WHY BANKS FAILED THE STRESS TEST

Andrew G Haldane*

Executive Director for Financial Stability
Bank of England

13 February 2009

The basis for a speech given at the Marcus-Evans Conference on Stress-Testing,
9-10 February 2009

* I would like to thank Geoff Coppins, Harry Goodacre, Nigel Jenkinson, Salina Ladha, Iman van Leyveld, Jack McKeown, Joe Noss, Adrian Penalver, Nick Vause and Matthew Willison for contributions and comments.

WHY BANKS FAILED THE STRESS TEST

By any historical standard, the financial crisis of the past 18 months has been extraordinary. Some have suggested it is the worst since the early 1970s; others, the worst since the Great Depression; others still, the worst in human history. Time will tell.

Risk managers are of course known for their pessimistic streak. Back in August 2007, the Chief Financial Officer of Goldman Sachs, David Viniar, commented to the *Financial Times*:

“We are seeing things that were 25-standard deviation moves, several days in a row”

To provide some context, assuming a normal distribution, a 7.26-sigma daily loss would be expected to occur once every 13.7 billion or so years. That is roughly the estimated age of the universe.

A 25-sigma event would be expected to occur once every 6 x 10^{124} lives of the universe. That is quite a lot of human histories. When I tried to calculate the probability of a 25-sigma event occurring on several successive days, the lights visibly dimmed over London and, in a scene reminiscent of that Little Britain sketch, the computer said “No”. Suffice to say, time is very unlikely to tell whether Mr Viniar’s empirical observation proves correct.

Fortunately, there is a simpler explanation – the model was wrong. Of course, all models are wrong. The only model that is not wrong is reality and reality is not, by definition, a model. But risk management models have during this crisis proved themselves wrong in a more fundamental sense. They failed Keynes’ test – that it is better to be roughly right than precisely wrong. With hindsight, these models were both very precise and very wrong.

For that reason, 2008 might well be remembered as the year stress-testing failed. Failed those institutions who invested in it in the hope it would transform their management of risk. Failed the authorities who had relied – perhaps over-relied – on the signal it provided about financial firms' risk management capabilities. And, perhaps most important of all, failed the financial system as a whole by contributing, first, to the decade of credit boom and, latterly, the credit bust.

That is a stark conclusion. But it is a conclusion which is hard to escape. When tested against real stress, large parts of the financial system seized-up and a number of financial institutions failed. Against that backdrop, now is as good a time as any for candour about what went wrong. That is the purpose of my comments today: to diagnose some of market failures or frictions in stress-testing practices highlighted by the crisis; and, more speculatively, to suggest some practical ways in which stress-testing might deliver answers which are "roughly right".

The Golden Decade

To understand the recent failures in risk management, some history is instructive. Prior to the current financial crisis, the previous two low tide marks for the financial system and risk management were the stock market crash of October 1987 and the failure of the hedge fund LTCM in September 1998. Both prompted a sea-change in risk management practices and technologies.

The October 1987 crash in many respects marked the birth of Value at Risk (VaR) as a key risk management tool in financial firms. By 1989, Dennis Weatherstone, JP Morgan's then-chairman, called for a "4:15 Report", which combined all of the firm's data on market risk in one place. That report should contain information sufficient to answer the question "How much could JPM lose if tomorrow turns out to be a relatively bad day?"

With this as the top-down edict, it is perhaps unsurprising that JP Morgan were an early-developer and early-adopter of VaR. By 1996, they had published their methodology and the detail of the parameterisation of their risk models. In 1998 RiskMetrics Group, an independent for-profit business, spun off the JP Morgan methodology and began offering consultancy services to the risk management community. And from 1997 onwards, VaR came to take a degree of prominence within the regulatory community, with first the US SEC and subsequently the Basel Committee on Banking Supervision (BCBS) giving further impetus to VaR, including through the design and implementation of Basel II.

By 2006, when Philippe Jorion published his famous textbook on VaR, relatively few would have disputed the claim in its title - "Value at Risk: The New Benchmark for Managing Financial Risk". The message was clear: the technological frontier of risk management had been shifted outwards decisively. A cursory search suggests that there have been more than 200 books published on VaR since the October 1987 crash, or roughly one a month.

The date of birth of stress-testing is harder to trace. Early mention is made of it in a technical note by RiskMetrics in 1996. But it is clear that stress-testing was given considerable impetus by the failure of LTCM more than a decade after the October 1987 crash. Unlike VaR, which had private sector origins, the official sector appears to have been at least as much a driver behind the adoption of stress-testing. By 2001, under the auspices of its Financial Sector Assessment Programme (FSAP), the IMF was publishing details of its stress-testing methodology and experience. Today, the same cursory search reveals over 250 articles on stress-testing in the past ten years, or more than one a fortnight. We were experiencing a second wave of technological revolution in risk management.

This technological transformation contributed to what was, with hindsight, an extraordinary period of growth and success for the financial system and financial markets – a Golden Decade. Between October 1998 and June 2007, banks' share prices increased almost 60% and their balance sheets rose more than threefold. In some markets growth was little short of explosive, with the rise in volumes outstanding in the CDS market making Moore's Law look positively sluggish.

And why was this credit boom not destined to end in bust? Because this time was different. At the same time as returns were being boosted by bigger balance sheets and financed by higher leverage, risk was being held in check by a shift in the technological frontier of risk management. A new era had dawned, one with simultaneously higher return and lower risk. This miracle came care of a compelling combination of cavalier risk-takers and roundhead risk-managers. Or so ran the rhetoric.

