The ERI Stress Testing Tool: A Coherent Approach to Stress Testing

Riccardo Rebonato
EDHEC Business School
EDHEC Risk Institute

September 18, 2017

1 The ERI Stress Testing Tool: Background and Motivation

Risk management, prudential macro- and microregulation, portfolio allocation and, in general, the strategic analysis of financial and economic outcomes share the common unstated assumption that the past conveys useful statistical information about the future. Indeed, a large part of contemporary finance rests on modern portfolio theory, which in turn places at centre stage the statistically-determined vector of asset expected returns and their covariance matrix.

In normal market conditions, the frequentist techniques that underpin these statistical analyses work well, and are perfectly justifiable. In these contexts, the role played by domain knowledge and subjective inputs to the determination of the statistical quantities of interest is limited – and often regarded with suspicion.

In recent years policy makers, regulators, portfolio managers and economic agents in general have been faced more and more frequently with situation of quasi-Knightian uncertainty. Take, as a salient example, the possible demise of the Euro. It is not clear what patches of past history could be relevant to provide ‘objective’ (frequentist) guidance about the expected outcomes of economic and financial variables. In these situations, subjective judgement and expert domain knowledge are forced to the fore. One enters the relatively unchartered territories of stress testing and scenario analysis.

As unprecedented financial and macropolitical events (or at least the fear thereof) seem to have visited the third millennium with disconcerting regularity, it comes to little surprise that there should have been a renewed interest in tail events in general, and stress testing in particular. And, of course, the financial crisis of 2007-2009 has put under the spotlight the failures of traditional (frequentist) risk management techniques such as VaR.

An additional reason for this renewed interest in stress testing can be traced to the association that the new regulatory regime has established between the
capital held by systemically important financial institutions (SIFIs) and the outcomes of stress testing exercises.¹ This development has been welcome, because purely statistical measures of risk by themselves have been proven to be inadequate to quantify the amount of capital financial institutions should hold. Stress testing can, in this respect, provide a useful complement to techniques such as VaR, and can fulfill additional functions (such as signalling in periods in market distress).

While there has been a wide consensus that more resources, and more thinking, should be devoted to stress testing, there have been relatively few new ideas on how to conduct a useful and believable stress testing programme. One of the most interesting proposals (see, eg, Rebonato (2009), Rebonato and Denev (2014), Denev (2016)) has been to harness the intuitional appeal and theoretical rigour of Bayesian nets for the task. Given the importance and the benefits of having at one’s disposal realistic, intuitively appealing and mathematically well-grounded stress testing and scenario analyses, ERI is launching a Stress Testing Monitor.

This note explains its rationale, its theoretical and practical underpinnings, and how the investment community can make use of its offerings.

2 Current Approaches to Stress Testing

Simplifying greatly, there are currently three main approaches to stress testing. For lack of an accepted nomenclature, I will refer to these three approaches as the extreme-tail approach, the vulnerability-driven approach, and the coherent-stress-testing approach. With a broad brush, they can be described as follows.

2.1 The Extreme-Tail Approach

I call ‘extreme-tail’ those statistical approaches based either on giving more weight to extreme events present in the data base, or on the extrapolation of a loss distribution obtained using traditional statistical methods.

When it comes to giving extra weight to extreme recorded occurrences, this is often done by choosing ‘blocks’ of past data that are supposed to reflect particularly stressful conditions (as it is done, for instance, with ‘Stressed VaR’ or ‘Stressed Historical Simulation’). See Meucci (2005, 2008, 2010a and 2010b) for related techniques.

As for the techniques of choice in carrying out the extrapolation, these are often based on Extreme Value Theory (EVT). (See Kotz and Nadarajah (2000) for a review of the theory and general applications of EVT and MacNiel(1999) and MacNiel, Frey and Embrechts (2005) for specific applications of EVT to risk management.)

