Stress Testing and Macroprudential Regulation

A Transatlantic Assessment

Edited by Ronald W. Anderson

Systemic Risk Centre

Financial Markets Group
RESEARCH CENTRE

CEPR Press

Stress Testing and Macroprudential Regulation:

A Transatlantic Assessment

CEPR Press

Centre for Economic Policy Research
33 Great Sutton Street
London, EC1V 0DX
UK

Tel: +44 (0)20 7183 8801
Email: cepr@cepr.org
Web: www.cepr.org

ISBN: 978-1-907142-98-7

Stress Testing and Macroprudential Regulation:

A Transatlantic Assessment

Edited by Ronald W. Anderson

CEPR Press

Centre for Economic Policy Research (CEPR)

The Centre for Economic Policy Research (CEPR) is a network of over 1,000 research economists based mostly in European universities. The Centre's goal is twofold: to promote world-class research, and to get the policy-relevant results into the hands of key decision-makers. CEPR's guiding principle is 'Research excellence with policy relevance'. A registered charity since it was founded in 1983, CEPR is independent of all public and private interest groups. It takes no institutional stand on economic policy matters and its core funding comes from its Institutional Members and sales of publications. Because it draws on such a large network of researchers, its output reflects a broad spectrum of individual viewpoints as well as perspectives drawn from civil society.

CEPR research may include views on policy, but the Trustees of the Centre do not give prior review to its publications. The opinions expressed in this report are those of the authors and not those of CEPR.

Financial Markets Group Research Centre (FMG)

The Financial Markets Group Research Centre (FMG) was established in 1987 at the LSE. The FMG is a leading centre in Europe for policy research into financial markets.

It is the focal point of the LSE's research communication with the business, policy making and international finance communities. The FMG works alongside the Department of Finance to understand problems in financial markets and in the decision-making processes of corporations, banks and regulators.

Systemic Risk Centre (SRC)

The Systemic Risk Centre (SRC) was set up in 2013 to study the risks that may trigger the next financial crisis and to develop tools to help policymakers and financial institutions become better prepared. Based at the London School of Economics and Political Science (LSE), the Centre is generously funded by the Economic and Social Research Council (ESRC) [grant number ES/K002309/1].

The primary goal of the SRC is to build up research findings across the broad area of systemic risk and to use these to construct practical tools for policymakers and private institutions to achieve a better understanding of the risks they face. The unifying principle of the Centre's agenda is endogenous risk – the notion that financial risk is created by the interaction of market participants.

The SRC is rooted in a multidisciplinary approach, bringing together experts from computer science, law, political science and the natural and mathematical sciences, as well as from finance and economics. This enables our researchers to investigate how risk is created through feedback loops within and between the financial, economic, legal and political systems.

The SRC runs frequent conferences, seminars, masterclasses and other events to facilitate and enhance collaborations and exchange of ideas.

The Directors of the SRC are Jon Danielsson and Jean-Pierre Zigrand.

Contents

About the contributors

Viral V. Acharya is the C.V. Starr Professor of Economics in the Department of Finance at New York University Stern School of Business (NYU-Stern). He is a member of the Economic Advisory Committee of the Financial Industry Regulation Authority (FINRA), International Advisory Board of the Securities and Exchange Board of India (SEBI), Advisory Council of the Bombay (Mumbai) Stock Exchange (BSE) Training Institute, and Academic Research Council Member of the Center for Advanced Financial Research And Learning (CAFRAL, India). He is also an Academic Advisor to the Federal Reserve Banks of Cleveland and New York, and in the past, also to the Federal Reserve Banks of Chicago, Philadelphia and the Board of Governors, and to the European Systemic Risk Board.

Ron Anderson is a Professor of Finance at the London School of Economics and a member of the the Financial Markets Group at LSE. He is a Research Fellow at CEPR in the Financial Economics programme and a principal investigator in the Systemic Risk Centre, an ESRC research centre. He has broad interest in the field of financial economics. In recent years much of his theoretical work has centred upon dynamic corporate finance. Ron has an active interest in financial institutions. In this field, he carried out a study of finance in ex-communist countries (*Transition Banking: The Financial Development of Central and Eastern Europe*, co-authored with Chantal Kegels, Oxford University Press). More recently, he has studied how financial reforms since the Global Crisis is transforming the structure of the financial sector. Ron has undertaken consultancy and research projects in corporate finance, corporate governance, credit risk, and risk management. His recent executive teaching has been in the areas of risk management and corporate governance. He has been involved in expert testimony on a variety of capital markets cases.

Alex Brazier is Executive Director for Financial Stability Strategy and Risk at the Bank of England. Alex joined the Bank in 2001 as an economist, and has held a variety of increasingly senior positions over the course of his career, most recently as Principal Private Secretary to the Governor since November 2011. Prior to that, Alex served as Manager of the Bank's UK Forecasting Team from 2008-2011, as Economic Assistant and Speechwriter to the Governor from 2005-2008, and as a UK Economic Analyst in

the Bank's Monetary Analysis Directorate from 2001-2005. Alex earned an MSc in Economics from the University of Warwick in 2000, and an MA in Economics from Cambridge University in 1999.

Vítor Constâncio is Vice-President of the European Central Bank since 1 June 2010. He was Governor of the Banco de Portugal from 1985 to 1986 and from February 2000 to May 2010. He graduated in Economics from the Universidade Técnica de Lisboa. Before having been appointed as Governor of the Banco de Portugal, he was Executive Director of the Banco Português de Investimento (1995-2000) and non-executive Director of the Electricidade de Portugal (EDP), the Portuguese national power utility (1998-2000). He was Invited Senior Professor of Economics at the Instituto Superior de Economia e Gestão from 1989 to June 2010, culminating a long academic career.

Udaibir S. Das is Assistant Director and Division Chief at the IMF's Monetary and Capital Markets Department. Mr. Das was at the BIS between 2012–2015, where he worked on Basel III implementation and analytical studies on risk weighting of assets, and helped set up the Regulatory Consistency Assessment Program (RCAP). At the IMF, he manages the IMF's Financial Sector Assessment Program (FSAP) worldwide and oversees the development of IMF policies and tools for financial system surveillance, including risk analysis, stress testing, and systemic risk monitoring. His hands-on policy and operational experience spans 30 years across advanced, emerging and frontier economies. Widely published, Mr. Das' recent publications include those on sovereign risk, debt restructuring, public balance sheet risks, debt capital markets, sovereign wealth funds and Basel III implementation. Mr. Das has also co-authored two books.

Rochelle Edge is a Deputy Associate Director in the Office of Financial Stability Policy and Research at the Federal Reserve Board (FRB) in Washington, DC. In this position, she has led teams to design the macroeconomic scenarios used in the supervisory Comprehensive Capital Adequacy and Review (CCAR) and Dodd-Frank stress tests; to operationalize the Basel III countercyclical capital buffer in the United States; and to study interest rate risk at banks. She currently represents the FRB on the Financial Stability Board's Analytic Group on Vulnerabilities and has also represented the FRB on several macroprudential policy implementation working groups for the Committee on the Global Financial System. Dr. Edge's current research interests are in the areas

of financial stability, banking, and macroeconomics. Her papers have been published in the *Journal of Monetary Economics, Review of Economics and Statistics, Review of Economic Dynamics, Journal of Economic Dynamics and Control, Journal of Applied Econometrics,* and *Brookings Papers on Economic Activity*.

Charles Goodhart, CBE, FBA is Emeritus Professor of Banking and Finance with the Financial Markets Group at the London School of Economics, having previously, 1987-2005, been its Deputy Director. Until his retirement in 2002, he had been the Norman Sosnow Professor of Banking and Finance at LSE since 1985. Before then, he had worked at the Bank of England for seventeen years as a monetary adviser, becoming a Chief Adviser in 1980. In 1997 he was appointed one of the outside independent members of the Bank of England's new Monetary Policy Committee until May 2000. Earlier he had taught at Cambridge and LSE. Besides numerous articles, he has written a couple of books on monetary history; a graduate monetary textbook, *Money, Information and Uncertainty* (2nd Ed. 1989); two collections of papers on monetary policy, *Monetary Theory and Practice* (1984) and *The Central Bank and The Financial System* (1995); and a number of books and articles on Financial Stability, on which subject he was Adviser to the Governor of the Bank of England, 2002-2004, and numerous other studies relating to financial markets and to monetary policy and history. His latest books include *The Basel Committee on Banking Supervision: A History of the Early Years, 1974-1997* (2011), and *The Regulatory Response to the Financial Crisis* (2009).

Thomas F. Huertas is a Partner in EY's Financial Services Risk Advisory practice and he chairs its Global Regulatory Network. He joined EY in 2012 with over 35 years' experience as a regulator, business executive and banker in Europe and North America. As a regulator, Dr Huertas served as a member of the Executive Committee at the UK Financial Services Authority (FSA). During his time at the FSA, he was responsible for the prudential and conduct supervision of wholesale firms in the banking, insurance and asset management sectors in the UK. He also led the FSA's work on recovery and resolution, remuneration and the analysis of risks facing the banking sector. In addition, he served as a member of the Basel Committee and as Alternate Chair of the European Banking Authority. Prior to joining the FSA, he held senior positions at Citigroup in the

US and Europe. Dr Huertas has published extensively on banking and financial issues, including his recent book, *Safe to Fail: How Resolution Will Revolutionise Banking*.

Philipp Hartmann is Deputy Director General of the research department at the European Central Bank, which he helped building up from its start. He also coordinates the ECB's work on financial integration and is a Fellow of the Centre for Economic Policy Research. Mr Hartmann published research on financial, monetary and international issues in several books and numerous journal articles. He serves as an associate editor of the *Journal of Financial Stability*. His policy work has been published in many official reports and discussed in fora including the ECOFIN Council, the ECB Governing Council and the Basel Committee on Banking Supervision.

Andreas Lehnert is Deputy Director of the Office of Financial Stability Policy and Research at the Federal Reserve Board in Washington, DC. Andreas joined the Fed after earning his PhD in economics from the University of Chicago. He started in the household finance research group where he worked on a variety of topics in consumer and mortgage credit. During the financial crisis, he contributed to several projects including various mortgage modification initiatives, the TARP, the 2009 bank stress tests and the TALF. In November 2010 he moved to the Board's newly created financial stability group where he participates in a variety of ongoing initiatives to promote financial stability, including regulatory reform, the periodic assessment of financial vulnerabilities, the development of macroprudential tools and the design and oversight of the bank stress tests; in addition, he supports the Federal Reserve's role on the Financial Stability Oversight Council and the Financial Stability Board. His research focuses on financial stability, macroprudential policy, banking, and finance.

Sascha Steffen is Professor of Finance and the Chair of Financial Markets at the University of Mannheim. He is also the head of the research group "International Finance and Financial Management" at the Center of European Economic Research (ZEW) in Mannheim. Previously he was an assistant professor at the University of Mannheim. From 2008 to 2009 Sascha was a visiting scholar at the Stern School of Business at NYU and ESMT. He received his doctorate degree in finance with honors (summa cum laude) from the Goethe University Frankfurt in 2007. His research is in the area of banking and financial intermediation and has been published in leading finance

journals such as the *Journal of Finance, Journal of Financial Economics* and *Review of Financial Studies* and has received several Best Paper Awards. His research has been supported (among others) by the FDIC and the Deutsche Forschungsgemeinschaft (DFG) and the Lamfalussy Fellowship program.

Til Schuermann is a partner in the management consulting firm Oliver Wyman where he advises financial institutions on stress testing and capital planning. Formerly, he was Senior Vice President in Research and Bank Supervision at the Federal Reserve Bank of New York. He is also a Research Fellow at the Wharton Financial Institutions Center and an associate editor for the *Journal of Risk* and the *Journal of Financial Services Research*.

Foreword

Assembling work from an expert group of specialists, this eBook presents various stress testing methods and an overview of macroprudential regulation for banking supervision. It brings together the main policy presentations from a two-day conference organised by Ron Anderson at the London School of Economics in October 2015, which was co-sponsored by CEPR, the Systematic Risk Centre and the Financial Markets Group.

Academics, policymakers, regulators, supervisors and financial analysts discussed stress testing methodologies over two days of speeches, panels and presentations. This study collects the main policy presentations from this conference in which the authors focused on stress testing experiences within the UK and US; its role in macroprudential and supervision policy; and its overall effectiveness in both historical and current markets.

Since the Global Crisis of 2007-08, stress testing for banks has become a binding constraint in terms of the way in which they are regulated. With this in mind, stress testing has shifted to the forefront of the media's coverage on the banking industry, making it essential to have constructive discussions on how best to implement these policies. This new eBook facilitates a debate on some of the major issues in this extremely important area.

CEPR is grateful to Professor Ron Anderson for taking editorship of this eBook. Our thanks also go to Simran Bola and Anil Shamdasani for the excellent and swift handling of its production. CEPR, which takes no institutional positions on economic policy matters, is delighted to provide a platform for an exchange of views on this topic.

Tessa Ogden
Deputy Director, CEPR
April 2016

1 Stress testing and macroprudential regulation: A transatlantic assessment

Ronald W. Anderson[1]
London School of Economics and CEPR

Introduction: Stress testing then and now

Since the onset of the financial crisis in 2007-2008, stress testing has emerged as a major component of the supervisory toolkit. For most of the large global banks in the US and Europe, meeting the standard to pass their annual supervisory stress tests is *the* binding regulatory constraint. So the outcome of banks' stress tests is headline news these days. This has been a remarkable development for a tool which, ten years ago, was little known apart from among a small fraternity of banks' risk modellers and their supervisory counterparts.

A conference held at the London School of Economics in October 2015 assembled an expert group of specialists to take stock of how stress testing has developed. Academics, policymakers, regulators, supervisors, and financial sector analysts drawn from both sides of the Atlantic discussed stress testing experiences and methodologies over two days of speeches, panels, roundtables, and presentations of recent academic research. This volume collects the main policy presentations from this conference, and this first chapter gives an overview of the main themes that emerged, thereby serving

1 I would like to thank Charles Goodhart and Malcolm Knight for detailed comments on an earlier draft. Responsibility for views expressed and any errors is my own. The support of the Economic and Social Research Council (ESRC) in funding the SRC is gratefully acknowledged (grant number ES/K002309/1).

as a statement of the current state of the art of stress testing for bank supervision and macroprudential regulation.

Financial stress testing emerged in the 1990s as a risk management tool for banks and other financial firms. Stress testing is a forward-looking, quantitative estimation of the losses that would likely occur in a portfolio or financial institution if it were exposed to very adverse conditions in the future. Of course, planners – including financial planners – have long used techniques of scenario analysis to take precautionary steps against bad events occurring. However, in the 1990s improved data and new modelling techniques such as value at risk (VaR) combined with greater computer power to provide bank analysts much greater scope to make quantitative assessments. In some banks, risk managers began to carry out stress tests to better understand the possible consequences of shocks to important risk factors, especially if it appeared plausible that such risk factors might stray outside normal ranges based on past experience. Furthermore, stress testing was sometimes incorporated into banks' own economic capital models and otherwise used in the capital planning process.

Stressing testing also began to be incorporated into banking regulation and supervision. Under Basel II's rules for calculating regulatory capital (Pillar 1), stress tests are included in adjusting probabilities of default for certain credit exposures and for assuring the robustness of parameterisations in some internal models. Stress tests are also included in Pillar 2 as part of the overall supervisory review of a bank's ability to withstand adverse changes in market conditions. However, within the Basel II framework there are no clear guidelines as to how adverse scenarios should be chosen, thus leaving this to the banks' and supervisors' judgement.

The fact that in drafting Basel II, policymakers felt unable to take a strong stand on what is the proper way to formulate adverse scenarios for stress tests indicates a basic difficulty involved with the methodology – that there is no easy way to set the standard of what is a 'reasonable' level of adversity. As any new chief risk officer quickly understands, by setting the bar of severity of stress tests too high, he will quickly burn through his stock of political capital within the firm should the future prove benign and traders can point to the attractive profits that have been foregone because of excessive caution.

The lack of a settled standard for choosing the right scenario is a widely recognised drawback of stress testing. But at the same time, the flexibility in exploring a wide range of possible scenarios is a major strength of the methodology. Under the Basel II internal models approach, regulatory capital calculations use recent historical data. Numerous commentators have pointed out that this tends to result in procyclical risk assessments. Thus following a period of recession-free, low-volatility growth, risk is assessed as being low, resulting in relatively low regulatory capital and high leverage. Stress testing potentially can overcome this tendency by allowing analysts to impose relatively severely adverse scenarios after sustained booms and the opposite when in the midst of a recession.

History of macro stress testing since the crisis

Thus even in the immediate aftermath of the agreement of the Basel II guidelines and while these were being implemented by national authorities, Basel working groups continued to explore the possible ways that stress testing methods could be used to improve the framework.[2] Later, as the crisis began to unfold in 2007, interest in stress testing grew more pressing. This gave rise to the Basel Committee's *Principles of Sound Stress Testing Practices and Supervision* (BCBS 155) in 2009. In the EU, the weaknesses of past stress testing practices were also highlighted in the 'de Larosière Report', which called for testing under more adverse scenarios (de Larosière Group 2009). Furthermore, calls for greater reliance on stress testing also could be heard from the banking industry itself (IIF 2009). Ultimately, these initiatives did bear some fruit through the Basel process. In Basel III, the use of stress tests for some asset classes is strengthened in Pillar 1 calculations of regulatory capital, and stress testing receives increased emphasis in the Pillar 2 review of a bank's internal capital adequacy assessment process (ICAAP).

Nevertheless, this did not amount to a clear mandate to make stress testing central to banking regulation and the supervisory process. The real game-changer came with

2 See the studies in this Basel Committee work stream reported in a special issue of the *International Journal of Central Banking*, as summarised by Van Lelyveld (2009).

the decision in the US to use stress testing as a major component of the economic recovery programme implemented by the incoming Obama administration in 2009. This choice of the tool intended to restore financial stability in the US was not obvious, given the previous experience with stress testing. Stress testing as used in the Basel rules for regulatory capital had been insufficient to protect banks from falling into grave difficulties in 2007 and 2008. Furthermore, stress testing was part of the regulatory framework for Fannie Mae and Freddy Mac, and these institutions were effectively insolvent in 2008 (Hirtle and Lehnert 2015).

In his personal account of the crisis, Timothy Geithner gives insight into to how the decision to adopt macro stress testing came about (Geithner 2014). Following Obama's election in November 2008, Geithner had been selected as the incoming Secretary of the Treasury. He and Larry Summers, who had been chosen to head up the President's National Economic Council, were charged with developing a programme of economic recovery, and in this the main responsibility of deciding how to restore stability to the financial system fell to Geithner. He had served as the president of the Federal Reserve Bank of New York since 2003 and had been deeply involved in dealing with the Lehman Brothers collapse and the subsequent attempts to contain the raging financial crisis. In this, he had worked closely with Hank Paulson, his predecessor at the Treasury, and was well acquainted with the difficulties involved in getting Congress to agree the Troubled Assets Relief Program (TARP) and other parts of the government's recovery programme. Despite the huge amounts of funds that had been mobilised to plug solvency holes in Fannie Mae and Freddie Mac, in the large banks, and in AIG, Geithner could see from bank equity prices and credit spreads that the market did not believe that the government had done enough to avoid a wholesale collapse of the financial system. Given this backdrop, in a telephone call over the Christmas holiday break, Geithner put forward to Summers the idea that a concurrent, forward-looking valuation of the assets of the major banks under stressed conditions could serve to assure the markets and the public generally that the government had a realistic assessment of the scale of the problem and would respond to it so as to avoid general failure of the banking system. This idea that macro stress tests should be a central part of the government's strategy was debated within the Obama economic policy team and Summers, among others, voiced doubts as to whether the markets would find the tests credible. A further

criticism was that the tests would take too long to execute and that the market was demanding more decisive, immediate action.

The main alternative to Geithner's stress test proposal was a pre-emptive infusion of federal funds to recapitalise the largest banks, effectively nationalising them. In the end, this alternative was deemed to be too politically costly to be feasible. When the plan for stress tests was announced in late February 2009, few details were given and the markets reacted negatively. Subsequently, the administration made it clear that banks without adequate capital under the adverse stress test scenario would be required to raise capital and, importantly, if they failed to raise capital in the markets, the federal government stood ready to inject capital. Another crucial feature that was decided in the course of carrying out the tests was to provide detailed information to the public both about the methodology used and about the specific valuations obtained for individual banks. When the results were announced in early May 2009, it was found that ten firms would need to raise a combined total of $75 billion of fresh capital. Ultimately, most of this amount was raised from the markets.[3]

The reaction to the stress tests in May 2009 was generally very positive. And this view of the tests, officially called the Supervisory Capital Assessment Program (SCAP), has largely been upheld in subsequent assessments. The markets reacted positively to the stress tests, and initial critics of the test, including those advocating pre-emptive nationalisation of banks, expressed their surprise and relief that stabilisation of the banking sector was done without more public funds. In reviewing Geithner's book, *Stress Test,* Paul Krugman, who was among the vocal 'nationalisers' in 2009, writes:

> *"A principal part of Geithner's argument against nationalization was the belief that a 'stress test' of banks would show them to be in fairly decent shape and that publishing the results…would, in conjunction with promises to shore up banks when necessary, end the crisis. And so it proved. He was right; I was wrong; and the triumph of the stress test gave him the title for his book."* (Krugman 2014)

3 The main exception was GMAC, which required a $7.5 billion infusion of capital from the US Treasury.

Furthermore, the use of macro stress testing in 2009 is viewed as having made a valuable and lasting contribution to the practice of banking supervision. In a 2013 speech, Ben Bernanke stated:

> *"In retrospect, the SCAP stands out for me as one of the critical turning points in the financial crisis. It provided anxious investors with something they craved: credible information about prospective losses at banks. Supervisors' public disclosure of the stress test results helped restore confidence in the banking system and enabled its successful recapitalization. The resilience of the U.S. banking system has greatly improved since then, and the more intensive use and greater sophistication of supervisory stress testing, as well as supervisors' increased emphasis on the effectiveness of banks' own capital planning processes, deserve some credit for that improvement."* (Bernanke 2013)

While he does not emphasise this in his book, one factor that may have entered into Geithner's decision to push for using stress testing was that he was familiar with the methodology from his time at the Federal Reserve Bank of New York. Indeed, Fed economists provided him with rough estimates of the possible scale of solvency problems in the banking system by doing some high level, 'top-down' calculations of a macro model under stressed conditions.

Building on the experience in SCAP, stress testing has become central to the supervision of systemically important institutions – including some non-banks – in the US, and the methodology employed has been extended and modified in many respects. These developments are described in the contribution to this volume by Rochelle Edge and Andreas Lehnert (Edge and Lehnert 2016). Among the highlights is that banks' capital planning continues to be the major focus, notably in the Comprehensive Capital Assessment Review (CCAR). Another is that a second, parallel stress test (DFAST) is run as required by the Dodd-Frank Act of 2010. The two tests have much in common. However, one important difference is that they make different assumptions about a bank's capital structure, with DFAST based on the current capital structure and CCAR based on projecting capital using the bank's proposed capital plan (including new issues, dividends and share buy-back). In addition, following a DFAST report there

is no supervisory action that follows automatically, whereas following CCAR the supervisor states whether the bank's capital plan is approved or not. That is, CCAR is the test that banks either pass or fail. A further notable development is that in CCAR, a bank may fail the test on qualitative grounds based on the supervisor's assessment of the bank's ability to produce clear and convincing projections under the stresses specified in the test. Finally, an important development is that the Federal Reserve has increasingly relied upon its own internal supervisory models. The effect of this is that banks are free to use their own internal models to evaluate stresses, but the Fed will do parallel calculations of its own. If the two sets of calculations differ substantially, the bank may find that it fails the test.