With hindsight, this Golden Decade and its aftermath has all the hallmarks of, in Charles Kindleberger's words, Manias, Panics and Crashes. Enthusiasm about return gave way to hubris and a collective blind eye was turned to the resulting risk. This was a latter-day version of the Hans Christian Andersen fairy-tale, "The Emperor's New Clothes". In a classic collective delusion, the Emperor's new clothes, you will recall, were admired by all. Conferences like this one became catwalks for banks and the authorities alike, parading their new garments through the streets in all their finery. Risk modelling became high fashion for the pointy-heads, haute-couture for the anoraks.

The past two years have rather changed all that. The sub-prime market has played the role of the child in the fairytale, naively but honestly shifting everyone's perceptions about how threadbare the financial system had become. The madness of crowds, as Charles Mackay so vividly put it, became visible to all. The resulting unravelling of the Golden Decade has been little short of remarkable.

Asset prices have collapsed – for example, world equity prices have lost more than three-quarters of their gains during the Golden Decade. Prices of banks' shares have fared even worse, losing almost 60% of their value and are now lower than at the start of the Golden Decade. In the face of these falls, risk management systems across virtually all institutions have been found badly wanting. A survey of 500 risk managers by KPMG in October last year found that 92% intended to review their risk management practices.

Estimated losses within the financial sector since the start of the crisis lie anywhere between a large number and an unthinkably large one. Today, managers of risk – the authorities just as much as banks – find themselves struggling to preserve their dignity, with risk management systems a combination of sack-cloth and fig-leaf. This year, stress-testing conferences like this one are more doghouse than catwalk.

Diagnosing the Market Failures

So what were the failures, specifically of stress-testing and other risk management tools, that contributed to this credit boom and subsequent bust? It is useful to try and identify the micro-economic friction – the market failure – that was the root cause of these risk management problems. Doing so better enables both financial institutions and the authorities to pinpoint what needs to change and how. These market failures fall roughly into three categories:

- disaster myopia;
- network externalities; and
- misaligned incentives.

All three have impeccable microeconomic credentials and potentially disastrous macroeconomic consequences.

<u>Disaster Myopia</u>

In a nutshell, disaster myopia refers to agents' propensity to underestimate the probability of adverse outcomes, in particular small probability events from the distant past. That makes it sound like a rather unworthy informational failure. In fact, it is well-established in cognitive psychology that economic agents have a tendency to base decision rules around rough heuristics or rules of thumb.[1] The longer the period since an event occurred, the lower the subjective probability attached to it by agents (the so-called "availability heuristic"). And below a certain bound, this subjective probability will effectively be set at zero (the "threshold heuristic").

If the period of stability is sufficiently long – a Golden Decade perhaps? - this subjective approach to evaluating probabilities looks increasingly like a fully-rational, Bayesian approach to updating probabilities. As time passes, convincing the crowds that you are not naked becomes progressively easier. It is perhaps no coincidence that the last three truly systemic crises – October 1987, August 1998, and the credit crunch

[1] For example, Kahneman, Slovic and Tversky (1982).

which commenced in 2007 – were roughly separated by a decade. Perhaps ten years is the threshold heuristic for risk managers.

Models of disaster myopia have been used to explain a number of phenomena, including the tendency for drivers to slow down having witnessed an accident and then speed up once the accident has become more distant in their memory, and for people to under-insure against low frequency natural hazards such as earthquakes and floods. In the context of financial crises, disaster myopia has been used to explain the LDC debt crisis, the US savings and loans debacle and various commercial property crises.[2] The credit crunch of the past 18 months is but the latest in a long line of myopia-induced disasters.

Such disaster myopia is not of course confined to the private sector. The official sector is just as likely to succumb to cognitive biases borne of long periods of stability. With hindsight, the stress-tests required by the authorities over the past few years were too heavily influenced by behaviour during the Golden Decade. Many risk management models developed within the private sector during the Golden Decade were, in effect, pre-programmed to induce disaster myopia. These models were often data hungry. Improvements in data and IT technology were able to feed these beasts with vast, high-frequency datasets. This provided, in the statistical jargon, ample degrees of freedom for modellers, enabling them to devise risk frameworks which, on the face of it, were very precisely calibrated *in-sample*.

And there's the rub. The sample in question was, with hindsight, most unusual from a macroeconomic perspective. The distribution of outcomes for both macroeconomic and financial variables during the Golden Decade differed very materially from historical distributions.   Charts 1-8 illustrate this small sample problem. They look at the distribution of a set of macroeconomic and financial variables, comparing the Golden Decade with a sample stretching back in some cases to the 17th century. Even visually, these distributions plainly suggest that the Golden Era distributions have a much smaller variance and slimmer tails. More formally, Table 1 looks at the first four moments of these variables, comparing the Golden Era with the full sample.

[2] For example, Guttentag and Herring (1986a) and Herring (1999).

For the macro time-series, the differences in variability are striking. The long-run standard deviation of UK GDP growth has on average been 4 times greater than during the Golden Decade; for unemployment 5 times greater; for inflation 7 times greater; and for earnings 12 times greater. Put differently, as part of the Basel II regime the FSA require banks to simulate the effects of a 1-in-25 year stress. In 2007, the worst such GDP growth outcome over the preceding 25 year period was -1.4%; the average 1-in-25 year stress over the full sample is -3.8%.