When used by a financial institution for stress-testing purposes, the extreme-tail approach typically makes use of the risk technology already used for ‘normal

¹For a survey of the stress-testing-related regulatory initiatives of the post-Basel-II period, see, eg, Cannata and Quagliariello (2011), Chapter 6 and Part III in particular.
risk management’ (such as the mappings from risk factors to profit and losses, the risk engine, the construction on the univariate distribution of losses, etc), but attempts to ‘fatten’ or ‘to look further out in’ the tails.

The main problem with Stressed VaR or Stressed Historical Simulation is that there still is a full reliance on what has happened in the past, and, in particular, on how different risk factors moved during the past crises in the data set. As a consequence, truly novel event (such as, say, the break-up of the Euro mentioned in the introductory section) cannot not be handled in a satisfactory manner.

Approaches such as Extreme Value Theory (which has often been advocated in this context – see, eg, McNeil (1999)), have two advantages: the first is that they do offer a probabilistic assessment of the likelihood of losses of different magnitude, and in this sense tells us whether ‘we should worry’. The second is that they are extremely difficult to ‘game’ because it is well-nigh impossible to extract what combinations of risk factors give rise to the largest losses.

This lack of transparency, however, is also their Achilles’ heel. Since it is virtually impossible to associate a ‘narrative’ about what conjunction of events could be generating a given loss, they require an extremely strong leap of faith on the part of a senior decision-maker, who may be required to carry out expensive hedges, raise fresh equity capital, or impose costly restriction to a business, based on the output of the ultimate ‘black box’.

It is probably for this reason that Extreme Value Theory, for all its promise, has to date found little practical application in the management of stress risk.

2.2 The Vulnerability-Driven Approach

This approach has been advocated by the Bank of England (2007), and describes the practice of focussing on the identification on ‘a small numbers of key vulnerabilities.’ Once the vulnerabilities have been identified, ‘[t]he next step is to identify stress scenarios that could expose these vulnerabilities’.2

Their approach has mainly a macroprudential focus, and the argument is made that ‘[t]here are considerable benefits from rigorous, quantitative assessment of the materiality of financial stability vulnerabilities.’ Their 2007 paper advocates a ‘systematic and analytical approach to assessing these vulnerabilities, including a broad-based attempt to assess their materiality in terms of probability and impact’.3

As we mention above, however, it is difficult to reconcile an assessment of materiality in terms of probability when a vulnerability-based approach is adopted, unless one relies on the comovements in the risk factors obtained using pre-crisis data and on unconditional probabilities – and, as highlighted above, if one does so the approach is only partly ‘forward looking’. Absent this narrative about what gives rise to an exceptional move in the vulnerability, it is difficult to engage the attention of decision makers, and to gain their ‘buy-in’ if corrective

---

2 page 8
3 emphasis added
actions are to be taken. This consideration is of less concern if the stress test is only used for macroprudential purposes (as the ‘consumers’ of its output are regulators and/or the monetary authorities), but becomes very important if the mandated stress tests are to be perceived by the banks that have to run them as something more than a compliance exercise.\(^4\)

Second, the same Bank of England paper suggests that once should work one’s way backwards from the identified vulnerability to ‘stress scenarios that could expose these vulnerabilities’. However, this is an approach more akin to reverse stress testing than to a true scenario analysis. As in the case of reverse stress testing, for a complex portfolio there are in general a myriad of scenarios that can stress the same vulnerabilities, and each one has a different likelihood of occurring. Under rather restrictive assumptions, Saroka and Rebonato (2015) and Rebonato (2016) show how the ‘most likely’ (or Maximum Entropy) scenario can be identified, but also their approach is not without problems (see Rebonato (2017) for a possible solution to these problems).