Starting in July 2007 with the disclosure of subprime losses in some German *Landesbanken,* and increasingly through 2008, it was clear that the same problems that were felling US banking giants were also undermining the solvency of European institutions. So it was natural to ask what was the extent of the damage to the European banking system. In May 2009, very shortly after the SCAP results came out in the US, the EU announced it was undertaking stress tests of the European banking system.

This coincidence in timing makes a comparison with US stress tests inevitable. However, the comparison is somewhat unfair. The institutional framework of the EU is very different from that of the US. Whereas in the US the stress tests were being carried out by the Federal Reserve, a long-established, strong institution that is the authorised supervisor of the major banks across the US, in Europe the 2009 stress test was the responsibility of the Committee of European Bank Supervisors (CEBS), an advisory body with no direct authority to supervise any bank. When the results of these first tests came out, it was announced that under the adverse economic scenario, '…the aggregate Tier 1 ratio for the banks in the sample would remain above 8% and no bank would see its Tier 1 ratio falling under 6%'.[4] There was very little explanation given about the methodology used. And CEBS made no comment on any individual bank, stating instead that this would fall to the competent national authority to do so if they wished. In fact, almost none did.

4 "CEBS Press Release on the Results of the EU-Wide Stress Testing Exercise", 1 October 2009 (http://www.eba.europa.eu/risk-analysis-and-data/eu-wide-stress-testing/2009).

Given this lack of transparency or of any linkage to identifiable supervisory actions, it was not surprising that the 2009 EU stress test was largely discounted. However, as problems in Irish banks came to light shortly after the 2009 results were announced, doubts about the general health of the European banking system only became more pronounced. In 2010 CEBS was again charged with organising a stress test, however, this time with more direct operational support from the European Central Bank (ECB). The tests covered 91 banks from 20 EU member states both inside and outside of the Eurozone. Only seven banks failed, and the overall stress test showed that in aggregate, the required Tier 1 capital in states covered could be about 9.5% under the adverse scenario, well above the 6% threshold set for the exercise. Much more detail about the methodology was given, and the exercise also included an attempt to take into account sovereign risk, notably through an application of haircuts to the sovereign debt holdings of banks. Nevertheless, the scenarios used were widely interpreted as being too rosy, and following the publication of the 2010 test, the credit spreads on European periphery sovereign debt widened.

In 2011, a third round of European stress tests was carried out, this time under the new European Banking Authority (EBA), which had replaced CEBS and which, in principle, had greater power to push national supervisors to take specific actions. The results were announced in July 2011, and eight banks were found to have fallen short of the 5% core Tier 1 equity capital (CET1) hurdle set in the exercise. The EBA specifically required these banks' responsible national authorities to take corrective actions to bring them into line with the standard. Furthermore, the stress testing methodology and results were explained in considerable detail. Despite these tangible efforts taken by the EBA to increase the credibility of European stress tests, they soon were overtaken by the events in the worsening sovereign debt crisis, notably the *de facto* failure of DEXIA, a bank which had passed the 2011 test.

Manifestly, European macro stress testing had not proved to be the tool that policymakers needed in order to build confidence and end the banking crisis in Europe. Why was this experience so different from that of the US? One explanation is that it was a problem of 'too little too late'. Even though transparency and severity were increased in each successive round of tests, they continually fell short of being credible. For this reason, there were no system-wide tests in 2012 and 2013, giving policymakers time to develop

a plan for a more credible stress testing regime. This came only after the EU agreed the single supervisory mechanism (SSM), in which the ECB is given a direct role in the supervision of systemically significant banks within the Eurozone and, potentially, some other EU member states. In preparation for the start of operations of the SSM in early 2015, it was agreed that a comprehensive assessment would be carried out of the major European banks, including all banks that were to be supervised by the SSM as well as some large banks from outside the Eurozone that were not sure to be supervised by the SSM. This would involve two parts: a) an asset quality review (AQR) involving a thorough and detailed audit of bank exposures and the adequacy of collateral and provisioning; and (b) a stress test to examine on a forward-looking basis the ability of the bank to withstand macroeconomic stresses.

The purpose was to put to rest the doubts about the solvency of major European banks and, in the process, to establish the credibility of European stress testing. In 2014, in parallel with the AQR, European banks raised some €50 billion of equity capital (CET1) in order to strengthen their balance sheets. The 2014 stress tests covered 123 banks accounting for 70% of European banking assets. The stress test methodology was thoroughly explained, and results for individual banks were reported in detail. In the end, 23 banks failed to meet the required standard of 5.5% CET1 under the stress scenario, and these banks were required to take remedial actions with their competent supervisor, which is the ECB in the case of those banks within the SSM.

The reaction of the financial markets following the release of the results of the stress test in October 2014 suggests that the tests were viewed as informative and credible. Equity prices and credit default swaps spreads moved substantially – improving for banks that were found to be healthy, and deteriorating for those that looked weak under the adverse scenario. Given the extreme scepticism about European stress testing following the first three rounds of tests, the 2014 exercise has not silenced all critics. However, the results have convinced policymakers to continue to use the basic framework of the comprehensive assessment, including an AQR and stress test annually on an ongoing basis.

Lessons

This brief summary of the stress testing experiences in the US and EU highlights two of the main themes that were discussed in our LSE stress testing conference in October 2015. These are the *transparency* of the stress tests and the need for a *credible backstop* to deal with capital shortfalls revealed by the tests.

Much of the discussion of stress testing in the financial press has focused on the transparency of such tests, or the lack thereof. However, transparent stress testing in the midst of a crisis will not necessarily restore financial stability. The case of Cyprus illustrates this point.[5] In 2011, because of a fear of contagion from the Greek fiscal crisis, confidence in the Cypriot banking system began to erode, and as part of negotiations with the EU and the IMF it was agreed that stress tests would be carried out. The tests were eventually carried out in mid-2012, using a well-known fund management firm. The results reported a €5.8 billion a loss, of which €4.5 billion was the consequence of applying a 79% haircut of bank holdings of Greek government bonds. The Greek bond write-down represented some 25% of GDP. There was obviously no way that banks could be recapitalised either from international markets or from the Cypriot government, which was already running a very large primary deficit. Ultimately, a bailout package was agreed with the EU and the IMF in which the Cypriot part, set at €5.8 billion, came in the form of bail-in of Cypriot bank creditors, including some insured depositors.

Thus transparency is not enough. In itself, a transparent stress test can lead to a bank run if it reveals a solvency hole that nobody knows how to fill. This accounts for the difficult time European authorities had in establishing credibility in the stress tests between 2009 and 2011. They could not put to rest the scepticism of the markets and the press, no matter how much technical detail was given about stress test methodology or of individual bank results because no one could explain how banks would be recapitalised if very large losses were found. This was particularly the problem at the height of the sovereign debt crisis in 2011, when the hazard that everyone feared most, namely Greece falling out of the Eurozone, was not part of the stress scenario.

5 The following discussion is based on Michelides (2014).

The AQR of 2014 took place after the financial conditions had improved relative to the depths of the sovereign debt crisis. And the EU had taken measures to be able to deal with large-scale solvency problems should they emerge in the future, notably through the European Recovery and Resolution Directive and the Single Resolution Mechanism (SRM). Whether this will suffice to assure the credibility of future European stress tests is uncertain. The question may only have an answer once the SRM itself is tested with a large bank insolvency which it satisfactorily resolves.[6]

Another theme that was discussed repeatedly at the LSE conference was *whether stress testing by the public authorities is a tool of microprudential or macroprudential regulation*. From the discussion above, it is clear that since SCAP in 2009, stress testing of large banks in the US is intended to serve as an instrument of *both* microprudential and macroprudential policy. Furthermore, I have argued (and this same view was a recurrent theme in the LSE conference) that the credibility of stress tests is enhanced if these two functions are linked. That is, the markets, the public generally, and, for that matter, the policymakers themselves take stress tests more seriously when the failure of the test will lead to concrete supervisory action to deal with solvency issues that have been identified, either by recapitalising weak banks or by resolution. However, the needs of stress testing for micro and macro purposes are not identical, and the design of stress testing systems can and should be adapted to the particular roles that policymakers intend for the tool.

Banking supervision is the primary tool used to implement microprudential regulation. This involves the examination of a bank's financial condition based on recent performance, with a focus on capital and other regulatory ratios. Supervision will also include a review of the bank's internal processes, systems and models, including its use of stress testing to explore exposures to risks it perceives. Historically and in most cases still today, the results of a bank's supervisory review are not made public.

The development of CCAR in the US and the AQR in the EU has changed this supervisory process. For the large banks under these regimes, their financial condition is examined *on a forward-looking basis under stress scenarios set by the supervisor,*

6 For an analysis of the challenges facing the Single Resolution Mechanism in Europe, see Veron (2015).

their projected capital position under stress is compared to *a capital standard set in the test,* and any deficiency requires *remedial action* under a plan approved by the supervisor. Furthermore, as we have seen, the stress test results of individual banks are revealed, in some cases, in considerable detail.

This stress testing for microprudential purposes typically examines in detail different portions of a bank's banking and trading books, which are valued under adverse conditions using models that are specific to those exposures. This *bottom-up* approach to calculating a bank's overall capital position is, by its very nature, a resource-intensive process. And since each macro scenario needs to be mapped into its implications for the detailed parameters that are typically needed in the various exposure-specific models, the number of the scenarios that can be explored is limited.

In macroprudential regulation, focus is on the health of the financial system as a whole and the ability of the banking system to support investment and economic activity generally on a sustained basis. This is useful as an input in the formulation of monetary and fiscal policy. In addition, policymakers may also consider actively using specific macroprudential tools. There is no settled set of such tools. One that is often cited is dynamic provisioning, employed by the Banco de España for a time. While it was eventually abandoned because it went against IFRS accounting standards, it was considered as having a fairly strong effect because it fed through banks' income statements to have a direct impact on profit and loss. Another approach that has been used in a number of countries is to impose a cap on the loan-to-value (LTV) ratio on mortgage loans (HKMA 2011). Alternatively, an approach that does not directly involve the valuation of the collateral for mortgage loans was taken in Switzerland in 2013 when the authorities imposed a temporary increase in the regulatory capital charge applied to Swiss residential mortgage loans (Danthine 2014). This is an example of a tool that is applied to introduce a counter-cyclical element into capital regulation. This has been generalised in Basel III, which envisages a counter-cyclical capital buffer of between 0% and 2.5% of risk-weighted assets (RWAs) to be phased in between 2016 and 2019. The activation of the buffer is the responsibility of the national authorities based on their assessment of business conditions. However, Basel guidance puts an emphasis upon credit-to-GDP ratios.

Given its flexibility, stress testing can be used in setting macroprudential regulation in a number of ways. Adapting stress tests to macroprudential needs can have implications for the specification of the scenarios, the models used to calculate values under stress, and in the policy used in disseminating results.

Different scenarios can be chosen to explore *system-wide vulnerabilities that can be either cyclical or structural in nature.* Cyclical risk factors would typically reflect a credit build-up based on a boom in collateral values (e.g. housing prices). This could be captured in the size of a hypothetical decline of housing prices considered in an adverse scenario. In periods judged to be near a peak in the credit cycle (e.g. as indicated by average LTV or the comparison of rental rates to housing purchase prices), a relatively large housing price decline might be used. As described by Edge and Lehnert in this volume, the Fed's stress tests have introduced a counter-cyclical element into capital regulation by imposing macro scenarios that envisage an increase of unemployment rates to at least 10%. With economic recovery and the fall of unemployment rates, this implies the increasing severity of stressed scenarios applied to the banks.

An example of a structural risk factor might be the heavy dependence of an economy and its banking system upon the export of single commodity (e.g. crude oil). To capture a structural vulnerability, a stress test could consider a commodity price decline, with a larger decline hypothesised for countries with a larger export concentration in that commodity. Policymakers may be interested to learn what the banks themselves believe may be their most important vulnerabilities. As described by Edge and Lehnert, CCAR calls for firms to evaluate their capital plans against their own stress scenario, as well as those specified by the supervisors. By comparing the scenarios proposed by all the banks in the exercise, supervisors may glean useful information that affects the system as a whole either directly or indirectly.

Using stress testing for macroprudential purposes has implications for the kinds of models needed for effective testing. This comes most directly from the fact that the scenarios that supervisors impose on the banks are specified in a set of hypothesised paths for a relatively small number of macroeconomic variables. From this, banks need to project losses across various categories of their loan and trading books. For example, in stress testing the largest US banks with large market exposures, an overall macro strategy is translated into stressed values of some 40,000 detailed market risk drivers

(Edge and Lehnert 2016). A similar mapping must be made for the banking book (e.g. to loan loss rates for different categories of loans). If need be, this might be done in a somewhat *ad hoc* way based on supervisory or banker judgement. Increasingly, however, analysts have sought greater consistency by combining macro models with financial sector models, often in a hierarchical or recursive manner. This approach can be followed with three or more layers of models (e.g. macro, sectoral, asset class) until detailed projections can be made at the level of the banks' own disaggregated data. Detailed loss projections can then be aggregated up to obtain implications for losses and capital at the bank level, at the group or holding company level, at the level of the set of banks being tested, and finally for the banking system as a whole. It is clear that this top-down plus bottom-up approach involves a lot of time, effort and resources. However, it is not clear that the resulting system-wide projections are necessarily superior to those obtained using a more aggregative approach.

These themes of how to use stress testing in supervision and in the analysis of system stability are discussed in Vítor Constâcio's contribution to this volume presenting the stress testing framework developed at the ECB (Constâcio 2016). As I have noted, the ECB was not directly responsible for the first three European stress tests, starting in 2009. Instead, these were coordinated by the CEBS and its successor, the EBA, and executed by the national supervisory authorities. The role of the ECB was in providing guidance on the setting of the macro scenarios used, and in this it built upon its considerable past work in macro and sectoral modelling at the ECB. This arrangement changed fundamentally with the introduction of the SSM, which for the first time places a central supervisory authority for the Eurozone in Frankfurt. Constâncio describes how the division of efforts of the ECB and the SSM are being coordinated. In particular, the SSM is charged with implementing the stress tests for microprudential purposes, including detailed analysis of exposures and implications of bank solvency ratios. The responsibility of the ECB is not limited to providing the framework for the formulation of consistent macro scenarios. Increasingly, it also has responsibility for assessing the implications of the stress tests for macroeconomic policy, for example, in examining stress test results for possible contagion effects using network models. Finally, the stress tests may guide decisions regarding macroprudential tools.

In the first three stress tests of the EU banking system, the national authorities retained responsibility for executing the tests on individual banks and for any supervisory actions that emerged from the tests. A number of representatives of national central banks and other authorities participated in the LSE conference. One approach that was discussed in some detail at the conference was that of the Bank of England, which had published a white paper on its approach to stress testing just prior to the conference (Bank of England 2015). The main features of the Bank of England's systemic stress tests are summarised in Alex Brazier's contribution to this volume (Brazier, 2016).

Somewhat similarly to the ECB, the Bank of England did not have a responsibility for banking supervision during the financial crisis. Supervision of UK banks was centralised in the hands of the Financial Services Authority between 2001 and 2013. This ended in 2013 with the return of responsibility for banking supervision to the Bank of England or, more specifically, to the Prudential Regulatory Authority (PRA), which operates as part of the Bank of England. With the creation of the PRA, Governor Mark Carney announced a policy of annual stress tests for the major UK banks that are designed both to assist the PRA in its supervisory role and also to inform the Bank of England in the formulation of its macroprudential policy (Bank of England 2013). While the first sets of concurrent stress tests in 2014 and 2015 focused principally on the examination of solvency of individual banks, internal discussions within the Bank of England were aimed at evolving towards a system that would be more explicitly linked to the formulation of macroprudential policy.

Under the framework set out in its 2015 White Paper, the Bank of England aims to use stress tests in developing an explicitly counter-cyclical policy that is an integral part of the capital regime for the UK's systemically important banks. Stress testing is to be used to assess how the risk exposures of the banks are changing with business conditions. When risk exposures have been rising for the system as a whole, the amount of capital the large banks will need to hold will be increased. In particular, the stress tests will inform the Bank of England's decision on the counter-cyclical capital buffer and sectoral capital requirements, which come on top of the capital conservation buffer that will be applied independently of macro conditions. Stress test can also inform the decision of the PRA to apply an additional bank-specific buffer if it deems that appropriate.

One noteworthy aspect of the new Bank of England approach is that the cyclical buffer will be linked to a macro stress testing framework that is intended to be stable over time. This means that the banks themselves should be able to anticipate the kinds of macro scenarios the Bank of England will apply in the next stress tests by looking at the evolution of business conditions. By running this through their own systems, banks will be able to have an idea of the likely evolution of their capital ratios, and they can build that into their capital planning process. While the banks themselves may not know in detail the models that the Bank of England uses in making its decisions, they should be able to anticipate in which direction capital buffers will be adjusted. In this way, the Bank of England's system is aimed at reducing regulatory uncertainty.

In his analysis in this volume of the Bank of England's proposed new approach to stress testing, Philipp Hartmann notes that most past thinking about marcroprudential tools has been 'asymmetric' in the sense that the focus is on imposing harsher, and therefore more prudent, standards in the upswing of the credit cycle (Hartmann 2016). The Bank of England approach is significantly more ambitious than this in that it proposes a 'symmetric' approach to macroprudential regulation whereby both increases and decreases to capital buffers are made using the same stress testing framework. The approach aspires to make the regulatory process foreseeable, although not the actual path of the buffer, which itself will depend on future economic conditions. In this way, it might be interpreted as being a 'rule-based' approach to macroprudential regulation. However, Hartmann points out that there is a major practical difficulty of determining in real time where you are in the cycle, and this inevitably will require policymakers' judgement. For this reason, he suggests 'constrained discretion' may be a more apt description.

Another view of the Bank of England's White Paper and Alex Brazier's summary is given by Thomas Huertas in this volume (Huertas 2016). He points out an important feature of the UK approach – namely, that it focuses on the ability of large UK banking groups to sustain lending to the UK economy during periods of macroeconomic stress. This implies that it takes a particular, and in some ways limited, view of UK banking on the one hand, and of what is required to support the UK economy on the other. In particular, UK banking groups to varying degrees are international in their scope, with lending to UK entities constituting only a fraction of their balance sheets. Stress

tests focusing principally on UK macroeconomic conditions can easily miss important sources of risk for these banks. Furthermore, given the importance of capital markets in the UK, credit provided by UK banks is only one source of funds for UK firms. Even if bank stress tests succeed in assuring the ability of those banks to maintain their lending during a cyclical downturn, this does not mean that they can compensate for the likely withdrawal of funding from capital markets. Finally, while noting that the new Bank of England policy represents progress in using stress testing as a tool for macroprudential policy, Huertas points out that the emphasis is placed squarely on how to regulate the banks' capital ratios through a combination of statutory minima and a number of buffers. The link to other, possibly more effective, tools of macroprudential policy, such as loan-to-value ratios, is largely incomplete.

The theme that supervisory stress testing places too much emphasis on regulatory capital ratios has been a recurrent theme in recent writings of Viral Acharya, with a number of co-authors. He returned to this theme in his presentation at the conference, in which he was critical of the EU's Comprehensive Assessment programme of 2014. At the heart of the critique is the view that focusing on capital ratios can be misleading because of problems with both the denominator (i.e. RWAs) and the numerator (i.e. regulatory capital) of the ratio. In this view, the problem with RWAs is that they are based on banks' own internal models, which are opaque from the point of view of the external market and which will reflect supervisory judgement that may vary across national supervisors within the EU. A problem with regulatory capital is that under Basel II, it was well known that different supervisors applied different criteria of what would qualify as Tier 1 or Tier 2 capital, and this lack of consistency has not been removed by Basel III's move to emphasise CET1.

Because of these weaknesses, Acharya argued that it is useful to check the results of the regulatory stress tests against an alternative measure of systemic risk based on publicly available information. His preferred alternative is SRISK, a measure of the capital a firm would need to raise in the event of a crisis, developed by Acharya and colleagues at New York University (Acharya et al. 2012). This measure is calculated as the expected shortfall of the market value of capital relative to the regulatory minimum level of capital conditional on the economy as a whole being in a crisis. In their contribution to this volume, the SRISK methodology is applied by Acharya and Sascha Steffin to a

subset of 39 publicly listed banks out of the set of 130 banks that were included in the 2014 EU Comprehensive Assessment (Acharya and Steffin 2016). Their definition of a crisis is a 40% fall in the stock market index. Based on this measure, they come to very different conclusions from those of the EU. In particular, they find a total shortfall under their stress scenario that is more than 20 times that of the capital shortfall found by the EU. The largest SRISK losses are in Germany, France, Spain, Italy and Belgium. In contrast, the included banks in the first three of these countries produced no shortfall whatsoever in the EU tests. Finally, there is a negative correlation between a bank's SRISK measure and its expected shortfall in the EU test.

These are strong and pointed criticisms not only of the EU's Comprehensive Assessment programme, but implicitly of all current methods of regulatory stress tests. In a sense, it is an extreme position based on the argument, made above, that greater detail and more elaborate model building do not necessarily produce more accurate estimates of aggregate quantities of interest for macroprudential policy. At its heart, the SRISK measure is driven by the historical co-variation between the market value of a bank's equity and the value of the overall stock market. This is an extremely simple reduced-form model, and the resource and time costs to produce it are negligible when compared to those of the regulatory stress tests of the EU, the Bank of England and the US.

Until now, policymakers have not moved to incorporate reduced-form, market-based assessments of financial strength into a central position in the supervisory process. In his contribution to this volume, Vítor Constâncio explains why not. First, by simply defining stresses based on a stock market index, these measures are divorced from any reflection of the broader macroeconomic context. Second, the market-based systemic risk measures such as SRISK are very sensitive to the choice of parameters. Different parameter values that are hard to distinguish statistically can translate into vastly different estimates of the amount of capital required to withstand a stress. Beyond this, the loss measures are extremely volatile over time and are, by their nature, procyclical. So, for example, in 2009 when SCAP was used to give clarity about underlying valuations based on projected future earnings under stressed conditions, a simple reduced-form, market-based measure would have tended to undermine confidence.

This is not to say that a major fall in a stock market index or in some other financial market benchmark is not of interest in stress testing and in macroprudential regulation

generally. Indeed, an extreme fall in financial markets can result in a fall in collateral values, and that can provoke a market reactions that can spread widely. Of course, this is at the heart of modern thinking about systemic risk.[7] Bank solvency and financial stability generally is exposed to a variety of risks that can be categorised as major economic shocks, contagion, and corrections in market imbalances (sometimes described as the 'bursting of bubbles'). As we have described here, macro stress testing places emphasis on examining the consequences of major macroeconomic shocks such as declines in GDP and increases in unemployment. Supervisors and central banks have built up considerable modelling capability to analyse risks of this nature. Regarding analysis of contagion, building on a burst of theoretical literature on this subject in recent years, supervisors have begun to develop practical models that can be calibrated to real world data, and, as discussed in the contribution by Constâncio, these are used in analysing the macro implications of the results found in supervisory stress tests. Examples of some of these models were presented at the LSE conference (Anand et al. 2015, Idier and Piquard 2015) However, at this stage there is no class of models that has proved to be a reliable workhorse in this area. Further work is underway, and that is a welcome development. Finally, in regard to the idea of integrating possible major corrections of bubbles or other market anomalies, there is some general recognition that these are potential hazards, but there cannot be said to be any settled official view on how they can be detected. As discussed by Alex Brazier in this volume, the Bank of England's new approach to stress testing envisages using 'exploratory' scenarios as well as more traditional macro stress scenarios. Potentially, this could be used to explore consequences of a major market correction. However, in light of resource costs involved in doing a thorough system-wide test, only one such scenario is planned every other year.