For financial time-series, small sample problems are even more acute, especially for events in the tail of the distribution. Measures of kurtosis – the fatness of the tails – of UK house price inflation are 6 times larger over the full sample than over the Golden Decade; for UK bond yields 10 times larger; and for UK equity returns 16 times larger. To bring these stylised facts to life, consider the distribution of equity returns in Chart 7. If we assumed the Golden Era distribution was the true one, the three worst monthly returns in history – the bursting of the South Sea bubble in September and October 1720, and Black Monday in October 1987 – would have been respectively 12.7, 6.9 and 6.5-sigma events. All three would have appeared to be once in a lifetime – of the universe – events.

Underestimation of risk, whether variances or tail outcomes, has consequences for the risks facing both individual firms and for the system as a whole. As an example of the former, Chart 9 plots some unconditional 90th percentile VaRs for a selection of UK banks, based on their equity returns up until end-July 2007 and then extended to include the present crisis.[3] These unconditional VaRs for UK banks increase, on average, by almost 60% once the sample is extended; and for some banks these risk measures more than double.

For the system as a whole, one way of illustrating the consequences of underestimating risk is to translate it into "fair value" insurance premia. For example, consider a financial firm offering insurance against moves in future equity prices by writing put options in mid-2007. Pricing of this insurance is assumed to be based on

[3] The banks themselves have been anonymised to protect the innocent.

the distribution of equity returns during the Golden Decade. If the "true" distribution of returns were its long-run average, by how much would this insurance have been under-priced in 2007?

Chart 10 provides some answers for a selection of strike prices for the option. The degree of under-pricing of risk is large and is larger for options designed to protect against tail risks (lower strike prices). For at-the-money options on UK equities, the insurance premium would have been under-priced by around 45%; for options well out-of-the-money – say, 50% below equity prices at the time – the mis-pricing would have been nearer 90%. This is risk under-pricing on a dramatic scale.

These examples are no more than illustrative. But they help illustrate that the quantitative consequences of disaster myopia were material ahead of crisis and may have contributed importantly to the price of risk being set too low. And that, in turn, helped sow the seeds of the credit boom.

Network Externalities

Any asset portfolio is, in essence, a financial network. So the balance sheet of a large financial institution is a network, with nodes defined by the assets and links defined by the correlations among those assets. The financial system is similarly a network, with nodes defined by the financial institutions and links defined by the financial interconnections between these institutions.

Evaluating risk within these networks is a complex science; indeed, it is the science of complexity.[4] When assessing nodal risk, it is not enough to know your counterparty; you need to know your counterparty's counterparty too. In other words, there are network externalities.[5] In financial networks, these externalities are often referred to as contagion or spillovers. There have been many examples of such spillover during this crisis, with Lehman Brothers' failure a particularly painful one.

[4] Gell-Mann (1994).
[5] Morris and Shin (2008).

That is why there have been recent calls to calibrate regulatory requirements to these risk externalities.[6]

These network uncertainties make it tremendously difficult for risk managers to identify and price, and hence manage, balance sheet risk. Consider first evaluating risks across the portfolio of an individual firm. There is evidence that firms find aggregation of risks across their balance sheet extremely difficult to execute.[7] To the extent this is done at all, it requires firms to make assumptions about correlations between asset prices. But at times of stress, asset correlation matrices are unlikely to be stable and correlations invariably head towards one. So pre-crisis measures of balance sheet risk are likely to be significant under-estimates. Chart 11 looks at asset correlations over the past few years. Note their instability and abrupt upward shift during crisis.

These risk externalities will tend to be amplified when aggregated across the network as a whole. This generates further underestimation of institutional risks. Consider again those 90th percentile VaR measures for UK banks. But instead of looking at unconditional VaR, consider now *conditional* VaRs (CoVaRs) – that is to say, VaRs conditional on other institutions in the network simultaneously facing stress.[8] As Chart 12 shows, this raises the median risk facing UK banks by around 40%; and for some banks, risk estimates almost double. For a financial firm leveraged 20+ times, those risk revisions could be the difference between success and failure.

Network risk externalities of this type impose formidable informational demands on banks. For example, understanding the full consequences of Lehman's failure would have required information on the entire topology of the financial network. This is unrealistic even for the authorities, much less an individual firm. Absent that knowledge, the financial system was seized by network uncertainty. If this informational failure is not easily rectified by the actions of individual firms, there is a

---

[6] See, for example, Brunnermeier, Crockett, Goodhart, Persaud and Shin (2009) and NYU Stern School of Business (2008).

[7] For example, a survey of stress-testing by the CGFS in 2005 found that only a small minority of firms considered the effects of multiple shocks on their balance sheet.

[8] Following the methodology of Adrian and Brunnermeier (2008).

case for the authorities attempting to provide that missing informational public good, however difficult that might be in practice.

Misaligned Incentives

Finally, and perhaps most contentiously, incentives and governance. Principal-agent problems crop up in all aspects of economics. But it is questionable whether there is any event in recent history where these agency problems have been exposed so frequently and extensively as during the current financial crisis. It is easy to see why. Financial innovation lengthened the informational chain from ultimate borrower to end-investor. The resulting game of Chinese whispers meant that, by the time information had reached investors at the end of the chain, it was seriously impaired.

In the narrower context of stress-testing, these principal-agent problems appear to have operated at two distinct levels. First, *internally*, through the relationship between risk managers and the risk-takers within financial firms; and second, *externally*, in the relationship between financial firms and the authorities. The former principal-agent problem has been rather less discussed, but appears to have been potent during the credit boom.