### 2.3 The Coherent-Stress-Testing Approach

The coherent-stress-testing approach consists

- in the assignment of one or more ‘root causes’ (e.g., a geopolitical event, a precisely identified macrofinancial or macroeconomic shock, etc);
- in the identification of transmission channels (typically probabilistic, but sometimes deterministic) that connect the root causes to the risk factors affecting a given portfolio; and
- in the linking of these transmission channels to the transactions (e.g., for a hedge fund), to the portfolio (e.g., for a complex bank or an asset manager), or to the system (e.g., for a regulator) of interest.

When the effect of the scenario under examination on a whole portfolio is of interest, typically one will have to determine the deformation of the unconditional joint distribution of factors which is consistent with the assumed stress test. This is particularly important for portfolio management applications, where the altered volatilities and codependencies among the component securities become crucial.

There have been some developments and concrete proposals on how such an ambitious task can be accomplished, and Probabilistic Graphical Methods (Denev, 2105) have been proposed.\(^5\)

\(^4\)Regarding Basel II, the then-Governor of Banco de España and Former Chairman of the Basel Committee, Mr Jaime Caruana said: “Basel II is not intended simply to ensure compliance with a new set of capital rules. Rather, it is intended to enhance the quality of risk management and supervision” (emphasis added). The same sentiment should inspire the use of stress testing as well. Alas, in case of Basel II the risk management practices it inspire ended up being very different from the regulators’ aspirations.

\(^5\)In recent applications, these probabilistic graphical methods have mainly, but not only, been Bayesian nets. See Rebonato (2010), Rebonato and Denev (2014).
For the approach to be feasible, conditional probabilities have to be assigned between the roots and the transmission channels and from these to the transactions or portfolios of interest. Since, in many, if not most, cases, there will be few historical precedents for the root scenario (again, think of the default of a Euro-member peripheral country), very often these probabilities will have to be assigned subjectively, or could be the output of structural models. When subjective probabilities are called for, it has been shown that, thanks to the causal interpretation with which these methods can be endowed, assigning them is an arduous, but cognitively feasible, test.

Moving to the implementation features of these approaches, the creation and maintenance of these coherent scenarios can require the assignment and updating of a large number of input probabilities. There exist elicitation techniques to facilitate the task and improve the quality of the subjective assessments (see, eg, Gigerenzer and Edwards (2003), Gigerenzer and Hoffrage (1995)), but the ‘cognitive investment’ should not be underestimated.

On the positive side, if well structured, these coherent stress tests, which can be naturally translated into a causal and temporal ‘narrative’, can be readily understood and questioned by an intelligent, but not necessarily mathematically strong, senior decision maker. From these perspective they lend themselves readily to intervention and remedial action. They also provide an approximate estimate of the probabilities attaching to the various losses, and allow the identification of the conjunction of events (configuration of states for the transmission channels) that give rise to the losses. This can provide a ‘sanity check’ prior to intervention.

Coherent stress tests can be used most easily at a transactional level (because the transmission channels are generally few, and their effects only have to be focussed on a specific transaction). When used at the portfolio level, they share some of the same problems highlighted with reference to vulnerabilities-based approaches, namely that it can be difficult to cascade the changes in a handful of salient risk factors that are stressed ‘by hand’ in a coherent stress test to the thousand of risk factors that will affect a complex portfolio. This is indeed one of the ‘engineering problems’ addressed in Rebonato (2017).

To summarize: if one agrees that intuitive appeal, ability to associate an approximate probability to the outcomes, cognitive resonance and mathematically solid foundations are all highly desirable in a useful and effective stress testing programme, then one is naturally lead towards a well-designed coherent stress testing programme. The rest of this note argues that Bayesian nets can offer the technology to implement this coherent stress testing programme.