The use of exploratory scenarios in stress testing at the Bank of England reflects the fact that by 2015, the UK banking system had emerged from the financial crisis which in the UK had started with the run on Northern Rock in 2007. So the task of the UK stress testers is to probe the economy to discover points of fragility that could threaten financial stability in the future. This is a very different use of stress tests than the Federal

7 For a good brief summary of the literature, see ECB (2009).

Reserve's SCAP exercise in 2009. This is a theme that is developed in Til Schuermann's contribution to this volume (Schuermann 2016).

In Schuermann's terminology, the SCAP was deployed in 'wartime' when the major hazard clearly was the possible onset of a severe recession, and the primary question for all was whether the major banks had sufficient capital to survive the recession were it to occur. In that context transparency and the availability of a credible fiscal backstop were crucial features that made SCAP a success.

In contrast, in 'peacetime', such as the current context for the Bank of England, there is no single clear threat to financial stability. Instead, there are many potential threats. The Bank of England intends to probe for signs of vulnerability by using exploratory scenarios in its stress tests. However, there are alternatives. Schuermann argues in favour of the approach taken in the US CCAR programme, which requires testing not only under the scenario imposed by the supervisor but also under the banks' own scenarios. Including banks' scenarios has a number of advantages. It allows the supervisors to learn something about the ability of an individual bank to probe its own potential weaknesses independently using its risk management tools. This helps to counter-balance a potentially dangerous tendency toward risk 'mono-culture', which could set in if banks simply viewed their stress testing as a compliance exercise in which they need to demonstrate that they meet a set of routine health checks. Furthermore, the supervisors learn from the banks' own scenarios something about the diversity of potential hazards faced by the financial system as a whole. They can then incorporate some of these insights into their own models and procedures so as to have a more comprehensive approach to macroprudential regulation.

In 'peacetime', full transparency of stress tests may not be an essential feature. For example, banks will open up to supervisors more if they know that market-sensitive information will not be made public. Furthermore, dissemination of information about

vulnerabilities may undermine risk pooling in the banking sector because of so-called Hirschleifer effects. [8]

The idea that stress testing may contribute to the supervision of banks outside of a crisis is also developed in the contribution of Charles Goodhart to this volume (Goodhart 2016). He argues that stress tests can supply a crucial tool currently missing from the framework for bank recovery and resolution that is being put into place in the US, Europe and in other jurisdictions. The new framework has involved the development of recovery and resolution plans, the so-called 'living wills'. It has also involved the creation of special resolution regimes designed to deal with banks at the point of insolvency. The two crucial contributions in resolution have been the designation of authorities specifically charged with the resolution decision and the introduction of bail-in that is intended to expedite resolution and thereby avoid contagion, as well as protect against bailouts by taxpayers. The main unresolved issue in this architecture is how to induce recovery in advance of the point of insolvency, when the going-concern value of the bank is worth preserving. Experience has shown that bankers are reluctant to undertake recovery, because of their aversion to share dilution and a fear for their jobs if their actions were to imply the failure of previously chosen strategies. Bank supervisors may be aware that the bankers are perhaps delaying recovery plans too long, but they have lacked concrete tools to force the bankers' hands. In effect, what they need is a tool that would have allowed them to intervene in Lehman in May 2008, when the bank could have been salvaged as a going concern, rather than in September 2008 when it was past hope and contagion was rampant. Stress testing potentially can be the tool that gives a formal trigger for the resolution process.

I have emphasised that one of the attractions of stress testing is its flexibility. This means that the analyst can adapt the tool to suit the purpose. Of course, the purpose will be shaped by both institutions and historical context. This chapter and most of the contributions to this book illustrate this using transatlantic experience since 2009. This includes crisis episodes first in the US and then in Europe, and post-crisis 'peacetime'.

8 Named after Jack Hirshleifer, who showed how introducing information can undermine the incentive for parties to enter into risk-sharing arrangement. For the application of this idea in the context of banking crises, see Alvarez and Barlevy (2015).

It includes the stable institutional context of the US and the experience in the EU, where responsibility for stress tests has moved from an advisory body (CEBS), to a coordinating authority without direct supervisory capabilities (EBA), to a rather unified supervisory framework (SSM). This has led to a variety of approaches to stress testing and to an agenda for future developments as the tool is refined in light of experience. However, there is a broad similarity in the approaches used in the US, the EU and the UK, with annual supervisory stress tests being the norm and with stress testing integrated with assessment of banks capital planning.

However, if you expand the field of view to go beyond the transatlantic focus, the diversity of institutions, historical contexts and approaches to stress testing is much greater. This is the theme developed in the contribution of Udaibir Das to this volume (Das 2016). The IMF has been using stress testing for more than 15 years, notably as part of its Financial Sector Assessment Programs. Based on hundreds of stress tests run for a wide range of countries, the IMF has experienced the diversity of contexts first-hand. In this way, stress testing has become an important part of its risk assessments and its surveillance functions. The IMF has developed a set of in-house models, generally of the macro top-down variety, that can be used in conjunction with bottom-up analyses provided by local institutions in the country being assessed. It has also seen that not all countries have the same quantity and quality of data, so that certain sources of risk may sometimes remain rather obscure. Finally, the effectiveness of stress testing approaches depends as well on a variety of what can be seen as governance conditions. These include clarity of the institutional coverage of stress testing (banks, insurance, markets, etc.), the degree of senior management buy-in, and the agreement on dissemination of stress testing results.

The governance of stress testing was a theme that was discussed at the LSE conference, specifically in the context of the accountability of the stress test regime. As discussed at the outset of this chapter, stress testing in major countries has emerged from being used principally as tool of banks' internal risk management to become the cutting edge of supervision for the systemically important banks. How has this come about? It seems to me that this rise of supervisory stress testing is founded largely on the prestige acquired through the justifiable perception that stress testing was crucial to stabilising financial markets in the US in 2009. However, reputation is ephemeral. It is an open question

whether stress testing will survive as the centrepiece of banking supervision if, in the future, another bank that passes the annual stress test soon after fails and provokes a crisis, even a minor one.

Stress testing is likely to maintain its place in banking regulation and supervision if the participants in the process understand both how stress testing can contribute to promoting financial stability and what it costs to achieve those benefits. Many aspects of banking supervisory regimes have been subjected to exchanges of view among stakeholders, with Basel consultation processes being a good example. There has been no comparable consultation process leading up to the current regime of supervisory stress tests. The banking industry has voiced some of its concerns. One complaint is that supervisory stress test models are a black box that leaves them in the dark as to what standard banks' risk management processes are being held to. Another is that annual stress tests, with changing types of scenarios each year, repeated in multiple jurisdictions but with different methodologies and data requirement, have contributed to the sharp increase in compliance costs in recent years. Policymakers have responded to some extent, by explaining the reasoning and research behind the supervisory models and by announcing the intended direction of stress testing regimes so that banks can plan accordingly. But probably more could and should be done to facilitate dialogue and to assure that supervisory stress testing regimes are accountable and adaptable, rather than arbitrary and rigid.

Acknowledgements

In this chapter, I have attempted to collect and synthesise some of the important themes that ran through the London School of Economics stress testing conference. I have sometimes introduced my own personal views, and in this case I have used the first-person singular in presenting an idea. The list of presentations made at the conference can be found in the programme in the appendix to this chapter. Thanks go to all of those appearing on the programme, but also to those who participated in the roundtable discussion (conducted under the Chatham House rule that ideas discussed can be reported but without specific attribution), namely *Franklin Allen, Alex Brazier, Jill Cetina, Rohan Churm, Jon Danielsson, Udaibir Das, Olivier de Bandt, Rochelle Edge, Morris Goldstein, Charles Goodhart, Piers Haben, Philipp Hartmann, Charlie*

Haswell, Jérôme Henry, Tom Huertas, Christian Julliard, Malcolm Knight, Andreas Lehnert, Massimo Marchesi, Matthew Pritsker, Til Schuermann, Alison Scott, Miguel Segoviano, Dimitrios Tsomocos and Jakob Winstrand.

References

Acharya, V. and S. Steffin (2016), "Making sense of the Comprehensive Assessment," in R.W. Anderson (ed.), *Stress Testing and Macroprudential Regulation: A Transatlantic Assessment*, London: CEPR Press.

Acharya, V., R. Engle, and M. Richardson (2012), "Capital Shortfall: A New Approach to Rankings and Regulating Systemic Risks", *American Economic Review Papers and Proceedings* 102(3), pp. 59–64.

Alvarez, F. and G. Barlevy (2015), "Mandatory Disclosure and Financial Contagion" (http://home.uchicago.edu/~falvare/disclosure_June2015.pdf).

Anand, K., C. Gauthierz and M. Souissix (2015)," Quantifying Contagion Risk in Funding Markets: A Model-Based Stress-Testing Approach", Bank of Canada Working Paper (http://www.bankofcanada.ca/wp-content/uploads/2015/08/wp2015-32.pdf).

Bank of England (2013), "A Framework for Stress Testing the UK Banking System", 30 September.

Bank of England (2015),"The Bank of England's Approach to Stress Testing the UK Banking System", approach document, 21 October (http://www.bankofengland.co.uk/publications/Pages/news/2015/076.aspx).

Bernanke, B. (2013), "Stress Testing Banks: What Have We Learned?", speech at the Federal Reserve Bank of Atlanta, 8 April (http://www.federalreserve.gov/newsevents/speech/bernanke20130408a.htm).

Brazier, A. (2016), "The Bank of England's approach to stress testing the UK banking system", in R.W. Anderson (ed.), *Stress Testing and Macroprudential Regulation: A Transatlantic Assessment*, London: CEPR Press.

Constâncio, V. (2016), "The role of stress testing in supervision and macroprudential policy", in R.W. Anderson (ed.), *Stress Testing and Macroprudential Regulation: A Transatlantic Assessment*, London: CEPR Press.

Danthine, J.-P. (2014), "Implementing Macroprudential Policies: The Swiss Approach", *Banque de France Financial Stability Review,* April.

Das, U. (2016), "Stress testing as a policy instrument: Some thoughts", in R.W. Anderson (ed.), *Stress Testing and Macroprudential Regulation: A Transatlantic Assessment*, London: CEPR Press.

de Larosière Group (2009), *Report of the High-level Group on Banking Supervision*, Brussels: European Commission (http://ec.europa.eu/internal_market/finances/docs/de_larosiere_report_en.pdf).

Edge, R. and A. Lehnert (2016), "Recent experience with supervisory stress testing in the United States", in R.W. Anderson (ed.), *Stress Testing and Macroprudential Regulation: A Transatlantic Assessment*, London: CEPR Press.

European Central Bank (ECB) (2009) "The Concept of Systemic Risk", *European Central Bank Financial Stability Review*, pp. 134-142

Geithner, T. (2014), *Stress Test: Reflections on Financial Crises*, New York: Broadway Books.

Goodhart, C. (2016), "In praise of stress tests", in R.W. Anderson (ed.), *Stress Testing and Macroprudential Regulation: A Transatlantic Assessment*, London: CEPR Press.

Hartmann, P. (2016), "Stress testing over the financial cycle: General approach, challenges and complements," in R.W. Anderson (ed.), *Stress Testing and Macroprudential Regulation: A Transatlantic Assessment*, London: CEPR Press.

Hirtle, G. and A. Lehnert (2015), "Supervisory Stress Tests*", Annual Review of Financial Economics* 7, pp. 339–55.

Hong Kong Monetary Authority (HKMA) (2011), "Loan-to-Value Ratio as a Macroprudential Tool – Hong Kong SAR's Experience and Cross-country Evidence," BIS Paper No. 57, Basel.

Huertas, T. (2016), "Do stress tests pass the test?", in R.W. Anderson (ed.), *Stress Testing and Macroprudential Regulation: A Transatlantic Assessment*, London: CEPR Press.

Idier, J. and T. Piquard (2015), "Pandemic in Financial Systems and Liquidity Emergency", Banque de France.

Institute of International Finance (IIF) (2009), *Restoring Confidence, Creating Resilience: An Industry Perspective on the Future of International Financial Regulation and the Search for Stability*, Washington, DC.

Krugman, P. (2014), "Does He Pass the Test?", *New York Review of Books,* 10 July.

Michelides, A. (2014), "Cyprus: From Boom to Bail-in", *Economic Policy* 29, pp. 639-689.

Schuermann, T. (2016), "Stress testing in wartime and in peacetime", in R.W. Anderson (ed.), *Stress Testing and Macroprudential Regulation: A Transatlantic Assessment*, London: CEPR Press.

Van Lelyveld, I. (2009), "Stress Testing: The Link between Micro and Macro", *International Journal of Central Banking,* September, pp. 1-8.

Veron, N. (2015), *Europe's Radical Banking Union,* Brussels: Bruegel.

Appendix

Stress Testing and Macro-prudential Regulation:
A Trans-Atlantic Assessment

London School of Economics
29-30 October 2015

Programme

29th October 2015

"Overview of Independent Supervisory Stress Testing in the United States"
Andreas Lehnert (Deputy Director in the Office of Financial Stability and Research, Board of Governors of the Federal Reserve System)

"Testing Macroprudential Stress Tests: The Risk of Regulatory Risk Weights"
Viral Acharya (Professor of Finance, New York University)

"In Praise of Stress Tests"
Charles Goodhart (London School of Economics)

"The Role of Stress Testing in Supervision and Macroprudential Policy"
Vítor Constâncio (Vice-President, European Central Bank)

"Stress Testing in Wartime and in Peacetime"
Til Schuermann (Partner, Oliver Wyman and Wharton Financial Center)

"Stress Testing and Systemic Stability"
Hyun-Song Shin (Economic Advisor and Head of Research, Bank of International Settlements and CEPR)

"Mitigating Systemic Risk: Is Stress Testing an Effective Policy Instrument?"
Udaibir Das (Division Chief, Sovereign Asset & Liability Management Division, Monetary & Capital Markets Department, International Monetary Fund)

30th October 2015

"The Bank of England's approach to stress testing the UK banking system"
Alex Brazier (Executive Director, Financial Stability and Strategy, Bank of England)

Discussion of "The Bank of England's approach to stress testing the UK banking system" by Alex Brazier
Thomas Huertas (partner, Ernst & Young)

Discussion of "The Bank of England's approach to stress testing the UK banking system" by Alex Brazier
Philipp Hartmann (Deputy Director General Research, European Central Bank)

"Finding Comfort in Stress Tests"
Deniz Igan (International Monetary Fund)

Discussion of "Finding Comfort in Stress Tests" by Deniz Igan
Alex Michelides (Professor of Finance, Imperial College)

"The Bank-Sovereign Nexus Across Borders"
Bernd Schwaab (European Central Bank)

Discussion of "The Bank-Sovereign Nexus Across Borders" by Bernd Schwaab
Sascha Steffen (European School of Management and Technology, Berlin)

"Bank Stress Testing: A Stochastic Simulation Framework to Assess Financial Fragility"
Giuseppe Montesi (University of Siena)

Discussion of "Bank Stress Testing: A Stochastic Simulation Framework to Assess Financial Fragility" by Giuseppe Montesi
Jon Danielsson (London School of Economics, co-director System Risk Centre)

"Choosing Stress Scenarios for Systemic Risk Through Dimension Reduction"
Matthew Pritsker (Senior Financial Economist, Federal Reserve Bank of Boston)

Discussion of "Choosing Stress Scenarios for Systemic Risk Through Dimension Reduction" by Matthew Pritsker
Christian Julliard (Department of Finance, London School of Economics)

"Quantifying Contagion Risk in Funding Markets: A Model-Based Stress-Testing Approach"
Kartik Anand (Research Department, Deutsche Bundesbank)

Discussion of "Quantifying Contagion Risk in Funding Markets: A Model-Based Stress-Testing Approach" by Kartik Anand
Pierre Chaigneau (HEC Montréal)

"Pandemic Crises in Financial System and Liquidity Emergency"
Julien Idier (Banque de France)

Discussion of "Pandemic Crises in Financial System and Liquidity Emergency" by Julien Idier
Aytek Malkhosov (Bank of International Settlements)

2 Recent experience with supervisory stress testing in the United States

Rochelle Edge and Andreas Lehnert
Board of Governors of the Federal Reserve System

Introduction

As described in Hirtle and Lehnert (2015), stress tests were used by banks, rating agencies and supervisors in the US prior to the financial crisis. In 2009, at the height of the crisis, stress tests were used as an important part of the official response to the crisis. Since then, stress testing has taken a central role in the supervision of the capital adequacy of the largest bank holding companies in the US. Supervisors project participating firms' *pro forma* capital ratios under stressful scenarios using models developed by supervisors and data supplied by the firms. The results are disclosed to the public and used to approve or reject firms' plans to distribute capital to shareholders. This quantitative exercise is tied to an associated qualitative exercise; the capital plans of firms failing to meet supervisory standards for modelling and governance may also be rejected despite passing the quantitative thresholds.

We document key features of the current regime, describe the evolution of supervisory stress tests from 2010 to 2015, and analyse the nature and results of the quantitative exercise, including the supervisory scenarios and projected losses.

Key features of the current regime

Before reviewing the experience over the past several years, it is useful to begin with a brief overview of the key features of the supervisory stress testing regime as of the

end of 2015. Of course, these features may change over time. We follow Hirtle and Lehnert (2015), who describe five design choices facing supervisors when putting in place a stress testing regime.

Supervisory scenario design. The US regime currently features three scenarios specified by supervisors: baseline, adverse and severely adverse. The severely adverse scenario is designed using a 'recession approach', in which the scenario will always feature a severe US recession. However, scenarios are somewhat countercyclical, with unemployment generally expected to rise at least four percentage points or to a level of 10%, whichever results in the higher level. (In addition, participating banks are required to specify a scenario that is tailored to their business model and risks, known as the BHC scenario.)

Disclosure. Supervisors disclose extensive details on their quantitative findings, publishing estimated loss rates for each participating bank under the adverse and severely adverse scenarios by major portfolio grouping (e.g. wholesale loans, junior lien mortgages, securities, and so forth) and estimated revenues. In addition, supervisors disclose firms' minimum and ending *pro forma* capital ratios under their original planned capital actions and, for firms that revised down their proposed capital distributions in light of the results, under the revised capital actions as well. Firms are required to disclose their own estimates of losses in a format comparable to that used by supervisors. Finally, supervisors disclose whether they objected (in some form) to firms' capital plans.

Capital policy. Firms' capital adequacy is evaluated assuming that they continue to make their proposed capital payouts even under the adverse and severely adverse scenarios. Firms whose *pro forma* capital ratios under the stress scenarios fall below regulatory thresholds can be required to submit revised capital plans that demonstrate how they will make up the shortfall (e.g. by reducing their planned payouts).

Balance sheets. The Federal Reserve independently projects firms' balance sheets and, perforce, their total and risk-weighted assets for the purposes of computing regulatory capital ratios. Balance sheet components such as loans and securities are assumed to grow at the same rate across all firms. Moreover, the projections are made under the assumption that credit supply does not fall in the scenario.

Models. The Federal Reserve develops or selects models used to estimate losses, income and balance sheets. These models form the basis of the supervisory estimates of *pro forma* capital levels. While the stress test disclosures contain extensive descriptions of these models, the precise formulas and assumptions are not made public.

Key developments, 2010 to 2015

The supervisory stress testing regime put in place following the financial crisis, known as CCAR, shared many features of the 2009 stress test, known as SCAP.[1] Supervisors publish a hypothetical stressful scenario, the largest banks participate simultaneously, *pro forma* bank capital ratios are computed under the scenario, these results are publicly disclosed, and supervisory actions are tied to the results.

Nonetheless, there are important differences. As described in Board of Governors of the Federal Reserve System (2011a):

> *The SCAP was focused on deriving supervisory stress test results to identify large bank holding companies with capital insufficient to weather a more adverse than anticipated economic environment while maintaining their capacity to lend. Firms identified as having such capital shortfalls were required to raise specific dollar amounts of capital within six months of the release of the SCAP results and a government backstop was in place if firms were unable to raise the required capital from private markets.*

By contrast:

> *…the CCAR is a broad supervisory exercise, part of the Federal Reserve's ongoing efforts to ensure that large bank holding companies have robust internal processes for assessing capital adequacy and carrying out capital planning, as well as active board of directors involvement in overseeing these processes and approving the*

1 CCAR is the acronym for Comprehensive Capital Analysis and Review; SCAP is the acronym for Supervisory Capital Assessment Program.

plans. In other words, the CCAR rounds out the Federal Reserve's assessment of the overall capital adequacy of these firms. This assessment includes, but is much broader than, an assessment of stress scenario results and a firm's sensitivity to different assumptions about potential losses.

Moreover, the post-crisis supervisory stress testing regime also underwent significant changes once in place. In this section, we review the key developments in the five annual exercises concluded from 2011 to 2015.

For the five stress tests that we review, the timing of key steps in a given year's test was as follows:

- Publication of instructions and the scenarios (generally in October to November). The scenarios jump off from the third quarter of each year, so that the fourth quarter of that year and the next two years (i.e. nine quarters in all) comprise the projection period.
- Submission by participating firms of data on key risk positions to be used in the test (generally in November to December).
- Submission by participating firms of their capital plans (generally in early January).
- Publication of results (generally in mid-March).

Thus, for example, the 2013 exercise actually started in October 2012. Starting with the 2016 cycle, this calendar was revised so that the cycle begins in January and ends in June of each year.

Table 1 shows, for each exercise, the number of bank holding companies that participated and, of those, the number that 'passed' CCAR, defined as not being objected to on qualitative or quantitative grounds and not receiving a 'conditional non-objection'.

Table 1 Participants and results: CCAR 2011 – CCAR/DFAST 2015

Year	Number of participants	Number with unconditional non-objections on both quantitative and qualitative grounds
CCAR 2011	19	Not disclosed
CCAR 2012	19	15
CCAR/DFAST 2013	18	14
CCAR/DFAST 2014	30	25
CCAR/DFAST 2015	31	28

CCAR 2011

The post-crisis supervisory stress testing regime began in earnest with the November 2010 publication of guidelines for evaluating bank capital actions that might diminish a firm's capital base. What would become the first CCAR, later dubbed 'CCAR 2011', formally began with these guidelines. Initially applied only to the 19 firms that had participated in the original 2009 stress test, the guidelines took the form of a supervisory letter directing the banks to consult with Federal Reserve staff before undertaking capital actions. This particular letter took the form of a temporary addendum to a previous letter (dubbed 'SR 09-4') and would itself be revised, hence the somewhat cumbersome and obscure title: "Revised Temporary Addendum to SR letter 09-4: Dividend Increases and Other Capital Distributions for the 19 Supervisory Capital Assessment Program Bank Holding Companies". The original letter ('SR 09-4') described general principles for determining capital adequacy, but focused on assessing the capital payouts of banks that had received public assistance in the form of capital injections during the financial crisis. The subsequent communication – the revised temporary addendum – specified that the firms should submit capital plans, described the elements that such plans should feature, and laid out the criteria used to evaluate these plans (Board of Governors of the Federal Reserve System 2009a, b, c, 2010).