Decision-making within firms is an arm-wrestle between risk and return, between risk managers and risk-takers. When returns are high and risks appear low, this arm-wrestle can become one-sided. Power switches from back to front offices and risk managers become the poor relation.[9] And what is true within individual firms is then amplified by behaviour across the system as a whole, as firms conduct their own arm-wrestle with competitors for higher returns on equity. The Bank's market intelligence suggested this "keeping up with the Jones's" was a potent force within financial firms during the upswing.

The second principal-agent problem, between firms and the authorities, is different in kind but similar in consequence. It arises because of a familiar public policy problem – time-consistency. If the ex-post failure of an institution risks destabilising the

[9] The KPMG survey of risk managers in October 2008 pointed to a similar conclusion, as does the FSA consultation paper on stress-testing published in December 2008.

system, any ex-ante pre-commitment by the authorities to let it fail will lack credibility. This is simply a variant of the old adage that if you owe the bank a small amount it is your problem, a large amount it is theirs. These days, if a bank owes a small amount it is their problem, a large amount it is the authorities.

This time-consistency problem weakens incentives for banks to consider for themselves large-scale risks to their balance sheet which might induce failure. The safety net becomes a comfort blanket, the backstop a balm. And the greater the risk these institutions themselves pose in the event of failure, the weaker the incentives to manage risk. These are topsy-turvy incentives from a public policy perspective, with risk management discipline weakest among those whom society would wish it to be strongest.

And the evidence? A few years ago, ahead of the present crisis, the Bank of England and the FSA commenced a series of seminars with financial firms, exploring their stress-testing practices. The first meeting of that group sticks in my mind. We had asked firms to tell us the sorts of stress which they routinely used for their stress-tests. A quick survey suggested these were very modest stresses. We asked why. Perhaps disaster myopia – disappointing, but perhaps unsurprising? Or network externalities – we understood how difficult these were to capture?

No. There was a much simpler explanation according to one of those present. There was absolutely no incentive for individuals or teams to run severe stress tests and show these to management. First, because if there were such a severe shock, they would very likely lose their bonus and possibly their jobs. Second, because in that event the authorities would have to step-in anyway to save a bank and others suffering a similar plight.

All of the other assembled bankers began subjecting their shoes to intense scrutiny. The unspoken words had been spoken. The officials in the room were aghast. Did banks not understand that the official sector would not underwrite banks mis-managing their risks?

Yet history now tells us that the unnamed banker was spot-on. His was a brilliant articulation of the internal and external incentive problem within banks. When the big one came, his bonus went and the government duly rode to the rescue. The time-consistency problem, and its associated negative consequences for risk management, was real ahead of crisis. Events since will have done nothing to lessen this problem, as successively larger waves of institutions have been supported by the authorities.

More recently, the Bank and FSA have been engaged in some practical work with banks, running stress-tests through their models on common scenarios. When asked to assess the consequences of a macro stress-test, the like of which we are currently experiencing, some banks have found it problematic. In defence, they have suggested that such an exercise was only conducted annually as part of their Basel II preparations and as such new stress tests would take months to conduct.

This too was revealing. If even the most obvious stress-test took many weeks to prepare and assess, how could these tests meaningfully be used to manage risk? The short answer, I think, is that stress-testing was not being meaningfully used to manage risk. Rather, it was being used to manage regulation. Stress-testing was not so much regulatory arbitrage as regulatory camouflage.

Prescribing Some Solutions

Each of these market failures has been exposed by events over the past 18 months. When risks materialised outside of calibrated distributions, risk models provided little guidance in identifying, pricing and hence managing them. This failure is not of purely academic interest. The breakdown of risk models is itself likely to have contributed importantly to crisis dynamics. Why?

First, the potential losses arising from under-pricing of risk are large. Consider the earlier example of a disaster-myopic writer of deep out-of-the-money put options on UK equities, priced using distributions drawn from the Golden Decade. Let’s say that, in June 2007, a five-year put had been written on the FTSE-100 with a strike price 40% below the prevailing market price. Today, that put would be at-the-money.

Hedging that position would crystallise a loss roughly 60 times the income received from having written the option in the first place.[10]

This example is far from hypothetical. These are essentially the same trades undertaken by a number of insurance companies and other investors ahead of crisis. In the go-go years, the insurance premia from them yielded a steady income stream. But when risk shifted, many insurers have suffered large-scale losses as premia have adjusted and investors have scrambled to hedge. The large US insurer AIG has so far suffered gross losses totalling over $60bn on CDS contracts alone. Losses by the monoline insurers have also totalled in excess of $60bn.

Second, the breakdown of these models had the consequence of turning risk into uncertainty, in the Knightian sense.[11] Once the models broke down, how were assets to be priced? Practitioners have a devil of a job pricing assets in the face of such uncertainty. So too do academics, though some attempts have been made.[12] The theory of asset pricing under Knightian uncertainty throws up at least two striking results. First, in the face of such uncertainty, asset prices are not precisely determined but instead lie in a range. This indeterminacy in prices is larger the greater is uncertainty and the greater agents' aversion to it. Second, asset prices exhibit a downward bias relative to fundamentals. Uncertainty gives the appearance of "pessimistic" expectations.

Both of these theoretical predictions match pretty closely the moments of many asset prices in the world today. Many appear to lack a clear compass relative to fundamentals. Most are excessively volatile. Among investors, pessimism is the new optimism, with talk of a lost decade in succession to the Golden one. Risk models – or the failure thereof – have played their part in generating these foggy outcomes.