3 Bayesian Nets in a Nutshell

In this section we provide a brisk-paced, yet gentle, introduction to the application of Bayesian Nets to stress testing. For more details on Bayesian nets in general, the reader is referred to Pearl (2009), Jensen and Nielsen (2007) for a thorough treatment, or to Murphy (2011) for a user-friendly, informal, yet
precise, discussion.\textsuperscript{6}

Graphical methods – of which Bayesian Nets are the best known and the simplest, but by no means the only, incarnation – satisfy many important desiderata mentioned in the opening sections. They are unsurpassed when it comes to intuitional appeal. They have a rigorous logical and mathematical underpinning. They allow the parsimonious factorization of a complex joint probability distribution into a handful of important conditional probabilities (ie, probabilities of the type ‘If A has happened, what is the probability of B happening?’). They speak the cognitive language of the human mind – ie, the language of causation. They lend themselves to ready sensitivity analysis – which means that the intelligent-but-not-quantitative senior professional mentioned above will be able to get a feel for how much she can trust the results. In short, they give us an idea of whether we ‘should worry’.

And, last but not least, understanding the intuition behind their power is disarmingly simple: suppose that you are interested in the probability of John slipping tomorrow on the pavement outside his front door – this is our ‘stress event’. The probability of John slipping will depend (inter alia) on whether the pavement is wet or dry. Whether the pavement is wet depends, in turn, on the probability of rain tomorrow, and on the probability of the garden sprinkler being active. This is shown in the Bayesian net in Fig (1), in which A represents the probability of rain tomorrow; B represents the probability of the of the sprinkler being on; C denotes the probability of the pavement being wet; and D is the probability of John slipping.

The central insight behind the probability factorization produced by Bayesian nets (an insight that goes under the name of conditional independence) is that, once we know whether the path is wet or dry, it does not matter if it is wet because of the rain or because of the sprinkler: the wet/dry status of the path contains all the necessary information needed to determine the probability we are interested in – the variable “John slips” is ‘screened’ from the variables rain and sprinkler by the variable “Pavement wet”.

If the simplification in the cognitive and computational burden required to arrive at the probability we are interested in seems in this case rather underwhelming, it is easy to see that the computational and cognitive savings could be much more substantial for realistically complex nets.

Quantitatively, a Bayesian net is a Directed Acyclical Graph (ie, a set of – usually Boolean – variables with arrows pointing in the direction of causal flow\textsuperscript{7}) equipped with conditional probability tables. Mathematically, the key to

\textsuperscript{6}For specific computational issues (such as how to simplify Bayesian nets in the case of causal independence) related to the efficient construction of Bayesian nets see, eg, Heckerman and Breese (1994) and Jurgelenaite and Lucas (2005). For specific applications to stress testing and portfolio management problems, see Rebonato (2010), Rebonato and Denev, (2010), Rebonato and Denev (2011), \textit{Coherent Asset Allocation and Diversification in the Presence of Stress Events}, Rebonato and Denev (2012). For an extension of Bayesian nets (which are Directed Acyclical Graphs) to more general probabilistic graphical methods, see Denet (2015).

\textsuperscript{7}Bayesian nets need not be endowed with a causal interpretation. If they are causally constructed, however, there are important computational and cognitive savings. See Rebonato and Denev (2014) in this respect.
Figure 1: The Bayesian net associated with the John-slipping-on-path scenario. 
A represents the probability of rain tomorrow; B represents the probability of 
the sprinkler being on; C denotes the probability of the pavement being wet; and 
D is the probability of John slipping. Once we know whether the pavement is 
wt or dry, we do not need to know whether it is wet because it rained or because 
the sprinkler was on in order to assess the probability of John slipping on the 
floor. The boxes associated with the two roots (‘Rain’ and ‘Sprinkler’) give the 
unconditional probabilities of rain or good weather tomorrow (0.3 and 0.7), and 
of the sprinkler being on or off (0.4 and 0.6). In the box labelled ‘Pavement’ the 
zeros and ones on the right-hand columns give the probabilities of the pavement 
being wet or dry given the different combinations of rain and sprinkler on or 
off. Graphical display courtesy of Black Goose (see www.avisnigra.com).
successful Bayesian net construction is the identification of variables that enjoy effective ‘screening’ (as in the case of John on the slippery pavement). Screening occurs whenever conditional independence applies, i.e., whenever
\[ P(A|BC) = P(A|B), \]
or, in words, whenever knowledge of the states of \( B \) and \( C \) does not permit a better prediction of \( A \) than if just \( B \) had been known.