These November 2010 guidelines contained a few elements that would be particularly important in later exercises:

1. The requirement that participating firms submit capital plans which described their planned capital actions, their processes for determining their needed capital levels, scenarios devised by them to be tailored to their own business model and their own estimates of their post-stress capital ratios under both their own stress scenario and a scenario supplied by the supervisors.
2. The commitment by supervisors to evaluate firms' capital adequacy using independent supervisory projections of losses and income (as well the firms' own quantitative results).
3. The specification that capital adequacy would be judged using *pro forma* capital levels under the various scenarios and, crucially, that firms' planned capital payouts under the baseline scenario would be assumed to continue under the stress scenario.
4. The nature of the supervisory decision, that is, whether to object to a firm's capital plan (or, as described in the supervisory letter, that '…the Federal Reserve may advise the BHC to revise and resubmit its capital plan for evaluation during a subsequent calendar quarter').
5. The basis for the supervisory decision, which could be driven by either qualitative deficiencies in a firm's capital planning processes, or by a quantitative finding that a firm had insufficient capital.

Elements (1) and (2) above can be thought of in terms of the identity given in equation 1 for the change in regulatory capital, which, when cumulated over time, represents the numerator of the capital ratio:

$$
\begin{aligned}
&\Delta \text{ Regulatory Capital} \\
&= \underbrace{\text{Net interest income} + \text{Non-interest income} - \text{Non-interest expense}}_{\text{Pre-provision new revenue (PPNR)}} \\
&+ \text{ Other revenues} - \text{Provisions for loan losses} - \text{Realised losses on securities} \\
&+ \text{ Realised losses on fair-value loans} - \text{Trading and counterparty losses} \\
&- \text{ Taxes} + \text{Extraordinary items net of taxes} \\
&- \text{ Net capital distributions to common and preferred shareholders} \\
&- \text{ Deductions \& additions to regulatory capital (e.g. OCI)}
\end{aligned}
\tag{1}
$$

Element (1) of the exercise – that is, capital plans submitted by banks – is the item in the penultimate line of equation 1. All of other items in equation 1 are estimated – conditional on a scenario – as part of the stress test. In the SCAP stress tests, all non-capital distribution items of equation 1 were projections that were made by banks, albeit with supervisory models, estimated on aggregate time-series data, being used to evaluate projections and highlight outliers. In the first CCAR exercise, a notable development was that supervisors estimated all of the non-capital distribution items in equation 1 with the exception of the revenue series (that is, pre-provision net revenue and other revenue).

The 2011 exercise featured three scenarios: two specified by each firm separately (a baseline and a stress) and applied by the firm to its own balance sheet, and one specified by the supervisors and applied to all firms. The supervisory scenario specified paths for nine variables – all of which were for US economic and financial variables – and was disclosed only at the end of the exercise, not at the beginning. (We defer discussion of the scenarios to the next section).

Supervisors did not publicly disclose either their quantitative projections, which were calculated conditional on the supervisory stress scenario, or their decisions regarding whether to object to firms' capital plans. Participating firms were told whether supervisors objected to their capital plans; several firms announced these decisions publicly, but some did not (Board of Governors of the Federal Reserve System 2011a).

CCAR 2012

The 2012 exercise was conducted under the terms of the Capital Plan Rule, finalised in November 2011. This rule, subsequently amended several times, has governed CCAR since it was put in place (Board of Governors of the Federal Reserve System 2011b). While a complete discussion of the elements of the rule is beyond the scope of this chapter, it is useful to note that it made CCAR an annual process and formalised many of the elements laid out in the November 2010 supervisory letter described in the previous section.

The exercise used two supervisory scenarios (a baseline and a stress), published at the same time as the instructions to the participating firms on 22 November 2011. Relative to the previous exercise, the scope of the scenarios was significantly expanded, including a description of global macroeconomic developments. This added 12 variables to the scenario – specifically, three variables for each country/country-block of the scenario. It is worth noting that the scenario was designed and published during a period of financial turmoil related to developments in European sovereign debt markets. As with the 2011 exercise, CCAR 2012 also features two firm-specified scenarios.

The results published in March 2012 disclosed supervisory estimates of firms' losses and income by portfolio category under the supervisory stress scenario, as well as supervisory estimates of firms' *pro forma* stressed capital ratios. This was the first such disclosure since the original SCAP results were publish in May 2009, which has continued for all subsequent stress tests. For the first time, the Federal Reserve disclosed whether it had objected to firms' capital plans on either quantitative or qualitative grounds.

CCAR/DFAST 2013

The key developments in the 2013 exercise included the implementation of the Dodd-Frank Act's supervisory stress testing requirement, the extension of independent supervisory estimates of key series to revenue generation, and the publication of a scenario design framework.

The Dodd-Frank Act, passed in July 2010, contained many provisions related to stress testing. Of particular note, the Act directed the Federal Reserve to undertake independent stress tests of bank holding companies with assets greater than $50 billion using three scenarios – a baseline, an adverse and a severely adverse scenario – and to publish the results. The final rule adopted by the Federal Reserve in October 2012 integrated the Dodd-Frank Act's stress tests, known as DFAST, with the existing CCAR framework.

The models and general framework are the same under DFAST and the quantitative aspect of CCAR. In addition, the scenarios are largely the same, with the sole difference being that DFAST has three scenarios (the baseline, adverse, and severely adverse

scenarios) whereas CCAR has two (the baseline and severely adverse scenario, also referred to as the 'stress scenario'). The key differences between DFAST and CCAR are in the assumed paths for capital distributions and the consequences of having minimum stress pro forma capital levels below regulatory thresholds. Under DFAST, firms are assumed to pay dividends at the average rate they had done in the past year under the stress scenarios. There are no supervisory or regulatory consequences if pro forma capital ratios fall below regulatory thresholds. Under CCAR, firms are assumed to make their planned capital payouts – dividends and share repurchases – under the stress scenarios. If pro forma capital ratios fall below regulatory thresholds, supervisors may object to the bank's capital plan. As part of an objection, capital payouts can cease.

When the results of the 2013 exercise were published in March 2013, the Federal Reserve announced publicly that it was objecting to the capital plans of two firms; a further two firms received a 'conditional non-objection', requiring them to remediate weaknesses in their capital planning processes and submit revised capital plans within six months. A 2013 paper described in some detail the qualities of good capital planning processes (Board of Governors of the Federal Reserve System 2013a).

As noted earlier, an important element of the new stress testing regime was the commitment by supervisors to evaluate firms' capital adequacy using independent supervisory projections of losses and income (as well the firms' own quantitative results). In the 2013 exercise, supervisors began projecting the revenues (that is, PPNR and other revenues) components of equation 1.

Since CCAR/DFAST 2013, the Federal Reserve has followed a clearly articulated framework for specifying the CCAR/DFAST scenarios. A proposed version of this framework – often referred to as the Policy Statement on Scenario Design – was published for comment in autumn 2012 and then finalised in autumn 2013 (see Board of Governors of the Federal Reserve System 2013b, and further discussion in the following section).

CCAR/DFAST 2014

The most notable development in the 2014 exercise was the use of independently projected supervisory paths for banks' balance sheets and risk-weighted assets

(RWAs). As numerous studies have pointed out (e.g. Greenlaw et al. 2012), if banks are permitted to project shrinking balance sheets under stress scenarios, *pro forma* capital ratios are, all else constant, higher and thus the stress test easier. But such projections are inconsistent with preventing capital losses at banks from feeding back to the macroeconomy.

In a December 2013 letter, the Federal Reserve described the range of balance sheet projections made by firms in CCAR/DFAST 2013: the median firm had projected a cumulative decline in total assets of 3.8% over the nine quarter planning horizon; more than a quarter of firms projected declines approaching 10% (Board of Governors of the Federal Reserve 2013c). The independent supervisory projections described in the letter assumed that credit supply did not contract under stress. Had they been used in CCAR/DFAST 2013, the letter noted that total assets would grow (rather than shrink) a cumulative 2 to 3% over the nine quarter planning horizon.

CCAR/DFAST 2015

The 2015 exercise proceeded without any particular innovations. At its conclusion, the stress test calendar was revised so that the annual cycle would begin with the publication of scenarios in January and results in June.

Quantitative results

In this section, we discuss the scenarios and quantitative projections of losses in the various stress tests that have been conducted since 2009 (i.e. SCAP, CCAR and CCAR/DFAST). Although the scenario used in CCAR 2011 was made public, the results were not. From CCAR 2012 onwards, results have been published in a consistent format.

The stress tests also features two additional notable components: a market shock, and the assumption of the default of a firm's single largest counterparty. These components are applied only to the largest firms. The market shock is designed to stress the trading books of banks with material trading activities and takes the form of moves to roughly 40,000 separate risk factors, such as points on the yield curve, spot and future exchange rates and commodity prices. It is typically based on the market moves that occurred

over the second half of 2008. Firms subject to the market shock (or with significant custodial activities) are required to compute their credit exposures to counterparties under stress and assume that the counterparty to whom they have the largest exposure defaults. These components are beyond the scope of this chapter; we focus here on the banking book and the macroeconomic scenarios.

Scenarios

The macroeconomic scenarios used in the stress tests described here have evolved in response to changing macroeconomic conditions and growing experience with stress tests. We describe in some detail the evolution of two of the most important scenario variables – the unemployment rate and nominal house prices – in each year's most severe scenario.

However, it is worth discussing briefly the experience with the scenarios of intermediate severity. As described in the previous section, the Dodd-Frank Act required the Federal Reserve to describe three scenarios: a baseline and two stress scenarios (labelled the 'adverse' and 'severely adverse' scenarios). In the framework adopted by the Federal Reserve, the intermediate scenario – the adverse scenario – is not intended to be the scenario that results in the binding decrease in capital. Thus it generally features a milder recession than the severely adverse scenario. However, within this constraint, the intermediate scenario has been used to explore qualitatively different types of scenarios. These scenarios have generally focused on interest rate shocks, featuring large increases in long rates, short rates, or both.

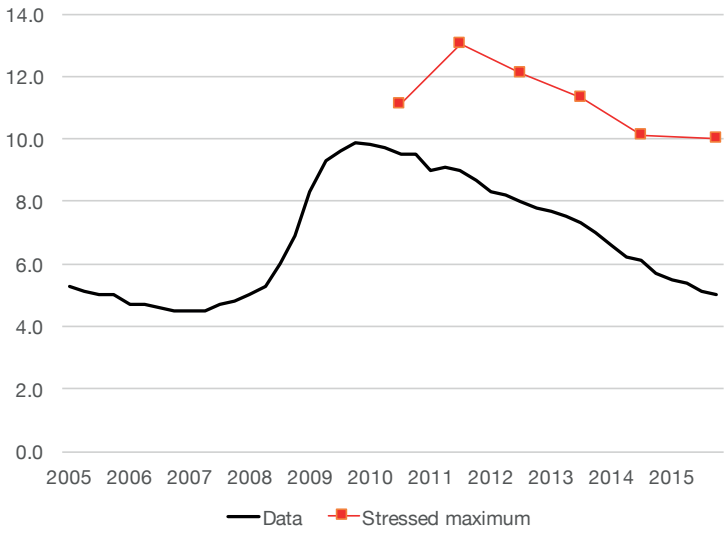
The severely adverse scenario is designed to be the scenario with the largest decline in capital. It features a severe recession in the US amid large declines in asset prices and tight financial conditions. The most recent exercises have described quarterly paths for 28 variables. However, earlier versions specified fewer variables (indeed, the original 2009 exercise only described paths for three variables: GDP, the unemployment rate and house prices).

It is instructive to focus on two variables: house prices and the unemployment rate. These have been described in each year's exercise and the rise in unemployment and

decline in house prices in the 2007–2009 recession were thought to contribute to the outsized bank losses of that period.

Figure 1 compares the actual unemployment rate to the maximum rate in the scenario used in each of the last five stress tests (and the scenarios for CCAR 2016, released in January 2016). The black line shows the actual quarterly US unemployment rate from 2005 to 2015. At the jumping-off points for each of the last five stress tests (that is, 2010:Q3, 2011:Q3, 2012:Q3, 2013:Q3 and 2014:Q3; plus 2015:Q4 for CCAR 2016), the figure plots the peak unemployment rate achieved in the stress scenario used in that year's exercise. Thus, for 2010:Q3, the red point is at 11.1% – the peak unemployment rate in the scenario used for CCAR 2011. Note that, in the scenario, this peak occurs in the sixth quarter of the scenario, or 2012:Q1.

Figure 1 Unemployment rate: Data and stressed maximum



The figure demonstrates important characteristics of the evolution of the scenarios over time. First, the scenario used in CCAR 2011 featured a smaller rise in the unemployment rate than did the scenarios in future exercises.

Second, in keeping with the scenario design framework, the scenarios for CCAR 2012 up to CCAR/DFAST 2015 featured an increase in the unemployment rate of four

percentage points. The increase in the unemployment rate in the scenario for CCAR/DFAST 2016 is larger, at five percentage points.

Third, until the CCAR/DFAST 2016 exercise, the peak unemployment rate under this framework has been declining in lockstep with the improving economy.

Specifying that scenarios will always hit the same *level* of the unemployment rate tends to produce countercyclical scenarios – the unemployment rate rises more when times are good and less when times are bad. Indeed, taken literally, the scenario could feature a declining unemployment rate if the jumping-off rate is above the target peak for the scenario. This emphasises the potential problem with specifying a level – it might not be sufficiently severe to produce credible estimates of 'how bad things could get' in the midst of very severe recessions.

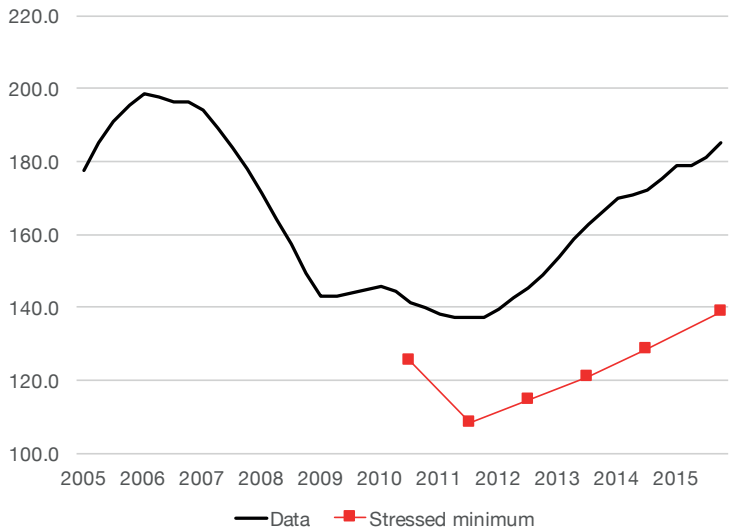
Specifying that scenarios will always feature the same *change* in the unemployment rate guarantees a known level of severity. However, it produces procyclical scenarios (as shown in Figure 1). When the unemployment rate is already high, in bad times, the peak unemployment rate in the scenario will be higher and the test thus, in some sense, tougher. When the unemployment rate is low, in times when financial vulnerabilities are likely building, the peak unemployment rate in the scenario will be lower and the test thus, in some sense, easier.

The scenario design framework envisions a hybrid rule in which the unemployment rate will always rise a certain minimum amount – three to five percentage points – and any additional amount necessary to hit a target peak level of the unemployment rate – 10%. In the first five exercises undertaken since the end of the 2007–2009 recession, the unemployment rate at each year's jumping-off date has been high enough that the increase has been the binding constraint rather than the 10% target. The CCAR/DFAST 2016 scenario, however, features a larger increase in the unemployment rate – five percentage points – than previous exercises, in order to hit the 10% target.

Turning now to house prices, Figure 2 compares the actual house prices to the minimum prices in the scenario used in each of the last five stress tests. The black line shows the actual quarterly house price index from 2005 to 2015. At the jumping-off points for each of the last five stress tests (that is, 2010:Q3, 2011:Q3, 2012:Q3, 2013:Q3 and 2014:Q3; plus 2015:Q4 for the most recent exercise), the figure plots the minimum price index

achieved in the stress scenario used in that year's exercise. Thus, for 2010:Q3, the red point is at an index value of 126, the minimum house price index in the scenario used for CCAR 2011. The black line at that point is at an index value of 142; the scenario features a decline in nominal house prices of 11.3%. (Technically, because house price indexes are revised each time they are published, we computed the percentage decline from the jumping-off point to the minimum in each scenario and applied this to the most recent available trajectories of the index.)

Figure 2 House prices: Data and stressed minimums



As with the unemployment rate, the percentage decline used in CCAR 2011 is smaller than those used in subsequent exercises – the minimum index value in CCAR 2012 is below that for CCAR 2011 even though actual prices had moved up in the intervening year. As with the unemployment rate, the house price path used in the scenarios demonstrates some procyclicality. However, the extent of this procyclicality is less. From 2011:Q3 to 2014:Q3, nominal house prices increased 25%. The minimum levels of the house price indexes used in the scenario have increased less, only 18%.

Formulating a rule that produces countercyclical house price paths encounters the same problems as does a rule for the unemployment rate: the trade-off between a minimum, credible severity and the need to make the path tougher in more buoyant times. In

addition, unlike the unemployment rate, there are only a limited number of cycles to drawn from in understanding the dynamics of house prices.

Projected income and losses

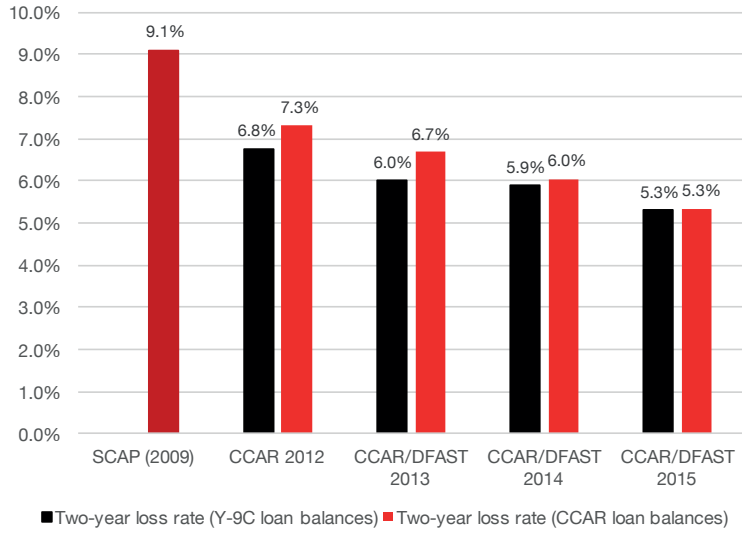
The ultimate objective of the bank stress tests described here is capital. The projected trajectories of capital are a function of banks' pre-provision net revenue, credit losses, losses from other sources such as unrealised gains or losses on securities, and their balance sheets.

Of these elements, credit losses on loans held in the banking book are easiest to compare over time. Indeed, the 2009 SCAP disclosure highlighted that the two-year projected loan loss rate in the SCAP, at 9.1%, exceeded any two-year loan-loss rate recorded, using data back to the 1920s. The methodology used to compute other important elements, as well as their treatment in regulatory capital accounting, changed over time. These elements include losses on securities, the trading shock, and pre-provision revenue.

Figure 3 shows the loan loss rate in each of the last five stress testing exercises as well as 2009 SCAP. The figure shows loss rates computed for the sample of 18 bank holding companies that were part of SCAP and have been in all subsequent exercises. (One of the original 19 firms in SCAP, MetLife, changed its legal organisation and was no longer subject to bank holding company regulation, including the stress tests.)

As shown, loan loss rates have been declining steadily from exercise to exercise. The two-year loan loss rate in CCAR/DFAST 2015 was 5.3%, well below the 9.1% rate from the original SCAP. Although a complete decomposition of the decline isn't available, it seems reasonable that it reflects two forces: decreasing risk on bank balance sheets – as evidenced by the declining delinquency rate on loans for example – and the decreasing peak unemployment rate in successive scenarios.

Figure 3 Loan loss rates for consistent sample of 18 banks

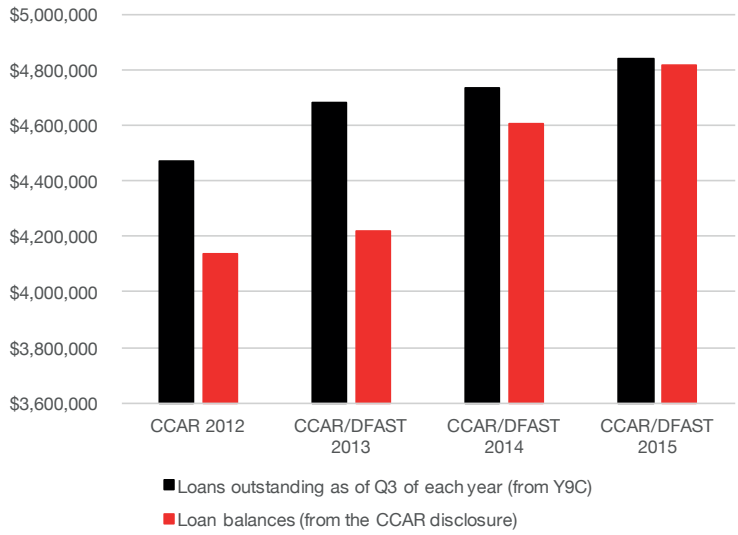


For each year's exercise, the figure shows two loss rates: one computed relative to average loan balances over the scenario horizon, and one computed relative to loan balances outstanding at the end of the third quarter of each year – the scenario jumping-off point. The former rates are reported in the public disclosures, whereas the latter are based on publicly available regulatory reports but have not been highlighted. As described in the various model disclosure documents, loss models typically deliver loss rates, which are then multiplied by balances to arrive at the dollar value of losses. Thus, the average balance projected over the CCAR horizon is the more useful denominator to understand losses.

However, from the perspective of a market participant contemplating the results of a year's exercise, loss rate relative to the balance sheet as it exists in reality, rather than in the hypothetical projection period, is an interesting concept as well.

Figure 4 shows loan balances as reported on the banks' regulatory reports (the Y-9C) as of the jumping-off quarter of each exercise (the third quarter of the previous year), and the average loan balances over the scenario horizon as published in the CCAR disclosure. Actual loan balances are greater than those used in the CCAR projections because they contain assets that are treated elsewhere in CCAR.

Figure 4 Loan balances for consistent sample of 18 banks



Notably, the difference between the two concepts shrinks between the 2013 and 2014 exercises. This is because, as described earlier, starting with the 2014 exercise, the Federal Reserve began using its own projections of firms' balance sheets. That is, the average balances shown for the 2012 and 2013 exercises represent the firms' forecasts, the average balances shown for the 2014 and 2015 exercises represent the Federal Reserve's forecasts.

The effect can be seen in Figure 3. Losses relative to firms' jumping-off loan balances have declined much less than losses relative to the average balances used in CCAR. The reason is that the loss rates projected by the models used in the stress tests are applied to substantially larger balances.

Conclusion

This chapter has described the experience in the US with a regime of supervisory stress tests. Bank supervisors independently project banks' *pro forma* capital ratios under stressful scenarios and tie the decision to object or not to banks' planned capital payouts, in part, to these quantitative results. The regime has evolved over the years as

all parties have gained experience using such a regime. We described the evolution of the macroeconomic scenarios and loan losses.

References

Board of Governors of the Federal Reserve System (2009a), "The Supervisory Capital Assessment Program: Design and Implementation", 24 April (http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20090424a1.pdf).

Board of Governors of the Federal Reserve System (2009b), "The Supervisory Capital Assessment Program: Overview of Results", 7 May (http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20090507a1.pdf).

Board of Governors of the Federal Reserve System (2009c), "SR 09-4: Applying Supervisory Guidance and Regulations on the Payment of Dividends, Stock Redemptions, and Stock Repurchases at Bank Holding Companies", 24 February (http://www.federalreserve.gov/boarddocs/srletters/2009/SR0904.pdf).

Board of Governors of the Federal Reserve System (2010), "Revised Temporary Addendum to SR letter 09-4: Dividend Increases and Other Capital Distributions for the 19 Supervisory Capital Assessment Program Bank Holding Companies", 17 November (http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20101117b1.pdf).