That is the diagnosis. What of the prescription? In their recent consultation paper, the Financial Services Authority has outlined some very good proposals for

[10] Roughly half of that loss represents under-estimation of the distribution of returns back in 2007.
[11] Knight (1921).
[12] For example, Epstein and Wang (1994).

improving stress-testing practices among financial institutions.[13] Based on my reading of the identified failures in stress-testing, let me put forward a complementary "five-point plan".

- **First, setting the stress scenario**. The key elements here are devising a multi-factor risk scenario that is sufficiently extreme to constitute a tail event. Arguably, designing such a scenario is better delegated to the authorities than to individual firms, in part because they ought to be more immune to disaster myopia. In its *Financial Stability Report (FSR)*, the Bank describes such stress scenarios. In future, the Bank aims to be able to offer through the *FSR* some greater clarity to financial firms about the sorts of vulnerability scenario it thinks they could use as one (and only one) input to their stress-testing machinery. This might include both solvency and liquidity-type scenarios. As the FSA have proposed, banks should also test to destruction their balance sheets through "reverse" stress tests, in order to identify potential areas of balance sheet weakness.

- **Second, regular evaluation of common stress scenarios**. Having banks conduct regular evaluations of their positions relative to a set of common scenarios (provided by the authorities) would be an improvement on current practices in several respects. First, it would allow some degree of benchmarking of results across institutions; second, it would allow a degree of benchmarking, and hence peer review, of models; and third, it would hopefully help in ensuring stress-testing exercises form an input to management decisions and are not an annual regulatory ritual. Comparing these bottom-up exercises with top-down evaluations conducted by the authorities – the like of which have appeared in recent Bank *FSR*s - can also help in benchmarking results and models.

- **Third, an assessment of the second-round effects of stress.** The results of these common stress evaluations should be the starting point, not the end point. These common stress tests need to be made dynamic, so that the second

[13] FSA (2008). See also Counterparty Risk Management Group (2008) for other useful suggestions.

and subsequent round interactions, and their consequences for system-wide risk, can be evaluated. This calls for an iterative approach to stress-testing in which banks' first-round results and management actions influence second-round stresses facing firms – for example, the effects of asset sales and liquidity hoarding. In effect, what we would then have is a hybrid stress test-cum-war game. This will better enable firms to assess the spillover and contagion consequences of their own and others' actions, so helping internalise to some degree the network externality problems which have been prevalent through this crisis. This dynamic, collective approach to stress-testing has already been attempted in one or two countries; it would be desirable if it became standard practice more widely. The recent Geneva report on financial regulation proposes greater use of such systemic stress testing. From the authorities' side, the Bank is developing a framework which will enable us to capture such network effects – for example, the effects of liquidity contagion and asset price disposals – on other firms in the network. The results from that framework could be used alongside firm-specific results to gauge network risks.[14]

- **Fourth, translation of results into firms' liquidity and capital planning**. The results from these exercises need to influence management outcomes if they are to be useful; the internal incentive problem needs to be overcome. So there should be a presumption that the results of these dynamic stress tests are taken, for example, to banks' risk committees. And banks' executives should periodically be asked how they intend to respond to these findings, including how effective their defensive responses are likely to be when the stress is system-wide and how the results affect liquidity and capital planning decisions.

- **Fifth, transparency to regulators and financial markets**. The bank-specific results ought to inform regulatory decisions about firms' capital and liquidity buffers. Indeed, there is a case for having these results set out regularly in firms' public reports. This would hopefully help exert a degree of market

[14] Aikman, Alessandri, Eklund, Gai, Kapadia, Martin, Mora, Sterne and Willison (2008).

discipline over management choices, as has been proposed by the Treasury Committee.[15] Existing disclosures by banks are a patchwork of different practices which make cross-firm comparisons of risk nigh on impossible. Having a standardised, published set of such stress-testing results would help improve financial markets' understanding and hence pricing of bank-specific risk – a particular problem during this crisis – thereby helping address the external incentive problem.

Working alongside the other Tripartite authorities, the Bank would be interested in exploring with financial firms the feasibility and desirability of putting this five-point plan into practice. This plan is about making stress-testing more robust but also more relevant. It is about providing that missing informational public good. In the arm-wrestle with management, it is about supplying power to the elbow of risk-managers.

Conclusion

Let me conclude. As after the previous two episodes of systemic failure, in October 1987 and August 1998, a third wave of technological transformation in the standards of risk management is now needed as a matter of priority. Firms themselves admit as much. That calls for a new agenda. I have outlined some elements of such an agenda, to address some of the failures exposed by the crisis. These measures involve a greater degree of engagement both between risk managers and senior management within firms, and between financial firms and the authorities. They would also involve much greater transparency to the wider world about risk metrics and accompanying management actions. These measures would not prevent a next time – nor should they – but they might help make risk management roughly right.

[15] House of Commons Treasury Committee (2008).

References

Adrian, T, and Brunnermeier, M.K (2008), 'CoVaR', *FRB of New York staff report no. 348*.

Aikman, D, Alessandri, P, Eklund, B, Gai, P, Kapadia, S, Martin, E, Mora, N, Sterne, G and Willison, M (2008), 'Funding Liquidity Risk in a Quantitative Model of Systemic Stability' paper presented at the 12th Annual Conference of the Central Bank of Chile on *Financial Stability, Monetary Policy and Central Banking* (Santiago, 6-7 November 2008).

Anderson H.C (1846), 'The Emperor's New Clothes', *Danish Fairy Legends and Tales* (1846), Pickering. London.