The computational saving alluded to above can be appreciated by recalling the Master theorem for Bayesian Nets, which says that the joint probabilities among \( n \) variables, \( \{x_1, x_2, ..., x_n\} \) (which encode all that we can know about the variables), are given by the product of the conditional probabilities, conditioned on their parents only:
\[ p(x_1, x_2, ..., x_n) = \prod_{i=1}^{n} p(x_i|\text{par}[x_i]) \]
where \( \text{par}[x_i] \) denotes the parents of variable \( x_i \), and the conditional probability of the root (i.e., of the variable that has no parents) coincides with its marginal.\(^8\)

It is this factorization that allows that great computational (and cognitive!) savings afforded by Bayesian nets.

As for the objections that Bayesian nets require the assignment of (supposedly difficult to specify) conditional probabilities, consider the following. You would be probably hard pressed if asked to provide the probability of you being hit by a car tomorrow; and you would find it no easier to assign the probability of ending up in hospital tomorrow. However, I think you can safely venture a guess of 40-60 per cent for the probability of you ending up in hospital, given that you have been hit by a car. Once again, when used in the causal direction, conditional probabilities are natural and cognitively ‘easy’ to assign.

Of course, many (or most) of the conditional probabilities a Bayesian net requires are subjective in nature. The agnosticism of the approach as to where the probabilities come from, should, however, be seen as an advantage, not a drawback. This is because the user of the net is unshackled by the constraints of a statistical (frequentist, past-data-based) approach that is of little or no use to analyze novel situations: where would we look up in our data series, for instance, if we wanted to glean information about what would happen if Greece left the Euro next Monday?

\(^8\) The second cornerstone of Bayesian Net construction is the No-Constraint theorem, which says that if one deals only with ‘canonical’ conditional probabilities – i.e., exactly the conditional probabilities required by the conditional probability tables – one can assign to them any number between 0 and 1, and rest assured that the resulting Bayesian Net will be mathematically consistent. See Rebonato and Denev (2013) for a proof of the theorem, and for the definition of canonical conditional probabilities. See Moskowitz and Rebonato and Kwiatkowski for the problems that can arise if one assigns non-canonical conditional probabilities.
4 An More Realistic Example

A concrete, if rather stylized, example, can help putting some flesh on the blueprint sketched above. We consider the scenario of a hard Brexit from the perspective of a UK asset manager who has under her watch both an equity and a fixed-income (real and nominal) sterling-denominated portfolio.

The transmission channels that link the root event (the occurrence or otherwise of a 'hard' Brexit) to the main proxy asset in the portfolio of interest (the FTSE 100 index, the 10-yr and 1-year Gilt yields, and the break-even-inflation of the 10-year Linker) are the following (see Fig (2)):

1. *Ceteris paribus*, the UK economy (node 5 — "GDP") is expected to enter a difficult period: the more so, the harder the outcome of the Brexit negotiations.

2. What could help the state of the UK economy would be a reduced current account deficit, and an export-led expansion (node 4 — "Trade").

3. The trade conditions would be a balance between the outcome of the negotiations (a hard Brexit could hinder trade with the EU: node 1 — "Hard Brexit"), and the weakness of the pound (node 2 — "Sterling") that could help exports to non-EU countries.
4. A weakness in the pound would be likely to give rise to ‘imported inflation’ (node 2 - "Inflation").

5. Inflation and the state of the economy could give conflicting signals to the Bank of England in its setting of the monetary policy (node 6 - "Monetary Policy").