Board of Governors of the Federal Reserve System (2011a), "Comprehensive Capital Analysis and Review: Overview and Objectives", 18 March (http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20110318a1.pdf).

Board of Governors of the Federal Reserve System (2011b), "Capital Plans", *Federal Register* 76(231).

Board of Governors of the Federal Reserve System (2013a), "Capital Planning at Large Bank Holding Companies: Supervisory Expectations and Range of Current Practice", August (http://www.federalreserve.gov/bankinforeg/bcreg20130819a1.pdf).

Board of Governors of the Federal Reserve System (2013b), "Policy Statement on the Scenario Design Framework for Stress Testing", 6 November (http://www.federalreserve.gov/bankinforeg/bcreg20131107a1.pdf).

Board of Governors of the Federal Reserve System (2013c), "Federal Reserve Independent Balance Sheet and RWA Projections", 16 December (http://www.federalreserve.gov/bankinforeg/independent-projections-letter-20131216.pdf).

Greenlaw, D., A. Kashyap, H.S. Shin and K. Schoenholtz (2012), "Stressed Out: Macroprudential Principles for Stress Testing", U.S. Monetary Policy Forum Report No. 5, Initiative on Global Markets, University of Chicago Booth School of Business (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2004380).

Hirtle, B. and A. Lehnert (2015), "Supervisory Stress Tests", *Annual Review of Financial Economics* 7, pp. 339–355.

3 The role of stress testing in supervision and macroprudential policy

Vítor Constâncio
European Central Bank

Stress testing and macroprudential regulation might seem like distinct topics. They are, however, in many ways intertwined. Even so, there is a need to further enhance the macroprudential function of stress testing.

Stress testing: The past and the present

Without the latest crisis that started back in 2007, there would likely not be so much discussion about stress testing. Nor would the area of macroprudential policy be so much under the spotlight.

The crisis, especially in Europe, triggered far-reaching structural changes. Decisive steps taken following the crisis include the creation of banking union in its various dimensions – not least, the establishment of the Single Supervisory Mechanism (SSM) and the corresponding joint approach to micro- and macroprudential policies. This dual approach is reflected, inter alia, in the lasting and standard use of system-wide stress testing.

Stress testing in Europe has changed substantially since the start of the crisis. Back in 2009, the Committee of European Banking Supervisors (CEBS) initiated the first EU-wide stress test. At the time, with its Supervisory Capital Assessment Program (SCAP), the US had also started with its by now regular exercise.

Placed under the aegis of the newly established European Banking Authority (EBA) in 2011, the EU coordinated exercise grew in size and scope, providing a common macro-financial baseline, an adverse scenario, and methodological guidance to the designated participating banks. However, in contrast to the US case, the estimated capital needs of individual banks were not integrated into centralised supervisory decision-making processes, as no supervisory powers had been attributed to the EBA. Any supervisory follow-up deriving from stress test results was left to the discretion of national authorities.

With the establishment of the SSM, now about to celebrate its first anniversary, there was, for the participating countries, a move to a new regime. Before its operational start, a financial health check of the banking sector was required. In synchronisation with the EBA's EU-wide stress test, this exercise was carried out by the ECB in a highly publicised manner. This 'Comprehensive Assessment' (CA) was an exercise of unprecedented scale. The sample covered comprised no less than 130 significant banks in the Eurozone. An additional and essential element was that the CA combined a system-wide asset quality review (AQR) with a stress test.

Owing to the inclusion of an AQR and to an in-depth 'quality assurance' exercise conducted at the ECB, the CA was seen as credible by market participants. This was evidenced by the fact that Eurozone banks' stock prices reflected the CA results as soon as they were published. More discrimination between banks ensued, notably regarding their valuation and risk assessment by the market, reflecting the outcome of the tests. Some further disconnect between sovereign and banks' CDS premia could also be observed at the end of the exercise – a very welcome outcome. Furthermore, the exercise involved a data-rich disclosure with comprehensive balance-sheet data and detailed final results, ensuring greater transparency and facilitating the assessment of banks by observers.

Both the 2014 stress testing exercise, including the tests carried out under the EU-wide exercise coordinated by the EBA, and the 2016 EU-wide exercise currently under preparation have a predominant micro-dimension, even though they are run on a common macro-financial scenario.

These EU-wide stress tests have some specific features worth highlighting. The tests can be described as balance-sheet-based, forward-looking assessments of bank solvency. They span a three-year horizon on a macro-financial baseline and an adverse scenario. The adverse scenario is designed on the basis of the main systemic risks to the banking sector identified as pertinent at a specific juncture. This set of risks is mapped into exogenous shocks. The calibrated shock profiles are then fed into dynamic macroeconometric models used to project macroeconomic and financial variables, which constitute the scenario output.

The framework comprises both exogenous and a form of endogenous shocks. The exogenous components are shocks to stock and bond markets, house prices, short- and long-term interest rates, sovereign bond yields, and to the structure and cost of funding, which are consistently linked to the macroeconomic scenario. The translation of the scenario into variables affecting the valuation of elements of banks' balance sheets has an endogenous dimension. It relates to the impact of the scenario on banks' credit risk (probability of default (PD) and loss given default (LGD)), market risk and other profit components. While the impact of liquidity stress is captured to a certain degree by the funding and liquidity shocks, the exercise remains primarily an assessment of solvency.

Exercises have so far been conducted under static balance-sheet assumptions, meaning that all balance-sheet elements are kept constant throughout the time horizon of the test. This is a simplifying feature, and clearly not very realistic. At the same time, it introduces a conservative bias to the exercise because it does not consider mitigating measures taken by the banks within the time span of the tests. The main purpose of the exercise is to identify banks with a capital shortfall in relation to a pre-determined hurdle rate. The outcome of the exercise is thus a list of capital needs or surpluses per individual bank, reflecting a predominantly microprudential objective.

Now that the SSM is up and running, there is a coordinated supervisory follow-up to the stress test results, since these feed into the SSM Supervisory Review and Evaluation Process (SREP). Following the CA, SSM Joint Supervisory Teams (JSTs) follow up with banks on their individual results, dig deeper into the issue of potential pockets of risk, and decide, where necessary, on additional bank-specific measures. The microprudential function of the ECB/SSM thus uses the results of system-wide stress tests to assess individual institutions and how they manage risk.

This type of stress test is also relevant from a macroprudential perspective for a number of reasons. First, the aggregate results are used for a second feedback exercise regarding their possible macroeconomic impact as assessed by ECB internal models. From this perspective, it may become apparent, for instance, that overly stringent capital requirements may be harmful to the economy to the point of being self-defeating. This could stem from banks' deleveraging by liquidating assets and reducing credit provision in order to meet higher capital requirements. When the economy has just entered the recovery phase or is languishing in a recession, such an approach could be damaging. Second, the test results are also used to assess contagion effects using network analysis. Finally, the overall results can be used in macroprudential analysis, informing discussions on the appropriateness of possible macroprudential measures.

Decisions on the follow-up to a stress test need to take into consideration all types of capital needs resulting from micro- and macroprudential requirements in a coordinated way. It is important to avoid overlaps between macroprudential measures and bank-specific Pillar 2 measures, which use bank-level stress tests as input. Similarly, potential cross-border spillovers of the effects of these decisions should be handled with care. The ECB has specific coordinating powers to deal with issues of cross-border effects and reciprocity.

Looking to the other side of the Atlantic, it is useful to compare the European experience with the US Comprehensive Capital Analysis and Review (CCAR). There are in fact significant methodological and design-related differences. The CCAR covers the largest US banks (approximately 30) and spans a shorter horizon of just nine quarters. The recent version included five different scenarios, of which three were generated by the Federal Reserve System and two were bank-specific and designed by each bank. Six bank holding companies with large trading operations were required to factor in a global market shock as part of their scenarios, and eight bank holding companies with substantial trading or custodial operations were required to assume that their large counterparty had defaulted. The exercise follows a dynamic balance-sheet approach. The Federal Reserve thus runs a stringent quantitative and qualitative assurance process, uses top-down models and may object to the capital plan of a bank even if it satisfies the minimum regulatory ratios throughout the exercise. Results from both the top-down and bottom-up exercises have been published, whereas only the latter are made public

by the EBA. CCAR hurdle rates in 2014 were the regulatory ratios, comprising Basel III transitional arrangements (4.5% CET1), G-SIB additional buffers where applicable, and a leverage ratio. The same hurdle rates applied to baseline and stressed scenarios. This was quite different from the EBA stress test, where hurdle rates were set at a higher rate than minimum requirements (5.5% CET1), which was tightened for the adverse scenario (8% CET1).

The CCAR is thus mostly a 'supervisory exercise'. If a bank fails the test it must submit a new capital plan, which may restrict dividend payments, for instance. In that respect, it differs from the EBA exercises, which can be considered a 'capital exercise'. In both cases, however, the microprudential dimension prevails.

Stress testing methodologies: Treatment of sovereign debt

Rules and methodological aspects need to be consistent and make economic sense within the macro-financial environment in which the stress testing exercises are conducted. A particularly controversial aspect is related to the nexus between banks and sovereigns and the treatment of sovereign exposures. This applies to both exercises conducted in the past and to future ones.

Banks in the EU, which follow the standardised approach for credit risk, benefit from a zero risk weight on sovereign exposures issued in local currency, as well as the exemption of those exposures from the existing large exposures (LE) regime. Banks that are authorised to use the internal risk model approach may also exclude sovereign exposures from risk models and use a zero weight. In the light of recent experiences, there is wide agreement that the existing regulatory treatment needs to be reviewed and the debate is thus underway.

Nevertheless, it should be underlined that in the context of EBA stress tests and our CA, sovereign exposures were adequately subjected to stress. In line with the EBA methodology, fair value positions, i.e. positions classified in the available-for-sale (AFS) category and under the fair-value-option through P&L (FVO), are subjected to market risk stress. As such, in the CA banks faced a large impact on capital on account of the revaluation of sovereign portfolios related to significant sovereign bond yield

increases. More importantly, banks are requested to estimate default and impairment flows for sovereign positions kept as loans and receivables or held-to-maturity (HTM), subject to a set of stressed default and loss rates for each scenario. In the CA, these losses were very substantial.

There is a broader ongoing policy debate about the treatment of sovereign debt exposures. Proposals have been made to impose quotas or quantitative caps on banks' holdings of the domestic debt of the country where they are domiciled. In this context, proposals have been made to use the large exposures regime and apply a cap of 25% of a bank's equity on all exposures to the sovereign.

This approach has serious flaws. First, well-functioning sovereign bond markets are essential for the functioning of financial markets and the pricing of risk. Banks hold sovereign exposures for a number of reasons, including their favourable credit and liquidity characteristics and use as collateral. An alternative to setting a hard limit for sovereign exposures is to use the pricing mechanism by introducing risk weights, which should be less intrusive and impose less distortion in financial markets. A price-based regulation would allow for more flexibility in banks' portfolio decisions and would also be less disruptive to the sovereign debt markets. Finally, preliminary analysis suggests that setting a hard limit under a large exposure regime of 25% of own funds would mean that, in order to accommodate their present holdings of domestic sovereign debt, EU banks would need additional Tier 1 capital of over €6 trillion or potential sales of those securities above €1.6 trillion. This is illustrative of the scale of disruption in sovereign debt markets, the banking sector and the economy as a whole that the introduction of such a regime would be likely to bring.

Some economists argue that the quantitative limit would generate welcome diversification by creating an active market of asset swaps among banks, with the banks from core countries taking debt from peripheral countries. However, the functioning of the sovereign debt market in past years raises doubts as to whether such a scenario could ever materialise.

The discussion is motivated by the need to put an end to the sovereign-bank nexus. It is worth recalling that the risk transmission between banks and sovereigns runs both ways.

It does not take much to conclude that, since the beginning of the crisis, government support to banks (in the form of capital injections or guarantees) has far exceeded total bank losses on sovereign exposures. To date, these losses relate only to the restructuring of private debt in Greece in 2012.

Banks' losses in the crisis derived primarily from excessive risk and concentration on residential and commercial real estate exposures, including to US underpriced and toxic assets. Should quantitative limits also be applied to such exposures? Regarding the feedback loop with sovereigns, the overwhelming bulk came from the bailout of banks, amounting in gross terms to €800 billion incurred by sovereigns (ECB 2015), dwarfing the €50 billion that banks lost from their Greek sovereign holdings. The Bank Recovery and Resolution Directive (BRRD) and the new bail-in regime in force as of 2016 have fully addressed this source of the feedback, as bailouts can no longer take place.

Finally, as referred to in the recent *Five Presidents' Report* on the future of the Monetary Union (Juncker 2015), any new regulation on sovereign debt risk weights should have a true worldwide agreement in the Basel Committee context. This could lead to a reasonable solution along the lines of a price incentive approach by which sovereign debt risk weights are increased. This approach should be sufficient to address the real concerns surrounding sovereign debt, also bearing in mind that sovereign debt will continue to be appropriately stress-tested in EU-wide exercises.

Alternative approaches to stress testing

Before moving on to examine the future of stress testing, it is worth reflecting on alternative approaches to the balance-sheet-based stress tests conducted by official authorities. The main alternative – known as 'market-price-based stress tests' – can be quite appealing given its simplicity, low data intensity and automaticity. At the same time, it places great emphasis on the wisdom or whims of the market. It can indeed be argued that these exercises are not usable for policymaking purposes and can create a dangerous false sense of comfort in quiet times.

They rely on market capitalisations and other market-based indicators to measure the safety and soundness of the financial sector. Quantitative analysis based, for instance, on the SRISK measure (Acharya et al. 2012), among others, was used to challenge the results of the CA in 2014, by comparing them to alternative stress test simulations based on market data. In October 2014, market-based metrics provided substantially higher estimates of capital shortfalls than the SSM CA results (Steffen and Acharya 2014, Steffen 2014).

The SRISK approach determines a bank's capital shortfall as a function of i) a bank's market capitalisation, ii) its asset size, and iii) the relationship of the bank's stock returns with the aggregate market returns.

The relevance of such approaches to support policy decision-making is questionable. They are detached from the macroeconomic and financial narrative that helps us lay the foundations for macroprudential policy and macroeconomic scenarios and their specific shocks. They are thus not comparable with the balance-sheet-based stress tests conducted by the EBA or the CCAR. The estimates vary widely, depending on the definition of the capital ratio, the reference threshold, and the underlying stress assumptions which are not linked to a macroeconomic scenario.

Additionally, the approach is subject to substantial volatility on account of its market-based dependence. Resulting capital requirements could vary substantially in a matter of months, according to market fluctuations, making the results unusable for either management or the authorities' purposes. At the same time, in periods of relative calm in the financial markets, such measures would provide false comfort – which is precisely what would need to be counteracted through the use of pre-emptive macroprudential policy. In contrast, during episodes of heightened market uncertainty, such measures would be likely to exacerbate fear and distress, pointing to astonishingly low bank market capitalisation levels and big capital shortfalls. We should view this kind of procyclicality with great concern and avoid taking policy decisions that would further amplify, rather than smoothen, the financial cycle.

Stress testing: The limitations of the current methodology

As can be seen from the description above, it is clear that system-wide stress testing exercises in the EU, conducted under the present methodology, have a number of limitations. This concerns, in particular, their macro dimension and their use for macroprudential policy purposes.

Featuring prominently among these limitations is the static balance-sheet approach, which is not suited to exercises that run for a horizon of three years. This may render the tests unduly conservative if the macro scenario is too severe.

A related limitation is the fact that no bank reaction is considered. It would be far more realistic to assume that market participants could react to adverse conditions, rather than assuming passive bank behaviour throughout the entire stress test period. Bank behaviour or reaction could take the form of deleveraging, straight capital increases or working out of non-performing loans.

While the EBA exercise entails some liquidity stress in the form of market liquidity (the risk that an asset cannot be sold or used as collateral) or funding liquidity (the risk that a bank is unable to roll over maturing funding), no thorough liquidity assessment is conducted. Given the strong, two-way interaction between liquidity and solvency stress, which was brought to the fore by the Global Crisis, proper liquidity stress testing should constitute an inherent part of the solvency stress testing framework.

Another limitation relates to the absence of interaction between banks and other specific sectors of the economy, notably the household and non-financial corporate sectors' balance sheet and financial positions. In an integrated model framework, these balance sheets would also be affected and would respond to the various macroeconomic shocks in the stress test scenario. In the same vein, the framework does not capture second-round effects and subsequent feedback effects within the financial sector and in the real economy at large.

All such features can clearly only be present in a centrally conducted, top-down stress test framework. This kind of framework would be difficult to adopt in the context of the EBA EU-wide stress tests, where bottom-up results are provided by the banks and cover a large sample. Top-down macro stress tests are thus a powerful tool that can be

employed in a range of exercises from the simplest – aimed at evaluating the direct impact of stress on each bank – to the most complex, which include a dynamic set-up and are combined with a macroeconomic model, taking into account second-round effects. They would have the potential to assist macroprudential policymakers in the design, calibration and assessment of the impact of macroprudential tools.

Macro stress tests: The future of the ECB framework

Macro stress tests are an indispensable tool of macroprudential policy. However, to be policy relevant, exercises need to be embedded in a coherent macroeconomic and financial set-up and must incorporate a macroprudential perspective. Overcoming the drawbacks highlighted earlier is a first step, and work in that direction is underway at the ECB (see also Figure 1). After incorporating the dynamic elements and the necessary feedback with the macroeconomy, we need to enlarge the framework to assess the impact of policy tools and expand the range of macroprudential tools used in the models. Developing proper indicators that can measure the system-wide level of systemic risk resulting from each scenario and the shocks considered is even more challenging. These indicators should relate to the various sources of systemic risk: macroeconomic shocks, endogenous financial imbalances, and contagion effects dependent on interconnectedness.

Staff at the ECB have been working on the characterisation of the financial cycle, which depends on several variables related to credit, leverage and selected asset prices, notably housing and bond prices. Going forward, the macro and financial variables – as they evolve over the stress testing horizon, including feedback effects – could be used in a framework to assess systemic risk levels over the financial cycle. This could provide valuable guidance for pre-emptive macroprudential policy decisions.

In 2013, the ECB published an occasional paper describing the framework it has developed to conduct stress tests and its various modules (Henry and Kok 2013). These had been used to support the CEBS, and later the EBA exercises. Since then, we have added new modules and tools (see Figure 1 at the end of this section), a brief description of which follows.

Regarding the dynamic treatment of certain variables, considerable progress has been made in accommodating the dynamic balance-sheet approach for macroprudential stress tests; in other words, the capacity to take into account banks' reactions to the stress. In our framework, the dynamic balance-sheet approach is implemented by allowing banks to re-optimise their portfolios according to the risk-return optimisation criterion (Hałaj 2013). 'Reaction functions' could also be informed by bank surveys. Typically, under adverse scenarios, reactions generate a negative feedback loop, whereby deleveraging leads to an aggravation of the initial stress. At the same time, bank results tend to be less acute, given the more realistic possibility to react to shocks.

Household sector

Concerning the attempts to assess household sector vulnerabilities, we have developed a framework for stress testing the balance sheets of individual households using the data from the ECB Household Finance and Consumption Survey (Ampudia et al. 2014). This approach, linking micro and macro data, has already allowed us to compute PDs and LGDs for mortgage exposures directly at the household sector level and link them to macroeconomic stress scenarios. This gave us valuable insight into heterogeneity across the Eurozone countries and non-linearities in responses to various macroeconomic shocks owing to different distributions of debt burden across households. We are currently working on extending this framework to account for a consistent stress scenario, as well as allowing for dynamic adjustments of individual households' balance sheets in response to shocks, and related second-round effects (Gross and Poblacion 2016).

Liquidity

The Global Crisis revealed major transmission channels from liquidity and funding to bank solvency, and vice-versa. Efforts are thus underway to integrate a macro liquidity stress test into the solvency stress testing framework.

Liquidity stress testing should be a tool to assess banks' capacity to withstand extreme liquidity shocks by looking at their liquidity positions. This goes far beyond compliance with the regulatory liquidity ratios such as the Liquidity Coverage Ratio (LCR). It

comprises measures of the depth, breadth and horizon of a shock, and takes into account the fact that liquidity stress usually unwinds very quickly and turbulently.

Brunnermeier et al. (2012) have proposed a new system of measuring risk and liquidity in the financial sector and recommend that policymakers focus on the risk topography of the economy. Their idea is to regularly obtain from financial companies the maturities and liquidity sensitivities of their assets and portfolios and build a total firm-specific liquidity indicator. Though very useful, calculating such an indicator on a regular basis would be demanding for the reporting institutions. Nevertheless, the implementation of their proposal in the context of point-in-time stress tests could be feasible.

Proper liquidity stress testing should constitute an inherent part of the solvency stress testing framework. Irrespective of the origins of the shock to the financial sector, which can be multiple and largely unpredictable, the major propagation channels between liquidity and solvency are broadly common and could be integrated within a liquidity stress-testing framework.

This nexus of solvency and liquidity includes four important propagation channels that we model (or intend to model) in our framework:

i. A fire-sale externality: Fire sales emerge because, during times of stress, incomplete markets impact upon the solvency of banks via realised losses owing to asset disposals at fire-sale prices, and mark-to-market losses on liquid assets held at fair value.

ii. Margin calls and the closure of funding markets: Negative feedback loops are triggered by illiquid banks calling in interbank facilities with other banks or raising margin requirements in illiquid repo or derivative markets. This may lead to increasing funding costs or even the closure of the funding markets.

iii. Credit rating: A deterioration of a bank's credit rating triggered by a worsened capital position may lead to higher funding costs, which, *ceteris paribus,* leads to a further deterioration of the solvency position.

iv. Asset quality: A worsening of the credit quality of assets leads to a worsening of the cash-flow (as non-performing loans do not generate cash-inflows), which in turn leads to an immediate deterioration of the liquidity and funding position.

We are aware that liquidity risk – in particular the fire-sale externality – cannot be assessed in isolation, but only within the whole banking sector. This will be explored further below.

Contagion

Financial contagion is another important element of a macroprudential stresstesting framework. In the 2007-2008 crisis, contagion via direct and indirect channels was key in propagating the initial, relatively small, subprime shock. In our top-down stress tests, we have been taking account of direct contagion via the interbank channel for many years now. The second-round effects of contagion are regularly published as part of our assessment of the resilience of financial institutions in the semi-annual ECB *Financial Stability Review*. However, by taking full account of contagion, we are expanding our framework to encompass indirect contagion as well. In particular, we are modelling the impacts of fire-sales[1] with the ultimate aim of making endogenous the price response within an integrated bank-shadow bank stress test framework using an agent-based modelling framework.