Blaschke, W, Jones, M.T., Majnoni, G and Martinez Peria, S (2001), 'Stress Testing of financial Systems: An overview of issues, methodologies, and FSAP experiences'. *IMF working paper WP/01/88*.

Brunnermeier, M, Crockett, A, Goodhart, C, Persaud, A and Shin, H (2009), 'The Fundamental Principles of Financial Regulation'**,** *ICMB-CEPR Geneva Report on the World Economy 11 (preliminary conference draft).*

Committee on the Global Financial System (2005), *Stress Testing at Major Financial Institutions: Survey Results and Practice,* Basel.

CPRMG (2008), *Containing systemic risk: the road to reform.*

Epstein, L.G and Wang, T (1994), 'Intertemporal Asset Pricing under Knightian Uncertainty', *Econometrica, Vol. 62, No. 2, pp. 283-322.*

Financial Services Authority (2008), *Stress and scenario testing: Consultation Paper 08/24.*

Gell-Mann, M (1994), *The Quark and the Jaguar: Adventures in the Simple and the Complex,* W. H. Freeman and Company, New York.

Guttentag, J.M. and Herring, R.J (1986a), 'Disaster Myopia in International Banking', *Princeton University Essays in International Finance, No. 164.*

Herring, R.J (1999), 'Credit risk and financial instability', *Oxford Review of Economic Policy 15:63-79.*

House of Commons Treasury Committee (2008), 'Financial Stability and Transparency, 2008: Sixth Report of Session 2007-08', *House of Commons Paper No. 371.*

Jorion, P (2006), *Value at Risk: The New Benchmark for Managing Financial Risk, 3rd Edition,* McGraw-Hill.

Kahneman, D, Slovic, P, & Tversky, A (1982), *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press

Kindleberger, C.P (1978), *Manias, Panics, and Crashes: A History of Financial Crises,* Palgrave Macmillan.

Knight, F.H (1921), *Risk, Uncertainty and Profit*, Houghton Mifflin Co.

KPMG/Economist Intelligence Unit (2008), *Never again? Risk management in banking beyond the credit crisis.*

Mackay C (1841), *Extraordinary Popular Delusions & the Madness of Crowds,* Wordsworth editions.

Morris, S, and Shin, H.S (2008), 'Financial Regulation in a System Context', *Brookings Papers on Economic Activity, mimeo.*

NYU Stern School of Business (2008), *Repairing the US Financial Architecture: An Independent View.*

Annex:

Chart 1: Probability density estimates for UK GDP Growth

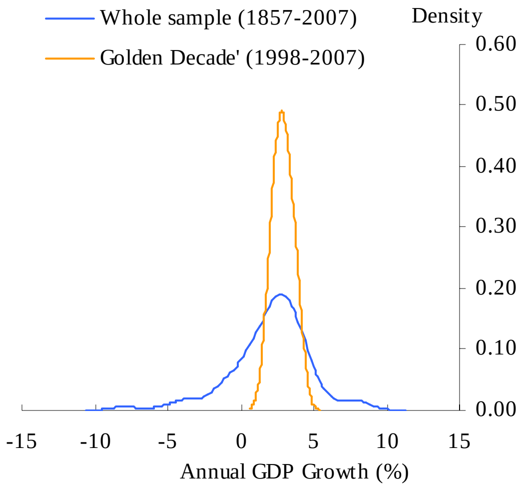


Chart 2: Probability density estimates for UK RPI

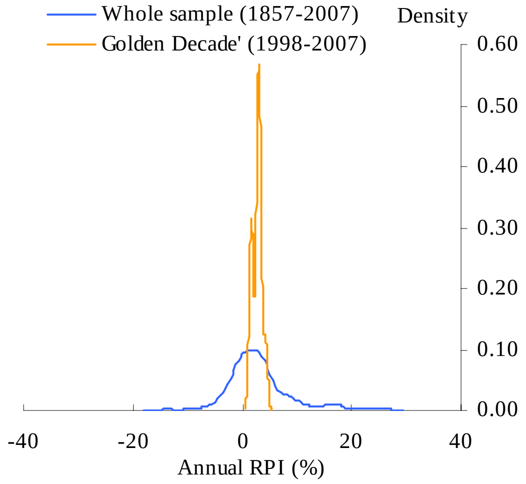


Chart 3: Probability density estimates for UK Unemployment

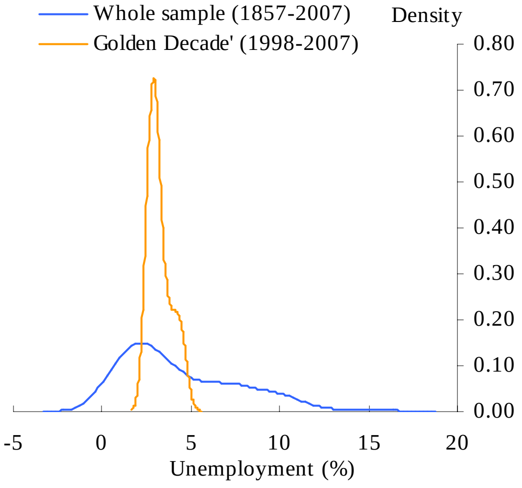


Chart 4: Probability density estimates for Annual Earnings Growth

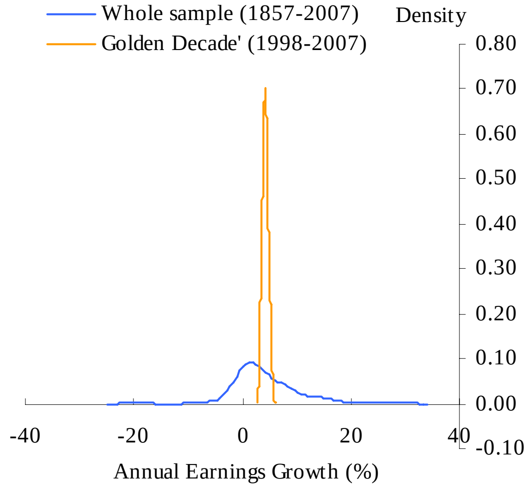


Chart 5: Probability density estimates for UK Base rate

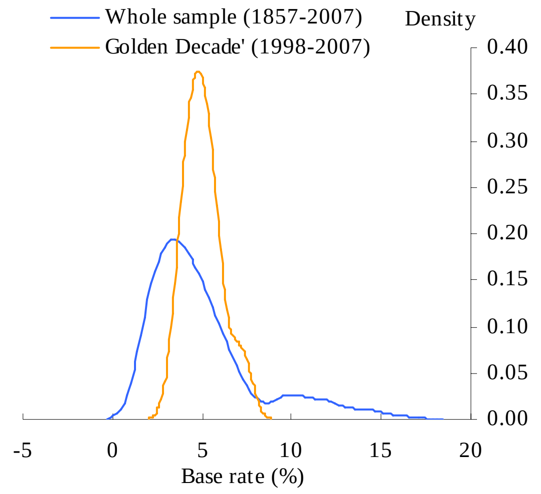


Chart 6: Probability density estimates for UK House Price Inflation

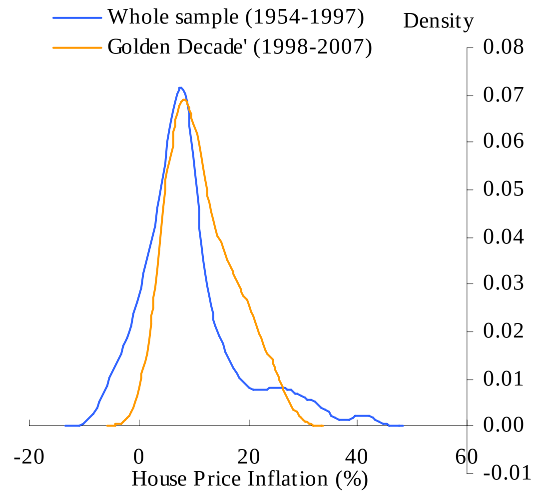
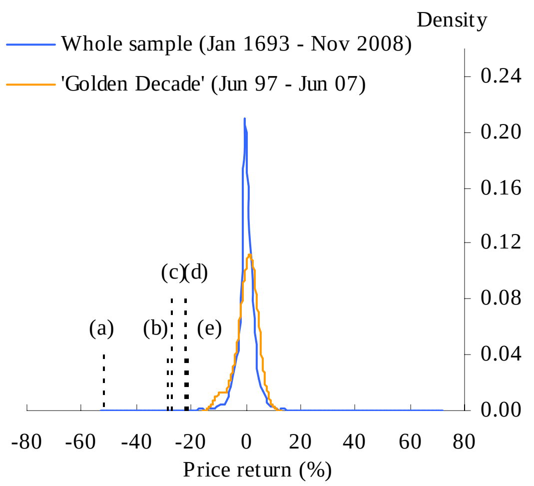
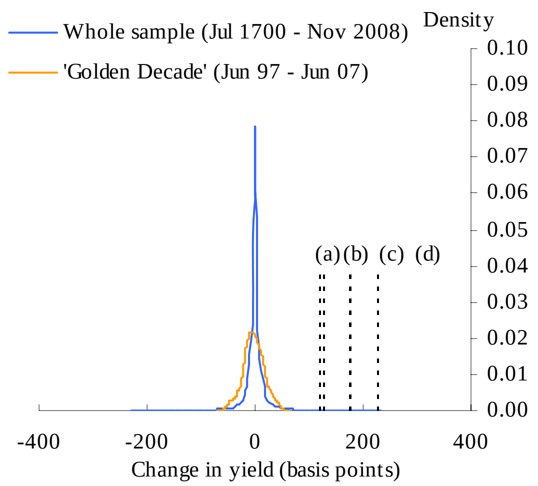


Chart 7: Probability density estimate for FTSE All-Share Index



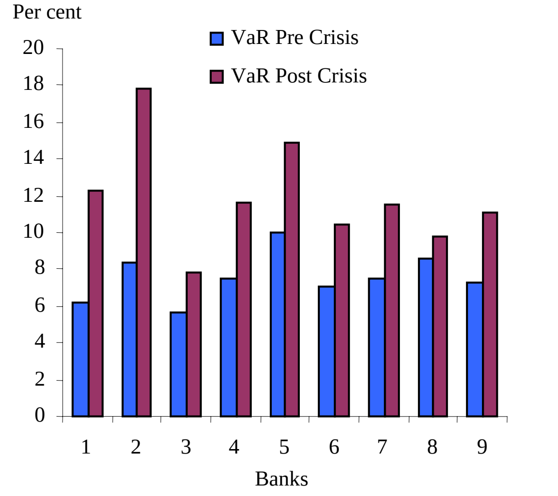
Sources: Global Financial Data and Bank calculations.
(a) September 1720 (South Sea Bubble)
(b) October 1720 (South Sea Bubble)
(c) October 1987 (Black Monday - portfolio insurance)
(d) July 1940 (WWII - merchant ships attacked)
(e) March 1974 (Price/wage controls, unions, etc.)