6. A particularly weak state of the economy could also convince the monetary authorities to engage in some form of non-conventional monetary policies, such as the purchasing of long-dated assets (node 7 - "QE").

7. Finally we reach the ‘leaves’ of the tree: the yield of the 1-year Gilt will mainly depend on the path of the monetary policy; the 10-year Gilt yield would also be affected by outright asset purchases; the FTSE 100 would be affected by a strong or weak economy, but also by the possible weakness in the pound (as the behaviour of the index in the aftermath of the Leave vote as shown); finally the break-even-inflation of the UK Linkers would be affected, among other things, by the level of inflation.

The analysis has been left very sketchy, but it should convey the nature of the building process. Noticeably, an intelligent non-mathematically-versed user can challenge the scenario: for instance, we have left out on purpose a direct link between "Sterling" and "Inflation" as an example of scenario modification that can be prompted by the critical analysis of the construction of the Bayesian net. Note that macro financial knowledge and understanding is needed to improve the scenario, not mathematical wizardry: this is the type of discussion that can, and should, be had at a Board meeting or Portfolio Management Committee.

As for the outputs, one intermediate advantage is the ability to carry out ‘sanity checks’. Once the conditional probabilities have been provided, one in fact obtains all the joint probabilities for the various possible combination of events. When one does this, what is produced is a table, a small section of which is shown in Fig (3).

This section shows (highlighted in red) the most likely configuration of occurrences. Does it make sense? Let’s see.

From the table we read that the most likely combination of events is the following:

We do not have a hard Brexit, and therefore the pound does not weaken significantly. Inflation remains low because the UK does not experience ‘imported inflation’, but trade conditions are poor because there is no ‘sterling boost’. The GDP remains weak and, because, of the subdued inflation, there is no need for a hawkish stance by the monetary authorities. However, given the weakness of the economy, and given the low levels of rates, the authorities engage in unconventional monetary operations, flattening the yield curve. Absent the sterling boost, and given the weak economy, the FTSE 100 performs poorly.

There is no point in asking whether this answer is ‘right or wrong’. However, this is a good example of how the output of a Bayesian net can be interrogated, and modified if found wanting. From practical experience in the investment
and portfolio management communities, the benefits from these discussions and analyses are in general at least as large as the quantitative outputs resulting from the Bayesian net.

5 Conclusions

To conclude. Both the industry and the regulators are paying more and more attention to stress testing. This is refreshing and encouraging. However, a stress-test-based review of capital adequacy and liquidity resilience can only be achieved if an approximate, but order-of-magnitude-meaningful, assessment of the probability and severity of a stress is possible: a stress test must pass the “Should we worry?” test to be of any use.

Similarly, when we move from stress testing to asset allocation, diversification tends to work well until it doesn’t, and it has the habit of stopping to function just when we need it most. What we need are therefore tools to make portfolios resilient not only to garden-variety market shocks, but also to rare, yet devastating, events. And, when we analyze scenarios which may be not necessarily cataclysmic, but simply not encountered in our data set, we must be able to understand (not just ‘calculate’) how different asset classes will interact, and how the portfolio will respond.

The ERI Stress Testing Tool offered by the EDHEC Risk Institute is a first
step towards making available to the investment community a tool to fulfill these needs. As Rebonato (2017) points out there are still many ‘engineering’ problems that stand between an enticing idea and an industry-strength application, but there are also interesting ideas on how to solve them.

The ERI Stress Testing Tool will progressively offer more and more ways to analyze portfolios under stress. What will be soon released should however already make a significant contribution, by showcasing the capabilities of the Bayesian-net technology and providing a solid foundation on which more sophisticated applications can be built.

6 References


International Monetary Fund, (2010), *Macrofinancial Stress Testing – Principles and Practices*, document prepared by the Monetary and CApital MArkets
Department, August 22, 2012


Meucci A, (2010a), *Factors on Demand*, Risk, 23, 84-89


13