Interaction with the real economy

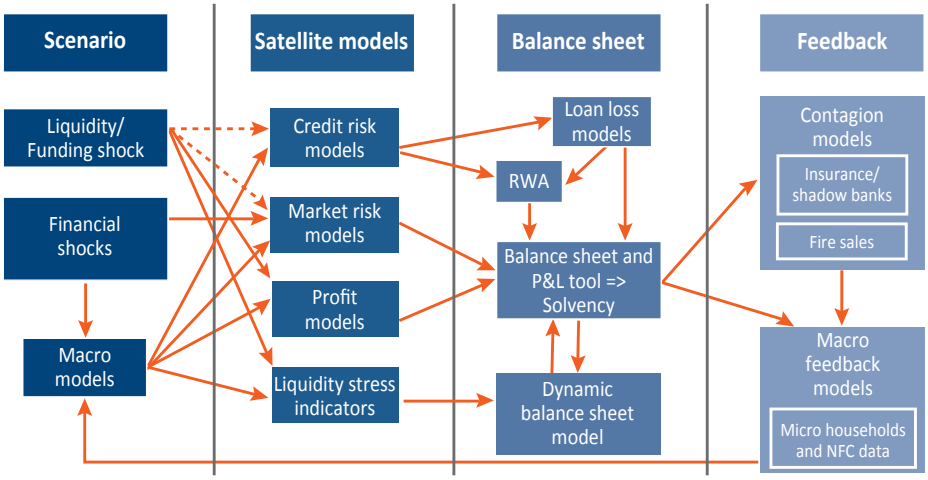
Macro-feedback analysis, which takes account of interactions with the economy, plays an important role in macroprudential stress tests. To this end, for the Eurozone, we have calibrated country versions of a stylised DSGE model (Darracq Paries et al. 2011), which we use to assess post-stress general equilibrium dynamics in each Eurozone country. Moreover, to take better account of the banking sector and default externalities, the ESCB's Macroprudential Research Network (MaRs)[2] has developed a general

1 Along the lines of Greenwood et al. (2015) or Cappiello and Supera (2014).

2 See ESCB Heads of Research (2014).

equilibrium model with a three-layer endogenous default, in the household, corporate and banking sectors (the '3D model') (Clerc et al. 2015). Country-specific versions of this model will soon be available for the regular policy evaluation. The macro-feedback loop has also been beefed up by econometrically testing banks' responses to capital shocks in time-series models, including GVAR[3] and FAVAR[4] models. These tools can also be used for macroprudential policy cost-benefit analyses, as well as for capturing cross-border effects. We are also working on integrating the GVAR with the early-warning models to take account of the medium-term boom-bust cycles by making endogenous the predictor variables used in the logistic early warning model (Behn et al. 2015).

Figure 1 The four pillar structure of the ECB solvency analysis framework



Note: "RWA" refers to risk-weighted assets.
Source: ECB.

3 See Gross et al. (2013), Gray et al. (2013) and Gross and Kok (2013).

4 See Budnik et al. (2015).

Stress testing of non-banking sectors

Let us now turn to an important aspect that needs to be taken into account for the purposes of building a fully-fledged stress testing framework, namely, the need to have a stress testing framework that integrates banks and the shadow banking sector.

With a steady growth in assets and the potential for substituting financial services from the regulated sector, shadow banks also constitute a challenge for any macroprudential policymaker. Recent ECB research suggests that the shadow banking sector has a natural tendency to grow until it becomes systemically important for the entire financial system and endangers the stability of the banking sector (Ari et al. 2015). To unveil the vulnerabilities in this sector and to identify the pockets of illiquidity and better understand the dynamics of the interactions of the sector with the rest of the financial sector, we need a stress testing framework that is conducted by the firms themselves but which is in line with guidance issued by the competent regulators, as recently recommended by the Financial Stability Board.[5]

Such a framework could start by assessing the resilience of the largest shadow banks to a number of stress factors, in particular to various asset price shocks, as well as their resilience to the materialisation of redemption risk. This should include a simulation of fire sales that accounts for the depth and liquidity of various asset markets. Ultimately, the shadow-banking stress testing framework should account for various layers of interconnectedness and identify direct and indirect contagion channels, thus requiring their integration in the banking sector stress test framework.

As a starting point for a financial system stress test which involves the firms themselves, 'guided' stress tests of non-banking entities, notably large asset-management entities, could play an important role, even if they were not yet fully integrated with banking sector stress tests.

To complete this picture, note should also be taken of stress tests on central clearing counterparties (CCPs). A proper macroprudential stress test of these institutions

5 See the press release for the Meeting of the Financial Stability Board in London on 25 September 2015 (http://www.financialstabilityboard.org/2015/09/meeting-of-the-financial-stability-board-in-london-on-25-september/).

should not only include solvency and liquidity stress, but should also account for interconnectedness via common exposures to clearing members as well as possible knock-on effects on the banking sector in the event that the guarantee fund of a CCP is wiped out and clearing members are required to cover the CCP's losses. In this vein, the ECB is contributing to the European Securities and Markets Authority's CCP stress tests by helping with the scenario design and the analysis of the network effects.

Conclusion

The standard use of system-wide stress tests has been an important consequence of the financial crisis. So far, the tests have focused mainly on banks and their solvency, reflecting a predominantly 'micro' supervision perspective. The macroprudential function has added a new dimension to stress testing. It necessitates going well beyond summing up the individual banks' results. Such a step is necessary to properly assess the macroeconomic impact of stress test results. As a by-product, the resulting enhanced stress testing capabilities should enable the quantitative assessment the impact of macroprudential measures. The underlying framework has to embed spillovers – within the banking sector and to other sectors, including the real economy – and allow for banks' own reactions, which can also spill over to other segments of the economy.

A new framework for stress testing must therefore combine the objectives of the two policy functions – micro and macro supervision – in a complementary manner. It is a challenging objective, but one the ECB is actively pursuing.

References

Acharya, V., R. Engle, and M. Richardson (2012), "Capital Shortfall: A New Approach to Ranking and Regulating Systemic Risks", *American Economic Review: Papers & Proceedings* 102(3).

Ampudia, M., H. van Vlokhoven and D. Żochowski (2014), "Financial fragility of euro area households", ECB Working Paper Series No. 1737, ECB, Frankfurt.

Ari, A., C. Kok, M. Darracq Paries and D. Żochowski (2015), "Shadow banking in general equilibrium", mimeo, ECB, Frankfurt.

Behn, M., M. Gross, T. Peltonen (2015), "Assessing capital-based macroprudential policy using an integrated Early Warning GVAR model", mimeo, ECB, Frankfurt.

Brunnermeier, M., G. Gorton and A. Krishnamurthy (2012), "Risk Topography", *NBER Macroeconomics Annual* 26, pp. 149-176.

Budnik K., L. Cappiello, H. Kick and G. Nicoletti (2015), "Financial shocks, bank balance sheet variables and financial spill-overs in the euro area", mimeo, ECB, Frankfurt.

Cappiello L. and D. Supera (2014), "Fire-sale externalities in the euro area banking sector", *Financial Stability Review*, ECB, Frankfurt.

Clerc, L., A. Derviz, C. Mendicino, S. Moyen, K. Nikolov, L. Stracca, J. Suarez and A. P. Vardoulakis (2015), "Capital regulation in a macroeconomic model with three layers of default", *International Journal of Central Banking,* June.

Darracq Paries, M., C. Kok, and D. Rodriguez Palenzuela (2011), "Macroeconomic propagation under different regulatory regimes: evidence from an estimated DSGE model for the euro area", *International Journal of Central Banking*, December.

ECB (2015), "The fiscal impact of financial sector support during the crisis", *Economic Bulletin* 6, ECB, Frankfurt.

ESCB Heads of Research (2014), *Report on the Macro-prudential Research Network (MaRs),* Frankfurt: ECB.

Gray, D., M. Gross, J. Paredes and M. Sydow (2013), "Modeling Banking, Sovereign, and Macro Risk in a CCA Global VAR", IMF Working Paper 13/218, Washington, DC.

Greenwood, R., A. Landier and D. Thesmar (2015), "Vulnerable Banks", *Journal of Financial Economics* 115(3), pp. 471-485.

Gross, M. and C. Kok (2012), "Measuring contagion potential among sovereigns and banks using a mixed-cross-section GVAR", ECB Working Paper Series No. 1570, Frankfurt.

Gross, M. and J. Poblacion (2016), "Assessing the efficacy of borrower-based macroprudential policy using an integrated micro-macro model for European households", ECB Working Paper Series No. 1881, February.

Gross, M., C. Kok, and D. Żochowski (2015), "The impact of bank capital on economic activity: Evidence from a Mixed-Cross-Section GVAR model", mimeo, ECB, Frankfurt.

Hałaj, G. (2013), "Optimal asset structure of a bank - bank reactions to stressful market conditions", ECB Working Paper Series No. 1533, Frankfurt.

Henry, J. and C. Kok (eds) (2013), "A macro stress testing framework for assessing systemic risk in the banking sector", ECB Occasional Paper No. 152, Frankfurt (available at http://www.ecb.europa.eu/pub/pdf/scpops/ecbocp152.pdf).

Juncker, J.-C., in close collaboration with D. Tusk, J. Dijsselboem, M. Draghi and M. Schulz (2015), *Completing Europe's Economic and Monetary Union*, Brussels: European Commission.

Steffen, S. (2014) "Robustness, Validity and Significance of the ECB's Asset Quality Review and Stress Test Exercise", SAFE White Paper No. 23, Goethe University Frankfurt.

Steffen, S. and V. Acharya (2014), "Making Sense of the Comprehensive Assessment", SAFE Policy Letter No. 32, Goethe University Frankfurt.

4 The Bank of England's approach to stress testing the UK banking system[1]

Alex Brazier
Bank of England

Last week, we at the Bank of England launched our new approach to stress testing the UK banking system.[2]

We've published the approach, which has been evolving with our experience of stress testing, to give banks, markets and others as much clarity as possible.

And as I'll come on to explain, clarity about our approach has the potential to make stress testing an even more powerful tool to ensure banks are safe and sound and that the banking system has the strength to support the real economy in all weathers.

Our approach has three core elements – the three Cs:

It's **countercyclical**, and systematically so. In general, it will be more severe in the good times and less so in the bad.

It's **consistent** with the capital framework for banks – an integral part of it, rather than an alternative.

And in assessing the results, it's **catholic** in approach **–** a broad church of banks' own models, as well as our own.

1 Speech given at the LSE Systemic Risk Centre, 30 October 2015.

2 Available at http://www.bankofengland.co.uk/financialstability/Documents/stresstesting/2015/approach.pdf

What is stress testing and why do it?

But before I elaborate on those, allow me to take a step back and ask: What do we mean by stress testing and why are we doing it?

We mean assessing the impact on bank capital of a hypothetical bad economic scenario. Across all the major banks at the same time, so we can assess strength of the system, as well as its parts. That marks a real departure from before the crisis, when there was, in the UK, no systematic approach to stress testing the banking system.

We want to know the damage the banking ship would sustain in an economic storm. But we're not just interested in whether banks stay afloat. We want the system to be strong enough to continue to serve the real economy, even in the storm. In this year's stress test, that means expanding lending to the UK economy by 10% over the next five years, even in the face of a synchronised domestic and global slowdown.

And we want to look forward in making that assessment, to where banks plan to be in the future, not just where they happen to stand today, and to the storms that could break tomorrow, not those that drenched us yesterday,

But why is stress testing needed to ensure banks have the requisite strength? Aren't the baseline Basel III capital requirements sufficient to do the job?

After all, once fully implemented, the most systemically important banks could face capital requirements that are ten times tougher than they were before the crisis.[3]

These requirements are designed to reduce the probability of a systemic banking crisis to a very low level. There is a reasonable debate to be had about whether a little more might be necessary given the costs of such crises. But with the advent of effective resolution regimes for failing banks and new requirements for loss absorbing capacity in resolution, which will dramatically reduce the economic consequences of failure, these requirements are not in the wrong ballpark.

3 See Caruana (2012).

However – and this is crucial – they are not designed for unusually risky banks in unusually risky situations. In particular, they are not designed for the more risky stages of the economic and financial cycles.

At the apex of such cycles, to be confident of being able to able absorb possible losses and resist crunching credit supply, banks would need to have capital ratios that are materially higher.

Requiring those sorts of capital ratios – which are not without their advocates – at all times would certainly be prudent.[4] But they would not make macroeconomic sense.

A banking system capitalised at all times for the peak of the cycle would result in inefficiently high borrowing costs for much of the time. That would harm the productive potential of the economy.

So we need a capital regime that is both prudent and makes macroeconomic sense: a capital regime that is macroprudential.

That can be achieved by varying capital requirements over time as conditions change. When the environment is unusually risky, including in the later stages of a financial cycle, banks should hold more capital. But when it's not, they shouldn't.

Though it is far from the only one, stress testing can be a powerful tool for turning that principle into practice.

It tells us whether banks are taking ever bigger bets on the economy. And if the stress scenario can capture the risks banks face, the test can help to calibrate whether, and how much, additional capital is needed for the current risk level.

Stress testing can help us to match the strength of the system to the degree of risk it faces.

4 A number of countries have exceeded minimum Basel III capital and/or leverage ratio requirements for systemically important banks, for example: Australia, Austria, Canada, Denmark, Finland, Hong Kong, Netherlands, Norway, Singapore, Sweden, Switzerland and United States.

This is an increasingly important exercise. In the aftermath of a crisis, the financial cycle has been in a muted phase. But as it progresses into a more normal phase, we will have to devote much more attention to whether capital requirements should be increased to keep pace with the risks facing the system.

Countercyclical

Which brings us to the first C of the new approach: **countercyclical**.

The stress scenario will become more severe as the risks get bigger and less severe as those risks either materialise or shrink.

This is ambitious and ground breaking. And if past cycles are anything to go on, we could be making the test more exacting just as lenders and their investors are thinking the world is *less* risky – that this time really is different. It will not be easy.

So how are we going to do it, first in principle, then in practice?

In principle we are making the design of the stress test scenario an exercise comparable to economic forecasting.

Forecasting typically amounts to making a projection that is in the centre of the distribution of possible future outcomes – a projection that, as time unfolds, is as likely to be too optimistic as too pessimistic.

Producing a stress scenario – in our approach – is also making a projection. But it is a projection not of a point near the centre but of a point in the tail of the distribution of possible future outcomes. In Figure 1, the central projection is point X; the tail projection – the stress scenario – is something like point Y.

It is a projection very likely to prove to be too pessimistic. The probability of it being too optimistic will be aligned with our appropriately low risk appetite for banks finding themselves so weakened that they crunch credit supply to the economy.

To produce that tail projection, we'll begin with the central projections made by the Bank of England and IMF of the UK and global economies. By using distributions

of past outcomes and analysis of forecast errors, we'll take the first step to making a projection of the point in the tail.

Figure 1 Stylised conditional distributions of expected movements in a variable

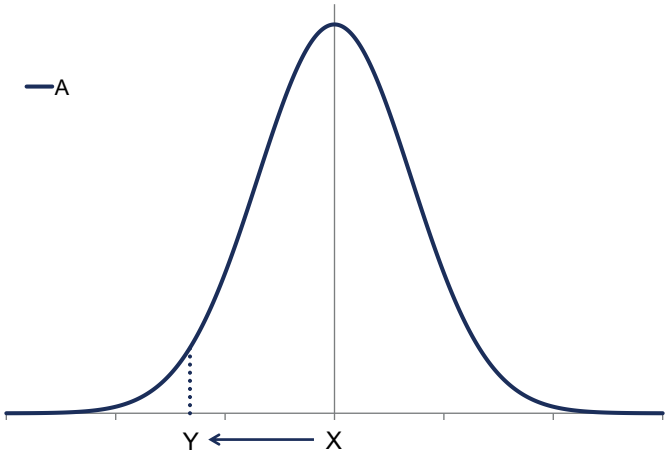
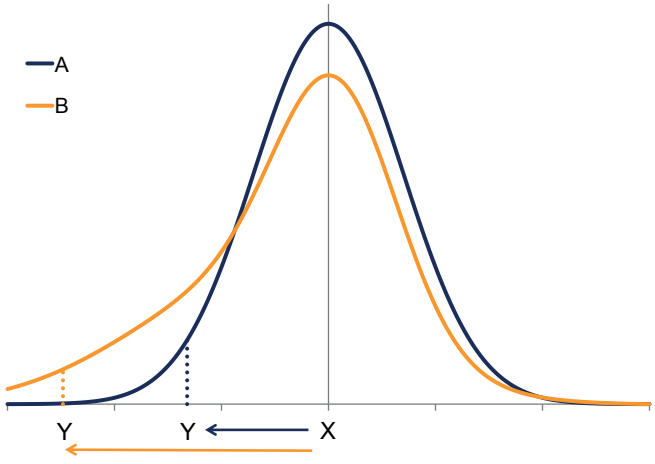


Figure 2 Stylised conditional distributions of expected movements in a variable



But it's the second step that's even more interesting, and even harder. We'll need to question whether the skew in the distribution of possible future outcomes is bigger than usual – more like profile B than profile A in Figure 2. If it is, the stress scenario – the same point Y in the tail of the distribution – should be more severe, as shown in Figure 2.

To gauge this, we will assess indicators of *potential* imbalances – in credit, in asset prices, and in household and corporate balance sheets. The emphasis is on potential because the exercise does not rest on spotting imbalances with any certainty. That is the strength of this approach. Adjusting the projection of a point in the tail of the distribution requires only a *possibility* that an imbalance *might* be present.

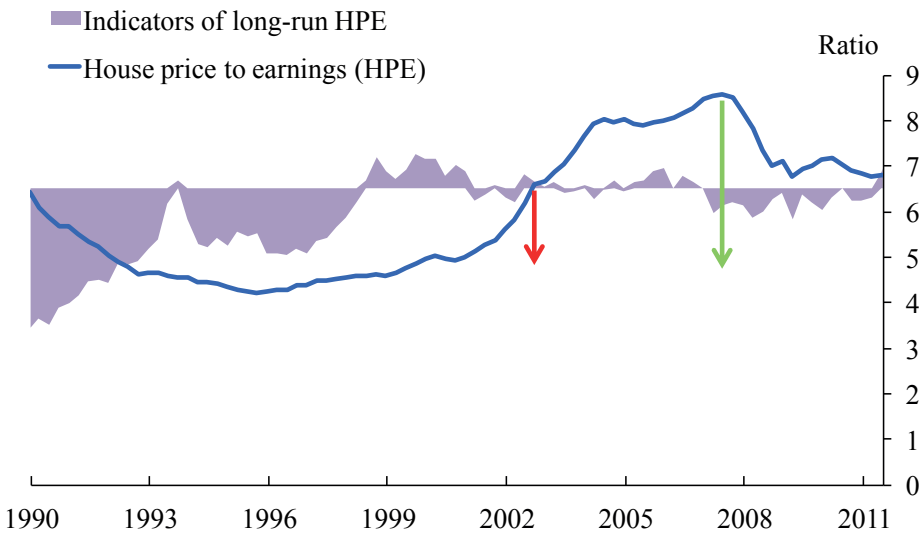
And where that possibility exists, we'll assess how, in the past, imbalances have affected the scale of any subsequent correction.

That's the principle. How might it have worked in practice if stress testing, and this approach to it, had been in place in the run-up to, and aftermath of, the financial crisis?

I'll use the example of the housing market, treating it in isolation for simplicity.

Figure 3 shows UK house prices relative to average earnings in that cycle. It also shows an estimate for what we might have expected it to be, given rents and interest rates.

Figure 3 UK house price to earnings ratio



Notes: House price to earnings ratio calculated as the ratio of the average nominal house price and average annual earnings. Indicators of long-run HPE include post-1990 median HPE and HPE estimated using a dividend discount model (DDM) that takes account of levels of rents and interest rates.

Sources: Bank of England, Halifax, Nationwide, ONS and Bank calculations.

At the turn of the century, house prices were aligned with these metrics. There wasn't an indication of an imbalance. But even from such an apparently balanced position, history tells us that shocks happen. There is always a distribution of possible future outcomes.

The stress scenario – the projection of the point in the tail of the distribution – might have included a double-digit fall in house prices relative to earnings.

As the decade progressed, house prices continued to outpace earnings, and by more than could be explained by rents and interest rates. An imbalance was possibly emerging and the probability of a correction growing.

So the projection of the same point in the tail of the distribution would have been more severe.

To calibrate how much, we look at previous housing cycles and assess how their severity relates to the starting imbalance.

At what we now know to have been the peak of the cycle, the stress scenario would have included a fall in house prices, relative to earnings, of almost 30% – twice as severe as earlier in the cycle.[5]

House prices then fell by almost 20%. And at that point, the severity of the test would have reverted back to where it started.

Now, projecting the middle of the distribution is hard enough. Projecting the tail is, as Sir Humphrey might say, 'very brave'. By their very nature, there isn't a large sample of tail events to go on.

But we've adopted this ambitious framework for two reasons.

5 This illustration treats the housing market in isolation. A real stress test would also include the knock-on effects of other economic imbalances. The 2014 stress test, for example, assumed that the UK current account closed and long term interest rates 'snapped back'. The knock-on effect of both of these elements of the stress scenario was incorporated in the scenario for the housing market, which included a fall in house prices of 35%.

First, the stress test is informing bank capital requirements, the setting of which has real economic consequences. It is incumbent on us not simply to ratchet up – or down – the severity of the stress scenario over time.

As policymakers, we have a responsibility to ensure that capital requirements change only because the risks change. As we move from stress test to stress test, the responsibility will be on us to explain the indicators that have caused us to change the severity of the scenario.

The side of the bed we exited on the morning of deciding the scenario should not feature in that dataset. It hasn't before. And this approach ensures it won't in future.

There is a second reason – more selfish on our part – for setting out a systematic countercyclical approach.

By looking at the economic situation, banks will increasingly be able to anticipate our actions. Although it can never be mechanical and entirely predictable, our aim is for changes in the severity of the test to become more predictable.

With that, they can adjust their capital and business plans accordingly. They don't need to wait for the results of the test. They don't even need to wait for the test scenario to be announced.

Systematic policymaking can shape the behaviour of the banking system, hardwiring into its DNA a capital strength that counters the cycle.[6]

It is a big investment that will take many years to bear full fruit. But the prize is great. The development of systematic monetary policy in the past 20 years shows quite how great.[7]

That is not to say we are throwing away any flexibility to test risks that are not part of the economic or financial cycle. A key element of our new approach is the addition of a second stress scenario in every other year.

6 By being more predictable, we can also help to ensure that, when more capital is needed in the system, it is added over a measured timeframe, not at the last minute, reducing any costs of transition.

7 See Woodford (2003), for example.

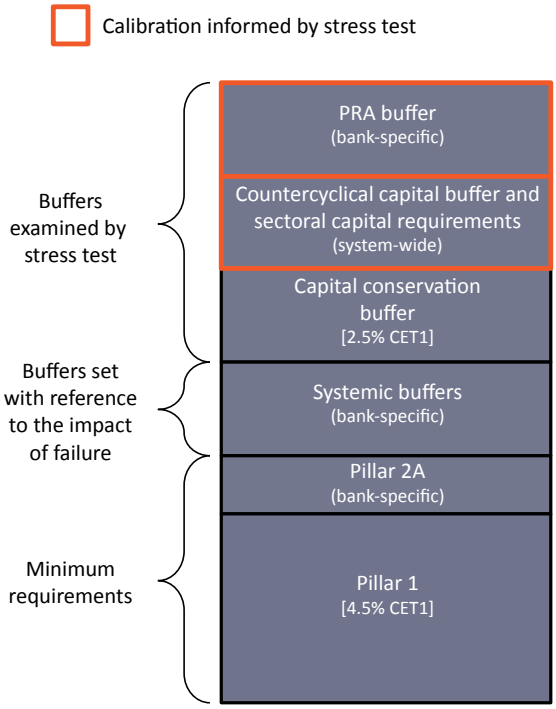
That 'exploratory' scenario is there to give us the opportunity to poke around in the corners, to test more structural risks to which UK banks might be exposed – risks that are not well captured by the state of the economic cycle. Those might include structural changes in overseas policy regimes, the possibility of sustained deflation and low interest rates at the global level, or changes in particular industries and financial markets.

That biennial exploratory scenario ensures we retain sufficient discretion to keep up with all the risks we might want to test, while at the same time ensuring consistency through time in our annual scenario to capture the state of the cycle.

Consistent with the capital framework

The second C of our approach is that it's **consistent** with the capital framework.

Figure 4 Components of the regulatory capital framework



Our approach to stress testing is not about building an alternative framework of capital regulations. We are integrating stress testing into the framework that's already in place.

