 0 & 0 & 1 & 1 & P(3) \\ \boldsymbol{I}_4 & 0 & 1 & 0 & 0 & P(4) \\ \boldsymbol{I}_5 & 0 & 1 & 0 & 1 & P(5) \\ \boldsymbol{I}_6 & 0 & 1 & 1 & 0 & P(6) \\ \boldsymbol{I}_7 & 0 & 1 & 1 & 1 & P(7) \\ \boldsymbol{I}_8 & 1 & 0 & 0 & 0 & P(8) \\ \boldsymbol{I}_9 & 1 & 0 & 0 & 1 & P(9) \\ \boldsymbol{I}_{10} & 1 & 0 & 1 & 0 & P(10) \\ \boldsymbol{I}_{11} & 1 & 0 & 1 & 1 & P(11) \\ \boldsymbol{I}_{12} & 1 & 1 & 0 & 0 & P(12) \\ \boldsymbol{I}_{13} & 1 & 1 & 0 & 1 & P(13) \\ \boldsymbol{I}_{14} & 1 & 1 & 1 & 0 & P(14) \\ \boldsymbol{I}_{15} & 1 & 1 & 1 & 1 & P(15) \end{bmatrix}$$

Recall that indicator variables represent the occurrence of combination of scenarios. Thus, for example, the indicator variable $\boldsymbol{I}_3$ represents the outcome that scenarios C and D occur and scenarios A and B and do *not* occur, and $P(3)$ is the corresponding probability (to be solved-for).

- We have 12 constraints of the form:

$$\sigma_{i|j} \sum_{\boldsymbol{I}:\{I_j=1\}} P[\boldsymbol{I}] - \sum_{\boldsymbol{I}:\{I_i=1,I_j=1\}} P[\boldsymbol{I}] \leq 0 \tag{36}$$

and 12 constraints of the form

$$\vartheta_{i|j} \sum_{\boldsymbol{I}:\{I_j=1\}} P[\boldsymbol{I}] - \sum_{\boldsymbol{I}:\{I_i=1,I_j=1\}} P[\boldsymbol{I}] \geq 0 \tag{37}$$

- For example, corresponding to $[D|A]$ we have (remembering from Equation (34) that $[D|A]$=0.5):

$$0.5\,\{P(8)+P(9)+P(10)+P(11)+P(12)+P(13)+P(14)+P(15)\} - \{P(9)+P(11)+P(13)+P(15)\} \leq 0 \tag{38}$$

ie,

$$0.5\,\{P(8)+P(10)+P(12)+P(14)-P(9)-P(11)-P(13)-P(15)\}$$
$$+s_{D|A}=0 \qquad (39)$$

where $s_{D|A}$ is the 'slack' variable corresponding to this constraint.

- Trivially, all the constraints, as formulated, can be satisfied by setting $P(\mathbf{0}) = 1$ (the probability of no scenarios occurring is one), and all other variables to zero. Clearly, however, for this solution the conditional probabilities are not defined and we do not have an economically interesting solution.

- Therefore we introduce non-negative 'artificial' variables, $a_{X|Y}$, of the kind:

$$a_{D|A} = 0.5\{P(8) + P(10) + P(12) + P(14) - P(9) - P(11) - P(13) - P(15)\} + s_{D|A} \quad (40)$$

and in order to get the algorithm started we initialize all of these to some arbitrarily small value, ε.

- Our initial (infeasible) solution is given by setting all the artificial variables to ε, and all others to zero.

- We immediately have an expression for the sum of the artificial values in terms of the currently zero variables[§], by summing the right-hand sides of all the equations exemplified by Equation (40), and we use the revised simplex algorithm to minimize this expression.

[§]This is what is required for the Revised Simplex Algorithm: at each iteration we have a set of 'basic' (in general non-zero) variables, and 'non-basic' (zero-valued) varaibales. The value of each iof the basic variables and trhe value of the 'objective function' (the expression to be optimized) are expressed as functions of the non-basic varibales only. See, eg, Press et al (1996)

- This gives the following solution:

$$\begin{bmatrix} \text{Indicator} & \textbf{A} & \textbf{B} & \textbf{C} & \textbf{D} & K \\ \boldsymbol{I}_2 & 0 & 0 & 1 & 0 & 1.606 \\ \boldsymbol{I}_1 & 0 & 0 & 0 & 1 & 0.336 \\ \boldsymbol{I}_{14} & 1 & 1 & 1 & 0 & 0.702 \\ \boldsymbol{I}_9 & 1 & 0 & 0 & 1 & 0.865 \\ \boldsymbol{I}_{13} & 1 & 1 & 0 & 1 & 0.946 \\ \boldsymbol{I}_{12} & 1 & 1 & 0 & 0 & 0.537 \\ \boldsymbol{I}_4 & 0 & 1 & 0 & 0 & 0.664 \\ \boldsymbol{I}_{10} & 1 & 0 & 1 & 0 & 0.734 \end{bmatrix} \tag{41}$$

where K is a quantity proportional to the probabilities and all other probabilities, except $P(\mathbf{0})$, have been set to zero.

- The corresponding solution in terms of conditional probabilities is:

Trial Solution	**A**	**B**	**C**	**D**
A	1.00	0.58	0.38	0.48
B	0.77	1.00	0.25	0.33
C	0.47	0.23	1.00	0.00
D	0.84	0.44	0.00	1.00

(42)

- Comparing this with Tables (34) and (**??**) we see that the largest violation of the constraints is for $[A|D]$, where the above solution lies 0.04 below its lower limit.

- We could

  1. accept this solution,

2. make *ad hoc* changes to the limits, or

3. increase δ and re-optimise.

9 Where Do We Get the Stand-Alone Stresses from?

- Starting point are the vulnerbailities of the portfolio

- We need an imperfect model of reality to make sense of the statistical interpretation

- Without such interpretative models statistical analysis is similar to chartism – at best, it is totally incapable of telling us whether the oberved price move (or price pattern, in the case of chartism) is *conditionally* more or less likely to occur today.

A Concrete Example

- At RBS we have developed a multi-state Markov model for vectors of variables (yields on a yield curve, FX rates etc).

- When applied to FX rates the analysis shows two states for JPY, one excited and one normal.

- In the excited state there is high volatility a pronounced strengthening of the Yen.

- In the normal state there is low volatility and modest weakening in the Yen.

- When Yen is excited, NZD and AUD are more likely to be excited.

- In the excited state for NZD and AUD these two currencies are weakening.

- These pieces of information taken together are *consistent with* a carry-trade model of reality.

- Having a model of reality can give us a conditional assessement of whether we are facing a danger period *today*.

Conclusions

- We have presented a methodology for aggregating in a coherent manner the ‘stress’ losses associated with a number of pre-chosen scenarios.

- The risk manager is required to provide a non-symmetric matrix of conditional probabilities conjoining the various scenarios.

- A procedure is then suggested to assess whether the suggested matrix is internally consistent. If not, we show how to find a ‘closest’ admissible solution.

- We provide a worked-out example to describe the procedure in detail.

- The procedure can be used to address the regulators' requirements for the Incremental Risk Charge.