Chart 8: Probability density estimate for UK 2.5% coupon consol yield



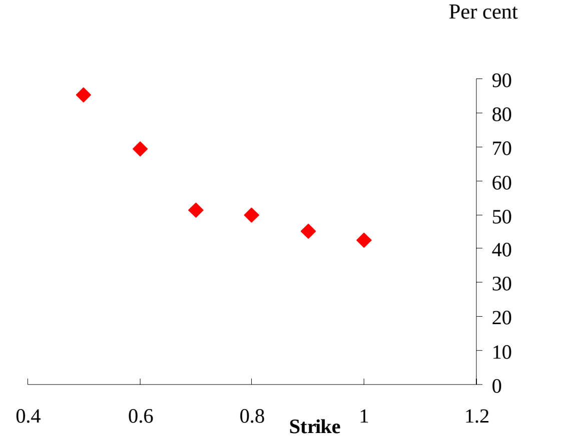
Sources: Global Financial Data and Bank calculations.
(a) March 1974
(b) October 1974
(c) June 1974
(d) January 1701

Chart 9: 90th Percentile VaR for a selection of major UK banks, pre and post crisis



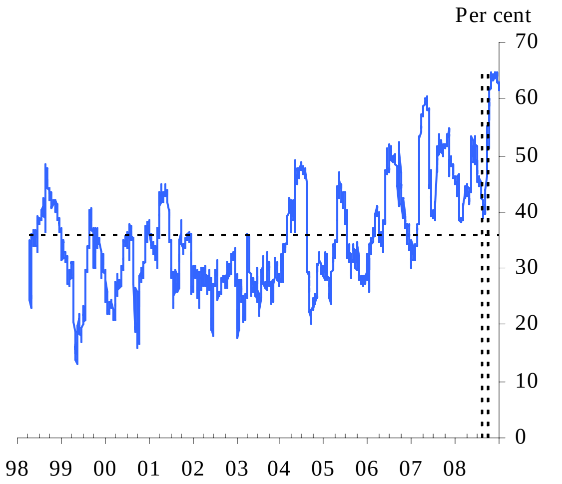
Sources: Bloomberg and Bank calculations

Chart 10: Percentage under-valuation of a put option on UK equities during the Golden Decade



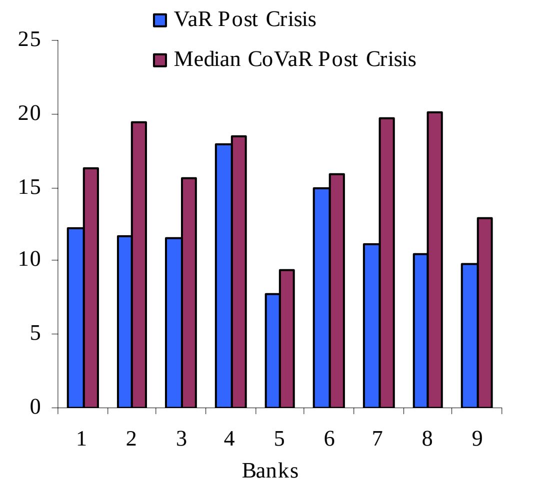
Sources: Global Financial Data and Bank calculations.

Chart 11: Common component in asset prices(a)



Sources: Bloomberg, Merrill Lynch, and Bank calculations.
(a) Proportion of variation in global equities, emerging market equities, high-yield spreads and commodities explained by a common component over a three-month rolling window.

Chart 12: VaR vs CoVaR for a selection of major UK banks, post crisis



Sources: Bloomberg and Bank calculations

Table 1: Distribution of UK Macroeconomic and Financial Time Series

		Golden Decade (1998-2007)	Long-run [(a)]
GDP Growth (%, annualised)	**Mean**	2.9	2.0
	Standard Deviation	0.6	2.7
	Skew	0.2	-0.8
	Kurtosis	-0.8	2.2
Retail Price Inflation (%, annualised)	**Mean**	2.8	3.1
	Standard Deviation	0.9	5.9
	Skew	0.0	1.2
	Kurtosis	-0.3	3.0
Unemployment rate (%, annualised)	**Mean**	3.2	4.4
	Standard Deviation	0.6	3.4
	Skew	1.2	1.0
	Kurtosis	0.4	0.3
Earnings Growth (%, annualised)	**Mean**	4.2	4.5
	Standard Deviation	0.5	6.4
	Skew	0.2	1.1
	Kurtosis	-0.1	3.8
House Price Inflation (%, annualised)	**Mean**	11.6	8.8
	Standard Deviation	5.8	9.2
	Skew	0.7	1.1
	Kurtosis	-0.4	2.0
Change in FTSE All-Share Index (%, monthly)	**Mean**	0.2	0.2
	Standard Deviation	4.1	4.1
	Skew	-0.8	2.6
	Kurtosis	3.8	62.3
Change in UK 2.5% coupon consol yield (basis points, monthly)	**Mean**	-2.0	-0.1
	Standard Deviation	18.5	18.4
	Skew	0.0	0.5
	Kurtosis	3.1	32.2

(a) The long-run time series for GDP growth, retail price inflation, unemployment, and earnings growth begin in 1857. The house price inflation series begins in 1954. The FSTE-All Share Index series begins in Jan 1693. The 2.5% coupon consol yield series begins in July 1700. All data are for the UK.