Here's how.

Capital requirements fall into two sets: minimum requirements that must be held in all circumstance, however stressed, and buffers on top of that minimum that can be depleted to absorb and cushion the impact of stress. That 'capital stack' is shown in Figure 4.

Those buffers really are there to be used. It would be self-defeating to require them to be maintained at all times or to be rebuilt at breakneck speed. Their usability is vital to avoiding the crunching of credit supply to the real economy.

Stress testing informs how big that buffer needs to be to deal with future storms. As risks grow and storm clouds gather, the stress scenario will get more severe and buffers will increase. If the storm hits, the buffers will be used.

Matters are complicated by the fact that the total capital buffer for each bank has many components.

All banks have a capital conservation buffer.

That can be augmented across all banks by macroprudential authorities (in the UK, the Bank of England Financial Policy Committee).

And it can be augmented further on a bespoke basis by microprudential supervisory authorities (the Prudential Regulation Authority of the Bank of England).

In the UK, we're fortunate. All the authorities sit in one place – the central bank – so the coordination to avoid over and underlap is quite achievable. But for the Kremlinologists, here's how we'll do it.

First, the macroprudential authority will identify in the bank-by-bank results of the test the common system-wide impact.[8]

8 This exercise will draw on the full set of results of the stress test, including results from smaller banks conducting the test outside the public exercise.

Where that is greater than the conservation buffer and any supplementary macroprudential buffers in place, it will inform a decision on whether, and how far, to raise those macroprudential buffers, including the countercyclical capital buffer.

Once these system-wide components are set, the microprudential authority will, using its supervisory judgement, set bank-by-bank top-up buffers.

Such a top-up will be appropriate for banks whose balance sheets are more sensitive to the stress than others, such as those who use highly procyclical risk-weight models.[9]

All of this will be going on internally. To the outside, we will be one central bank, working in concert to set a total capital buffer for each bank that's big enough to deal with a storm if it hits.[10]

Our new approach is also consistent with the soon-to-be-implemented additional capital buffer that globally systemic banks must hold. Four of the seven banks currently in our stress test will be subject to an additional buffer, but each to a different size. [11]

The principle behind these buffers is to ensure that, the more systemic is a bank, the bigger the stress it can absorb. The tolerance of risk of failure is lower for systemic banks.[12]

We had two choices to align stress testing with this principle. The first was to subject each bank to a different severity of stress scenario – a projection of a different point in the tail of possible outcomes.

9 An issue highlighted in the 2014 stress test results, available at http://www.bankofengland.co.uk/financialstability/Documents/fpc/results161214.pdf.

10 If the impact of the stress is to be captured by the capital buffer, it follows that the capital banks must show they maintain in the stress test – the 'hurdle rate' - is their minimum capital requirement. Until now, the Bank has stipulated a hurdle rate of 4.5% common equity relative to risk-weighted assets. But each UK bank and building society has an add-on in its minimum requirement to reflect its own balance sheet and business model. So from next year, those add-ons will be reflected in the stress test hurdle rate explicitly. Because they differ across banks, the hurdle rate will be bespoke to each participant. Across the participants, they'll average 6.3% common equity relative to risk-weighted assets.

11 See FSB (2014).

12 Cunliffe (2014) sets out the challenges involved with the failure of systemically important institutions.

The second was to subject banks to a common scenario reflecting our risk tolerance for less systemic banks, but to hold globally systemic banks to a higher standard in the test.

For practical reasons, we chose the second. The more systemically important the bank, the stronger its capital position must be in the test. We are not allowing the systemic buffers to be used to deal with the stress in our test.

But this does *not* mean we think the additional capital buffers that systemic banks must hold are unusable in a real stress. Quite the opposite: like the other elements of the capital buffer, they are there to be depleted if necessary.

It simply reflects that we want to have confidence that these banks are able to withstand a stress more severe than the one we are applying to all banks.[13]

And by doing so, we have a stress testing approach that's completely consistent with the principles of the capital framework for globally systemic banks.

Catholic in modelling approach

The third C of our new approach is that it is **catholic** in its approach to modelling the impact of the stress.

We'll continue to ask banks to perform the test themselves, on their own balance sheet. We are not intending to apply our own models exclusively.

This allows banks' own stress testing capability to be put to the test. We want assurance that they can test themselves, for the specific risks they face, and do so regularly.

13 The supervisory response to a breach of Pillar 1 and Pillar 2A CET1 requirements in the stress will be more intensive relative to a failure to meet systemic buffers. For banks that fall below their minimum Pillar 1 and Pillar 2A CET1 capital requirement in the stress, there will be a strong presumption of an intensive supervisory response to rebuild capital. Systemically important banks that fall into their systemic buffers, but not their minimum CET1 capital requirements, will still be expected to strengthen their capital positions. But the supervisory response will be less intensive across one or more dimensions, including the size, nature and timing of required remedial actions (see Section 1.4 of The Bank of England's approach to stress testing the UK banking system).

It also ensures a wide range of models is brought to bear on each test. Each bank brings its own approach, so right now, that's seven modelling approaches in the test. We can compare them, pick the best for the problem at hand, and adjust the results of the others.

Our focus will be on developing models in areas that all banks find most challenging. The modelling of net interest income is one such area. But we are *not* seeking to check every part of banks' own efforts every year. Our resource is better spent on capturing what we can't expect the banks to capture: the system-wide dynamics of the stress.

These are amplifiers and feedbacks that were so evident in the financial crisis.

The mechanisms that turned so-called six-sigma events into day-to-day events.

The mechanisms that prompted Ben Bernanke to argue in 2008 that central banks needed to 'widen our field of vision'.[14]

I'd highlight three mechanisms that should be brought more fully into our field of vision.[15]

First, the feedback loop between weakened capital and higher funding costs, not just for individual banks, but across banks with similar business models.[16]

Second, the feedback mechanisms associated with banks' liquidity positions.

Like capital buffers, liquid asset buffers are there to be used in stress. And central bank facilities are open for business.

But in a stress, weaker banks may see the tenor of their funding shorten, there may be a flight to (perceived) quality by investors, and derivative counterparties can call for more collateral. All of this will place demands on liquid asset buffers.

Those buffers may not be sufficient to absorb the full impact. Where that's the case, we'll need to factor in the impact on capital strength of asset fire sales and, where appropriate, the impact on funding costs of use of central bank facilities.

14 See Bernanke (2008).

15 For a full review of these mechanisms and the approaches to deal with them, see Demekas (2015).

16 In this, we have made progress already, but more is needed to put this into practice (see Aikman 2009).

Third, our stress testing field of vision should begin to extend to non-bank parts of the financial system.

Investment funds are one such, but far from the only, part.

They have grown rapidly in importance since the crisis and in many cases are offering – or are perceived to offer – the ability to redeem shares at short notice.[17]

On the back of that, the Bank of England and the FCA have together been surveying these funds and assessing the risks associated with their growing importance.

Individual funds rightly manage their liquidity and test their ability to meet redemptions under stress. They tend to have contingency plans. They have little debt. And their investors bear the market risk. So there's not really a question of whether they can withstand a storm.

But a macroprudential stress test would ask a different question.

It would ask: Would investors and fund managers behave in a similar way after a correction in financial markets?

Will investors try to use the redemption possibility they're being offered *en masse*?

And if that means funds together try to sell into the same markets at the same time, how will that magnify volatility, disrupt the real economy, and feed back to the core of the financial system?

So their inclusion in any stress testing would be very different from the way we treat banks. But if the investment fund community is collectively engaged in activity that could create systemic risk – and that is a matter for further analysis at domestic and international level – then those activities should be within our field of vision.[18]

17 Globally, assets under management in open-ended investment funds have doubled since the financial crisis.

18 The Financial Stability Board, at its meeting in London last month encouraged funds to "assess their ability individually and collectively to meet redemptions under difficult market liquidity conditions" (see FSB 2015).

Conclusion

That's our new approach: countercyclical, consistent and catholic.

It's an approach that will see the UK's bank stress testing regime fit for the next stage of the cycle.

An approach that increases the predictability of our actions; that hardwires stress testing into the DNA of the banking system.

An approach that slots stress testing into the capital framework, not one that rivals it.

An approach that encourages banks to develop their own capability.

And above all, it's an approach that gives the economy what it deserves:

A banking system that's prudent, makes macroeconomic sense, and that supports it in both calm seas and severe storms.

Thank you.

References

Aikman, D. et al, (2009), "Funding liquidity risk in a quantitative model of systemic stability", Bank of England Working Paper No. 372, available at: http://www.bankofengland.co.uk/research/Documents/workingpapers/2009/wp372.pdf.

Bernanke, B. (2008), "Reducing Systemic Risk," remarks at the Federal Reserve Bank of Kansas City's Annual Economic Symposium, Jackson Hole, WY, available at: http://www.federalreserve.gov/newsevents/speech/bernanke20080822a.htm.

Caruana, J. (2012), "Building a resilient financial system", available at: http://www.bis.org/speeches/sp120208.pdf.

Cunliffe, J. (2014), "Ending Too Big to Fail – progress to date and remaining issues", available at http://www.bankofengland.co.uk/publications/Documents/speeches/2014/speech727.pdf.

Demekas, D. (2015), "Designing Effective Macroprudential Stress Tests: Progress So Far and the Way Forward", IMF Working Paper WP/15/146, available at: http://www.imf.org/external/pubs/cat/longres.aspx?sk=43043.

FSB (2014), "2014 update of list of global systemically important banks (G-SIBs)", available at: http://www.financialstabilityboard.org/wp-content/uploads/r_141106b.pdf.

FSB (2015), "Meeting of the Financial Stability Board in London on 25 September", available at: http://www.financialstabilityboard.org/wp-content/uploads/September-Plenary-press-release.pdf.

Woodford, M. (2003), *Interest and prices*, Princeton University Press, available at: http://press.princeton.edu/titles/7603.html.

5 Do stress tests pass the test?

Thomas F. Huertas[1]
EY

Stress tests have two purposes. From a microprudential perspective, they aim to assure that banks have enough capital now to withstand the losses they might incur in the future, if the macroeconomic environment were to become markedly more adverse. From a macroeconomic perspective, stress tests aim to assure that banks will have enough capital to sustain lending in the downturn, so that any recession will be short and shallow. In other words, stress tests aim to assure that both banks and the economy at large will be resilient.

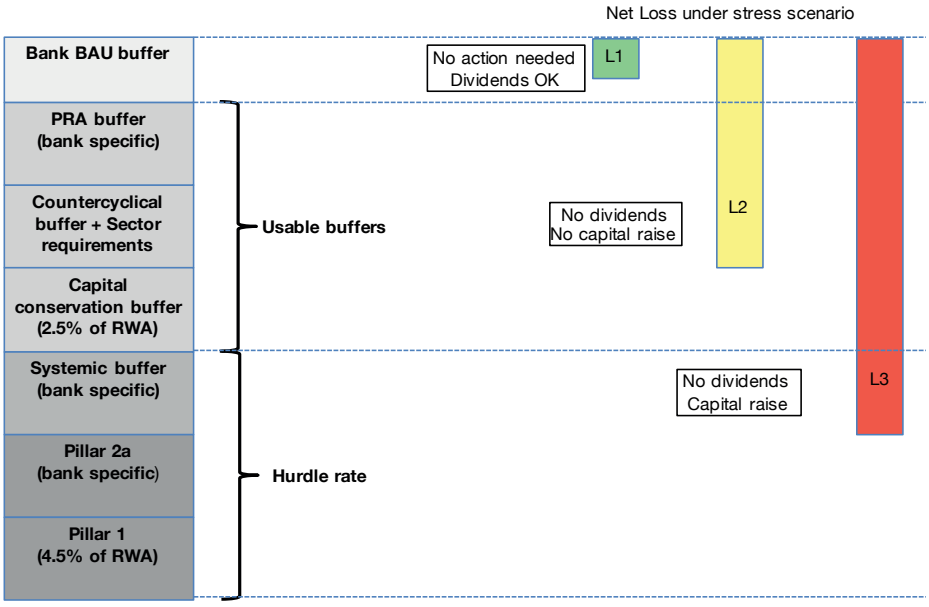
Can stress tests assure such an outcome? Do stress tests themselves pass the test? The answer to this question is important, for stress tests have become the cutting edge of microprudential bank supervision and control of credit, especially bank credit, and the main tool of macroprudential supervision.

Stress tests are the cutting edge of microprudential bank supervision

Failing one's stress test has serious consequences. The supervisor can forbid the bank from paying dividends or making distributions to shareholders, or order the bank to raise new capital (see Figure 1), even if the bank fully complies with current capital requirements on its current business.

1 The views expressed here are personal.

Figure 1 Stress testing is the the cutting edge of microprudential supervision



This possibility has certainly captured the attention of bank boards, and contributed to a simple form of shareholder discipline. Boards have generally made it very clear to their banks' executives that failing the supervisor's stress test will certainly slash the executives' compensation and may well shorten their careers. From a supervisory perspective, this is already a significant accomplishment. Stress tests have effectively elicited shareholder discipline in reinforcement of supervisory discipline. However, this may have come at the cost of focusing management's attention on passing the particular test set by the supervisor. And this may in turn literally create a concentration risk – namely, that concentrating on the supervisor's test may cause the bank to ignore questions not on the test, but that are nonetheless important for the bank.

The exam question

In concept, the exam question is simple: Does the bank currently have enough capital to withstand the losses that it might incur over the horizon covered by the test, if the environment became adverse? That in turn depends on the answer to two pairs of questions.

The first pair sets the framework for the test.

1. What 'bank' is taking the stress test?
2. What counts as capital?

The second pair sets the actual test.

3. How much capital is enough?
4. How should the stress scenario be set?

The second pair of questions is the main focus of a Bank of England paper (Bank of England 2015) and of Alex Brazier's speech (included as one of the chapters in this book; see Brazier 2016).

Before commenting on these, a few words on the first pair of questions, the framework for the test.

What 'bank' is taking the stress test?

An individual entity or the group as a whole? Increasingly, the answer is both. Banks are subject to multiple stress tests. Not only does the home jurisdiction supervisor test the group as a whole, but host country supervisors may each test material legal entities within the group and place restrictions on the ability of such subsidiaries to pay dividends or make distributions to its parent bank or parent holding company in the home jurisdiction.

Such a multi-test environment poses significant challenges, especially as each jurisdiction takes its own distinct approach to stress testing. Not only does the stress scenario differ from one jurisdiction to the next, but so do the test horizon, the data

requirements, the methodology and the date on which the stress test is due. This has implications for operations, capital planning and overall risk management (see below).

What should be the baseline or business-as-usual case to which the stress would be applied? The bank as it currently stands? This may accurately reflect the starting point for each bank and appear to facilitate comparisons across banks, but the static balance sheet used to implement such an approach is unrealistic. It not only fails to take into account where management plans to take the bank, but it fails to allow for what might be considered normal management response. Alternatively, should the baseline case be the bank as management currently plans to shape it over the test horizon? This has the advantage of conducting the test on the bank management expects to have over the test horizon, but will effectively assume that the supervisor has previously concluded (for example, via the bank's Individual Capital Adequacy Assessment Process, or ICAAP) that the bank's baseline forecast is realistic.

What counts as capital?

The simple answer is 'the amount of common equity tier 1 (CET1) capital as currently defined'. This, in turn, is the value of the bank's assets less its liabilities adjusted for items, such as goodwill, that count as assets but don't count towards CET1. Should the supervisor simply accept the bank's valuation, or should it commission an independent valuation of the bank's balance sheet? If the book value of the bank's assets is significantly above its economic value, there is a substantial 'garbage-in, garbage-out' risk: a bank could pass the stress test but fail in real life. This was the experience made in connection with the 2011 EBA stress test, and one of the reasons why the Single Supervisory Mechanism (SSM) initiated a comprehensive review of banks' balance sheets in 2014 prior to assuming supervision over banks in the Eurozone.

Over the test horizon, the definition of capital will change. Under Basel III and jurisdiction-specific implementation measures such as Capital Requirements Directive (CRD) IV, the definition of CET1 capital will harden over the horizon covered by the test, as the ability to count items such as deferred tax assets toward CET1 capital gets phased out. Most stress tests take this into account.

But what about changes in accounting standards, such as the shift to expected loss provisioning under International Financial Reporting Standard (IFRS) 9? Under stress, it can be anticipated that a much higher proportion of loans will fall into category 2, where provisions should be made for possible losses over the remaining life of the loan. This can result in much higher provisions than under the current impairment method. Is it reasonable to require banks to use the new methodology for the current (2016) stress tests for loans that are projected to be on their balance sheet in 2019 (when the new methodology takes effect)? That would be tantamount to accelerating the implementation date for the new accounting standard – something very few banks, if any, would be able to achieve.

And, what about fines, settlements and restitution charges that banks may incur over the test horizon? These are now large enough to have a material impact on CET1 capital. Indeed, in some cases such charges outweigh the provisions that banks have had to make for credit losses. As a result, for purposes of the stress test, banks may need to make allowances not only for current enforcement cases, but also for breaches incurred in the past but not yet reported as well as for any new breaches that the bank may commit over the test horizon.

We now turn to the second pair of questions, and to the Bank of England's answers.

How much capital is enough?

This is effectively the amount needed to pass the stress test. De facto, this so-called 'hurdle rate' has become the bank's minimum capital requirement.

In the Bank's view, there is no single number across all banks, and the number for each bank is likely to be considerably higher than the 7% of risk-weighted assets commonly cited as the minimum for banks under Basel III, or even the 9.5% minimum cited for globally systemically important banks. In order to be able to pay dividends or make distributions, the bank must meet all requirements (Pillar 1 and Pillar 2a) and fill all buffers (a systemic surcharge, a capital conservation buffer, a countercyclical buffer and a Prudential Regulation Authority buffer) – a figure that can easily exceed 15% of risk-weighted assets. If the bank falls below that boundary, it cannot pay dividends or make distributions, and if the bank falls below the sum of its requirements (Pillar 1 and

Pillar 2a) and its systemic surcharge, it must raise new capital or face the prospect of resolution (see Figure 1).

This 'hurdle-rate' minimum is not only significantly higher than the levels set under Basel III, but it has been introduced considerably earlier. This has effectively negated the transition period negotiated as part of Basel III and forced banks to improve the quantity and quality of capital much faster than originally anticipated. This in turn may have contributed to de-leveraging and constrained the growth in bank lending, particularly in Europe.

How should the stress scenario be set?

As Goldilocks might say, the scenario should be 'not too tough, and not too soft', and this is what the Bank of England will aim to achieve with the changes proposed in its consultation paper.

The first and most significant change is to make the degree of stress countercyclical. Expressed in terms of percentage decline in GDP, the stress scenario imposed at the top of the cycle will call for a larger fall in GDP than the scenario imposed at the trough of the cycle. In other words, there will be a limit to what might be called stress-on-stress. This diminishes the prospect that the stress test could turn into a self-fulfilling prophecy, but it ultimately depends on a view that the authorities have the ability to arrest a downturn.

That seems a bold statement to make in the current environment, for the Great Recession has reduced monetary and fiscal flexibility. If another recession were to develop over the coming year, central banks would have little or no scope to cut rates, and governments would generally have little or no scope to increase budget deficits. Therefore, the Bank's countercyclical approach may well depend on whether the current recovery lasts long enough and is rapid enough to restore monetary and fiscal flexibility.

The second aspect of the Bank's paper is the confirmation that it will continue to test the resilience of the banking system to specific stresses. Such a specific stress should be worthy of a 'deep dive': relevant for the banking system as a whole, if it were

to materialise; and not too likely to occur (otherwise it would merit inclusion in the general scenario), but not too remote either.

In sum, stress testing has certainly made banks more resilient. It has forced banks to increase the quantity and quality of capital far further and far faster than originally mandated under Basel III. It has also contributed to improving banks' risk governance and to restoring (some would say, introducing) shareholder discipline. All in all, from the perspective of microprudential supervision, stress testing should receive a grade of 'pass'.

The role of stress tests in macroprudential supervision

In contrast, 'incomplete' is the grade that stress tests should be given for macroprudential supervision. Macroprudential supervision rests on the premise that controlling credit can smooth the economic cycle. According to the theory, restraining credit growth as the upswing gathers momentum should help prevent the boom from getting out of hand and ending in a bust. Should a recession nonetheless develop, promoting credit growth in the downswing should help shorten the length and shrink the magnitude of the downturn. Stress tests should therefore serve as an input to macroprudential supervision – above all, they should assure that banks have the capacity to lend in a downturn.

However, there is not a complete match between micro and macro supervision. The former concerns the banks headquartered in the supervisor's jurisdiction; the latter concerns the state of that jurisdiction's economy. The two do not coincide. Banks headquartered in a jurisdiction may be active in other jurisdictions as well. Stress tests that probe the resilience of a bank as a whole need not indicate that bank's behaviour with respect to the terms or volume of credit that the bank will provide in the jurisdiction administering the test. In addition, banks are not the only source of credit – bonds are a prominent, and arguably more procyclical, alternative to loans. So controlling bank credit need not control the volume of credit in the economy as a whole.

All stress tests can do is assure that banks will have the capacity to lend in the downturn. But they cannot necessarily assure that banks will have the capacity to make up for the

decline in credit from other sources, such as the bond market. Nor can stress tests assure that creditworthy borrowers will demand credit. Placing an obligation on banks to 'support' the real economy could therefore compromise microprudential supervision. Such 'support' either implies that banks should relax their credit standards at the trough of the cycle, or base their credit decisions in the downturn on the assumption that the economy will recover.

Finally, stress tests focus only on the lender. They fail to take the borrower into account. If controlling leverage holds the key to controlling the quality of credit (and therefore to assuring financial stability), then surely macroprudential supervision should take leverage at the borrower level into account, not just leverage at banks. This would imply that macroprudential supervision concern itself as much or more with setting minimum loan-to-value and loan-to-income ratios as it does with setting bank capital ratios.

In sum, stress testing is far more likely to make banks resilient than the economy at large.

What banks need to do

Stress tests are here to stay. Stress tests are multiplying across jurisdictions. Stress tests are becoming more comprehensive. Supervisors are starting to look at liquidity as well as capital and to grade the test not just on the number generated by the test, but also on how the bank generated that result. To pass the stress test, a bank's capital ratio must exceed the hurdle rate, and the bank's data, data management and modelling must pass muster as well. Only then will the bank be able to pay a dividend to its shareholders.

Global banks face an additional challenge. They must integrate the stress tests from their principal jurisdictions. Indeed, the parent in jurisdiction A may not be able to pass its stress test, unless its significant subsidiary in jurisdiction B can freely pay dividends and make distributions to the parent – something the subsidiary can only do if it passes jurisdiction B's stress test.

Banks will therefore need to 'industrialise' the stress test process. This starts with streamlining the collection of data about the bank's portfolio and about the factors that

determine the risk of that portfolio. It continues with improvements in the ability of the bank to model the effects of changes in the underlying factors on the value of the portfolio, particularly in periods of market stress, when correlations among risk factors increase, spreads widen and liquidity contracts. It further proceeds by integrating finance and risk, so that the results of the risk modelling feed into the bank's finance models and capital plans. This capability will become increasingly important as group supervisors, rating agencies and investors in parent company TLAC-qualifying debt securities begin to demand details on the dividends and distributions that the parent holding company or parent bank can expect to receive from material subsidiaries in other jurisdictions.

References

Bank of England (2015), "The Bank of England's approach to stress testing the UK banking system", approach document, London (http://www.bankofengland.co.uk/financialstability/Documents/stresstesting/2015/approach.pdf).

Brazier, A. (2016), "The Bank of England's approach to stress testing the UK banking system", in R. Anderson (ed.), *Stress Testing and Macroprudential Regulation: A Transatlantic Assessment*, London: CEPR Press.

6 Stress testing over the financial cycle: General approach, challenges and complements

Philipp Hartmann[1]
European Central Bank

"The Bank of England's approach to stress testing the UK banking systems" (Bank of England 2015, Brazier 2016) contains the most lucid description I have seen to date about how to use stress tests in a dynamic prudential capital assessment framework. I would expect that it will influence many supervisory authorities' thinking about this issue around the world.

Stress testing over the financial cycle

Let me start looking at it from a macroprudential perspective. There are two variants in which the macroprudential policy objective has been phrased. One I would like to denote as 'symmetric' across the financial cycle, and the other as 'asymmetric'. Proponents of the asymmetric macroprudential objective suggest this novel policy branch should aim at preventing systemic financial crises, or at least keeping their likelihood at very low levels. This variant is truly prudential in the sense that it takes primarily an ex ante perspective, focusing in the first place on the upturn of the financial cycle, when the preventive measures should take place.

Proponents of the symmetric macroprudential objective are, I would argue, more ambitious. They suggest that macroprudential policy should smoothen the financial cycle in its entirety. By leaning against exuberance during the upturn of the cycle, a

1 Any views expressed are only the author's own and should not be regarded as views of the ECB or the Eurosystem.

successful implementation of this variant would prevent that the preconditions for a systemic crisis can emerge. Moreover, by relaxing regulatory constraints during the downturn of the financial cycle (ex post) it would try to help financial intermediation to come out of the trough and continue to support the economy with credit. Supporting the smooth flow of intermediation services in this way resembles other macroeconomic stabilisation policies, such as monetary or fiscal policy.[2]

Bank stress tests can contribute to any of the two formulations of macroprudential policy. For the asymmetric variant the stress tests could be designed in a way that banks are basically unlikely to fail at any point in time. If the stress tests are associated with the formulation of capital needs so that default probabilities are extremely low, the likelihood of a systemic banking crisis should be even lower.

In this chapter, I would like to argue that the Bank of England's approach to stress testing makes a good step in the direction of the more ambitious symmetric form of prudential policy.[3] Let me explain how I get to this conclusion through a simple graphical illustration. Figure 1 shows the financial cycle over time in a stylised form as a dashed curve. Time is indicated as t on the horizontal axis. The vertical axis measures the amount of financial intermediation services, such as credit provision for example. The annual concurrent stress tests, as described in the Bank of England's paper, will be based on scenarios that are conditional on the state of the financial cycle (see, for example, Box 2).[4] These cyclical scenarios will be 'harsher', i.e. assuming more drastic stresses, in the upturn of the cycle.

This is based on the observation that the upturn of the cycle is characterised by accelerating credit and/or asset price growth and compressing risk premiums. Although the measured instability is still low, underlying risks grow as financial imbalances may build up. As a consequence, the probability or size of potential downward corrections (e.g. unravelling of imbalances) becomes greater. State-contingent scenarios should

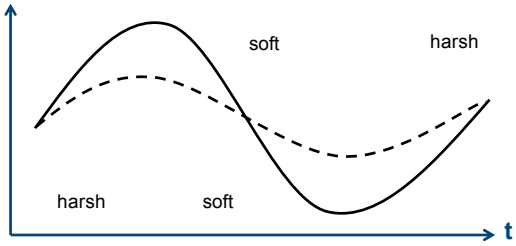
2 For a related but different discussion of the macro-prudential objective, see Bank of England (2009, section 2).

3 I focus only on a few aspects of the paper here. There are many interesting topics discussed in it that I cannot address in the present context.

4 The Bank's stress testing framework comprises these annual concurrent scenarios and biennial exploratory scenarios. Although also interesting, I cannot discuss the latter here.

reflect these greater risks, making the stress test harsher for the banks during the financial upturn. In principle, this provides incentives to banks for holding increasing capital buffers during the upturn, which reach their peaks at the turning point where the financial cycle declines again. A stylised representation of the desirable fluctuations of capital buffers is indicated through the solid line in Figure 1.

Figure 1 Stylised representation of the financial cycle, the severity of stress test scenarios and associated capital policies



As the financial cycle declines, losses may materialise and part of the capital buffers may be eroded. If the stress test scenarios stayed 'harsh', banks may have to replenish capital precisely when asset prices decline, credit provision decelerates and risk premiums are high. In other words, equity issuance will be costly and deleveraging on the asset side will come on top of the already decelerating credit growth. Such a policy could easily become procyclical, further reinforcing the downturn in the financial cycle.[5]

Avoiding this tendency, the Bank of England approach suggests softening stress test scenarios when the financial cycle passes its peak. This way, the associated prudential capital assessment acts in a countercyclical way (Bank of England 2015, Brazier 2016).[6] As banks can run down capital buffers (see the decline in the solid line of Figure 1) above a regulatory minimum and softer scenarios would usually not lead to the obligation to maintain high capital ratios via the disposal of assets or to conduct drastic recapitalisations, the risk that the stress test-related prudential capital policies

5 Already Crockett (2000) had described how a "fallacy of composition" could be associated with a purely microprudential approach. Trying to stabilise individual banks through capital increases in a general downturn, when risks materialise, would lead to procyclicality and could therefore destabilise the financial system.

6 See Kowalik (2011) for a general discussion of counter-cyclical capital policies (without particular reference to stress tests).

further reinforce the financial cycle is minimised. Quite the contrary, banks should have room for maintaining the provision of financial intermediation services if the build-up of capital in the upturn was large enough.

So far, I described the approach in a way that capital buffers would be high in the upturn, reducing the likelihood of failures and crises, and lower in the downturn, avoiding that supervisory constraints amplify the negative part of the financial cycle. In theory, such a policy could also be pushed to the point where it is attempted to smoothen the financial cycle. To illustrate this, let us now interpret the curves in Figure 1 differently. The solid black line now describes the financial cycle, before policy is applied, and 'harsh' and 'soft' continue to refer to the countercyclical stress test scenarios. If the changes in the severity of the scenarios and subsequent capital policies were pronounced enough, they would have the potential to push the upturn of the financial cycle down and the downturn up. As a consequence, the financial cycle would flatten from the amplitude of the solid line to that of the dashed line.

Such a more radical approach would amount to an implementation of the more ambitious symmetric macroprudential approach, the first variant I described earlier. The Bank of England explained that smoothening the entire financial cycle was not the intention of the approach described in their paper. It is intended to ensure that regulatory constraints do not act in a procyclical manner during the downturn, but not to flatten the financial cycle in the upturn.[7] Still, I found that the paper beautifully illustrated how – in principle – such an aggregate (symmetric) financial stabilisation policy could work.

Rules versus discretion

The Bank of England's approach to stress testing also relates to another conceptual policy debate, one that was hotly debated in monetary policymaking for a long time: the choice between rules and discretion.[8] In principle, it could be seen as a rules-based

7 Response by Alex Brazier, Executive Director Financial Stability Strategy and Risk, at the London School of Economics/ CEPR conference on "Stress Testing and Macro-prudential Regulation: A Trans-Atlantic Assessment" on 30 October 2015.

8 See, for example, Fischer (1990) for a survey.

approach. As the severity of scenarios is conditioned on the state of the financial cycle, prudential capital policies could follow from it in a quasi-automatic way. To the extent that the relevant financial cycle(s) is (are) captured well, some kind of formula could be defined that would clarify the 'policy reaction function' to bank managements and investors in bank liabilities. If the policy rule could be announced in a credible way, all parties involved would adapt their expectations to it for the different parts of the cycle. This would remove uncertainty for banks, investors and potentially even for other agents in the economy whose activities depend on the availability of financial intermediation services (such as credit) enhancing the effectiveness of prudential capital policies. It would also remove potential incentives for authorities to pursue actions that may avoid political costs in the short term, but that in terms of the overall net benefits may be inferior to a policy rule that has been optimised over the entire financial cycle.

Following a strict rule, however, faces a number of challenges. First, it is difficult to ascertain the state of the financial cycle in real time. Second, different aspects of financial intermediation might have different cyclical frequencies or amplitudes. Third, specific circumstances that do not follow the nice and well-behaved theoretical exposition above may require exceptions to or amendments of the rule. This is why I would describe the approach lined out in the Bank of England paper as falling in between the two extremes of strict rules and full discretion in policymaking.[9] Practically speaking, the approach could perhaps be denoted as one of 'constrained discretion'.[10] The general principles that the Bank of England plans to follow constitute the rule-like features, while it preserves enough discretion for responding to specific circumstances of a given financial cycle or bank.

Assessing the financial cycle

Let me discuss two of the challenges that can emerge in such a framework. The first is to ascertain the state of the financial cycle and be consistent in doing so over

9 Kowalik (2011) argues that a rule-based approach has more advantages than a discretionary approach.

10 Alex Brazier confirmed at the LSE/CEPR conference on 30 October 2015 that some form of constrained discretion is broadly what the Bank has in mind.

time. To start, there is not a single financial cycle that strictly dominates all financial markets and intermediaries at the same frequency. For example, securities markets, such as equity markets, tend to have more pronounced shorter cycles more closely aligned with the regular business cycle, whereas credit and housing markets tend to have more pronounced longer cycles (e.g. Drehmann et al. 2012, Table 1). The practical implication is that stress test scenarios may have to include different counterfactuals for different asset prices and financing quantities.[11] And, of course, there is a limit to how many different sub-scenarios can be practically pursued.

Second, not all financial cycles have the same implications for systemic financial instability. Historical research suggests that booms driven by bank credit and/or housing valuations more frequently lead to severe financial crises (Crowe et al. 2013, Schularick and Taylor 2012), whereas equity financed bubbles tend to have more limited implications in the downturn. This means practically that more emphasis may have to be given to scenarios for the systemically more important financial markets or intermediaries.

Third, financial systems are subject to structural change. New financial instruments and intermediaries emerge and old ones may lose importance. For example, as a consequence of the financial crisis and subsequent re-regulation, non-bank intermediaries ('shadow banks') may gain market shares and traditional banks somewhat lose. As the transformation of financial systems progresses, so will the way in which prudential authorities need to assess the financial cycle. All these factors make it difficult to base a state-contingent stress testing framework on a simple rule that could be applied relatively mechanically over an extended period of time. Quite some judgement has probably to be applied and be incorporated in the framework through a fair amount of discretion.

11 See, for example, the references to ranges of indicators, markets and sectors on page 12 of the Bank of England paper (Bank of England 2015). It should, however, be noted that even for single markets different indicators of the cycle and the existence of imbalances can easily give opposite results. See, for example, the case of residential real estate markets in Eurozone countries (Hartmann 2015).

Example of an exception: Recapitalisation in a downturn

The other challenge I would like to discuss is a particularly important exception to the 'rule' of letting bank capital erode in the downturn. I base my argument on a simulation of a model that we developed as part of the European System of Central Banks Macroprudential Research Network (MaRs) (ESCB Heads of Research 2014). This network had as its objective to improve the analytical underpinnings of macroprudential policies. One of its core outputs was the development of a macroeconomic model that could assess the overall benefits and costs of prudential capital policies (Clerc et al. 2015).[12] Featuring heterogeneous bank, firm and household sectors, which all exhibit default risks, the model was dubbed the '3D model' (for its three layers of default). The ECB is currently implementing this model as a quantitative assessment tool for supporting macroprudential policymaking.

The way I would like to make my argument can be seen from two perspectives. One is that if stress test scenarios are not harsh enough in the financial upturn, then banks may be left with too little capital in the downturn and letting capital run down further could become counterproductive. The other perspective is that if the downturn materialises more drastically than could be anticipated, then letting capital be eroded too much may become counterproductive. Both perspectives rely on the assumption that stress test scenario-building and associated capital policies may never be perfect. Or, alternatively, a cost-benefit analysis will never lead capital levels to cater for all contingencies and reduce the probability of a crisis to zero.

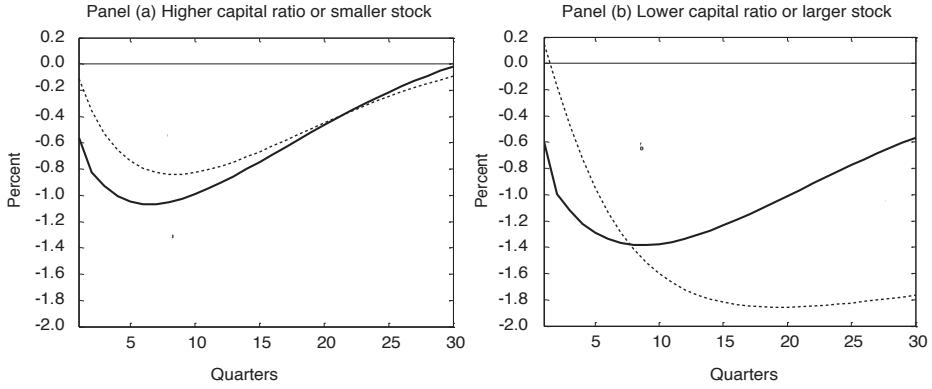
With these considerations in mind, look at the two panels of Figure 2, which both show time on the horizontal axis (measured in quarters of years) and the percentages by which the total output of the economy deviates from its steady-state equilibrium level.[13] The curves in both panels are impulse response functions describing the effects of persistent declines in house prices and firm valuations in the 3D model. The two

12 All of MaRs' results are summarised in its final report (ESCB Heads of Research 2014).

13 The two panels are excerpts of Figures 13 and 14 in Clerc et al. (2015), where they are used to discuss the role of the Basel countercyclical capital buffer. In this first paper, the simulations are run for a rough calibration of the 3D model to the Eurozone.

solid lines show the impulse responses under the assumption that no regulatory capital can be released, i.e. there is a higher threshold below which bank capital ratios cannot decline. The two dashed lines describe the impulse responses under the assumption that the losses caused by the financial shock on households and firms can eat deeper into bank capital.

Figure 2 Simulations of the dynamic effects of a persistent reduction in house prices and firm valuations on economic output in the 3D model



Notes: The four curves in the two panels show impulse response functions of economic output in deviations from the steady-state equilibrium level. In panel (a), the pre-shock bank capital ratio is assumed to be 10.5% of risk-weighted assets, which amounts to the optimal steady-state level in the calibration used. In panel (b), it is 8%. Solid lines describe impulse responses when capital ratios are bounded from below. Dashed lines describe impulse responses when the capital ratio is allowed to decline.

Source: Clerc et al. (2015), Figures 13 and 14.

The left-hand panel shows the two impulse responses for the case when bank capital ratios are relatively high before the shock. The right-hand panel shows them for a lower bank capital ratio. Let me, however, interpret the figure from the second perspective mentioned above, under which – equivalently – the left-hand side could describe the situation for a smaller and the right-hand side for a larger financial shock (with capital being the same across the two panels). One sees on the left that the economic downturn caused by the financial shock is less pronounced when capital is released (at least for about a year and a half). The situation is, however, different on the right. Whereas the dashed line starts above the solid line, after less than a year it moves increasingly below it. In other words, if too much capital is eroded then the fragility that this implies for the banking system exceeds the benefits of trying to maintain credit provision via

a softening touch in regulatory capital requirements. The recession becomes deeper instead of shallower.

The lesson for countercyclical stress testing is that there can be extreme circumstances in which a stress test scenario should not be softened in the financial downturn. For example, when a large negative shock has significantly weakened the capital position of (some) banks, it may sometimes be necessary to enforce decisive recapitalisations. In such circumstances the positive, confidence-building effects in the medium term of such a deviation from the regular policy rule may well exceed the potential short-term obstacles to the provision of financial intermediation services.

Whilst this consideration may enter some time-inconsistency element into the constrained discretionary policy approach, which may be hard to judge for banks and investors ex ante, it should not be forgotten that the post-crisis bank capital framework in the UK and in many other countries contains considerable safety buffers against such a scenario becoming relevant. For example, under Basel III banks have to maintain considerable minimum levels of high-quality capital and the variable parts that can be built up and run down after the peak of the financial cycle are strictly on top of those. So, the likelihood of scenarios as the one just described should be relatively low.

Caveats and broader issues

In this chapter, I have taken a quite narrow perspective on the role of bank stress tests and associated prudential capital policies. I particularly focused on the interesting feature of the Bank of England's approach to stress testing the UK banking system that make some tests state contingent, i.e. dependent on the state of the financial cycle. In my view, this is a strong contribution to the ongoing work of defining macroprudential policy frameworks. There are, however, a number of broader issues that should not be forgotten. Before leaving the topic, I would like to raise two of them. First, stress tests – as they are usually pursued by bank supervisory authorities – are only one tool among several, so their value and use should be put into perspective. Second, changing bank capital is only one among a number of policies that can be used to manage individual or systemic risks.

On the need to complement stress tests

It is perhaps not an exaggeration that stress tests have become by now something like the holy grail of prudential policy. For those who were not convinced before, their successful use by the IMF in assessing the financial sectors of their member countries and by US authorities in re-establishing confidence in their country's main banks in response to the financial crisis that broke out in 2008 illustrated how powerful this tool can be. In my view, one of the great strengths of the stress testing approach is the room it gives to initiative by the competent policy authorities. Public authorities can act decisively by proactively designing the scenarios, setting the parameters of the exercises or running their own models. They can have a substantial impact on the behaviour of the tested banks over a relatively short time horizon.

But there is also a risk that the stress testing approach becomes the victim of its own success in the future. It now consumes a significant share of the human resources in prudential policy authorities, of the risk management staff in banks and of the attention span of investors. But stress tests are not (yet) an all-encompassing tool. One way to see this is by looking to which extent they can capture the different forms of systemic risk, aggregate shocks, the unravelling of widespread imbalances and contagion (ECB 2009, De Bandt et al. 2014). Once stress tests moved beyond their micro application for assessing investment portfolio risks, they were particularly focused on assessing banks' resilience to aggregate shocks. Recently, some efforts were made by some authorities, including the ECB (Henry and Kok 2013), to incorporate contagion modules into the macro approach. And the Bank of England's approach to linking the scenarios to the financial cycle illustrates one way of capturing the unravelling of imbalances to some extent. But the current frameworks used are by no means complete for macroprudential purposes. So, it is important that regular stress tests are also accompanied by early warning tools for detecting widespread imbalances and by dedicated simulation tools for assessing contagion risks (ECB 2010).

Another area to watch carefully is the models that are used in stress testing. Decades of finance-less macro theory before the crisis left modellers without coherent analytical toolkits for combining aggregate economic developments and bank stability. The MaRs network mentioned before made a major effort to provide a series of novel macro

financial stability models.[14] I hope that over time this fundamental research effort, and the ones by a few others pursuing similar directions, will also find its way into practical stress test tools and thereby enhance their internal consistency. Against this background, I fully commend the Bank of England's mentioning of priorities for model development in sub-section 1.2 of their paper. Last, but by no means least, improved stress test frameworks make significant demands on data, as also emphasised by the Bank of England in sub-section 2.4. In my view, to combine the economy-wide perspective with banks and their interrelations requires particularly granular data (such as detailed data from microprudential supervision) that can be aggregated up in consistent ways (for macroprudential purposes).

In sum, the success of stress testing creates the risk that its shortcomings are overlooked and complementary approaches and tools neglected. It would be imprudent not to invest in improving current practices and not to complement stress tests with other tools.

Other prudential policies than capital requirements

The second caveat I would like to raise is the choice of policy instruments. Policies can address financial stability risks on the side of the providers of financial services, or on the demand side of financial services. To take the particularly important example of credit provision, in the upturn of the financial cycle either borrowers' risks could be managed by limiting their access to credit, or lenders' risks could be addressed by constraining their credit provision. Bank stress tests with associated capital policies belong to the latter category. Banks are supposed to withstand the failure of some borrowers without reducing their credit provision to other borrowers.

Borrower-based policy instruments include notably loan-to-value ratios (LTVs) and debt-to-income limits (DTIs). LTVs limit the size of a loan, for example, that a consumer can take from a bank relative to the value of the house that she intends to purchase with it. DTIs limit the size of a loan that a consumer can take from a bank relative to his regular income. Recent empirical research suggests that these borrower-based

14 For overviews, see, for example, Hartmann et al. (2013), Boissay et al. (2015) and ESCB Heads of Research (2014).

instruments show some effectiveness in limiting fluctuations in credit growth, whereas lender-based instruments are relatively ineffective in this regard (e.g. Claessens et al. 2014).

Whereas it is not the ambition of the new UK stress testing framework to generally smoothen credit cycles, it is still important to recall that there are other policy instruments than stress tests and capital requirements. Depending on the objective and the sources of financial imbalances, it may be advisable to also consider those.

References

Bank of England (2009), "The role of macro-prudential policy – A discussion paper", London, November (http://www.bankofengland.co.uk/publications/Documents/other/financialstability/roleofmacroprudentialpolicy091121.pdf).

Bank of England (2015), "The Bank of England's approach to stress testing the UK banking system", London, October (http://www.bankofengland.co.uk/financialstability/Documents/stresstesting/2015/approach.pdf).

Boissay, F., P. Hartmann and K. Nikolov (2015), "Financial instability in macroeconomics: a set of new structural models", ECB Research Bulletin No. 22, pp. 1-11 (https://www.ecb.europa.eu/pub/pdf/other/researchbulletin22.en.pdf).

Brazier, A. (2016), "The Bank of England's approach to stress testing the UK banking system", in R. W. Anderson (ed), *Stress Testing and Macroprudential Regulation: A Transatlantic Assessment*, London: CEPR Press.

Claessens, S., S. Gosh and R. Mihet (2014), "Macro-prudential policies to mitigate financial system vulnerabilities", IMF Working Paper No. WP/14/155, Washington, DC.

Clerc, L., A. Derviz, C. Mendicino, K. Nikolov, S. Moyen, L. Stracca, J. Suarez and A. Vardoulakis (2015), "Capital regulation in a macroeconomic model with three layers of default", *International Journal of Central Banking* 11(3), pp. 9-63.

Crockett, A. (2000), "Marrying the micro- and macro-prudential dimensions of financial stability", remarks before the Eleventh International Conference of Banking Supervisors, Basel, 20-21 September (http://www.bis.org/speeches/sp000921.htm).

Crowe, C., G. Dell'Ariccia, D. Igan and P. Rabal (2013), "How to deal with real estate booms: lessons from country experiences", *Journal of Financial Stability* 9(3), pp. 300-319.

De Bandt, O., P. Hartmann and J. L. Peydro (2014), "Systemic risk in banking after the Great Financial Crisis", in A. Berger, P. Mullineux and J. Wilson (eds), *Oxford Handbook of Banking* (2nd edition), Oxford: Oxford University Press, pp. 667-699.

Drehmann, M., C. Borio and K. Tsatsaronis (2012), "Characterising the financial cycle: don't lose sight of the medium term!", BIS Working Paper No. 380, Basel (http://www.bis.org/publ/work380.pdf).

ESCB Heads of Research (2014), *Report on the macro-prudential research network (MaRs)*, Frankfurt am Main: ECB (https://www.ecb.europa.eu/pub/pdf/other/macroprudentialresearchnetworkreport201406en.pdf).

European Central Bank (2009), "The concept of systemic risk", *ECB* Financial Stability Review, December, pp. 134-142 (https://www.ecb.europa.eu/pub/pdf/other/financialstabilityreview200912en.pdf).

European Central Bank (2010), "Analytical models and tools for the identification and assessment of systemic risks", ECB Financial Stability Review, June, pp. 138-146 (https://www.ecb.europa.eu/pub/pdf/other/financialstabilityreview201006en.pdf).

Fischer, S. (1990), "Rules versus discretion in monetary policy", in M. Friedman and F. Hahn (eds), *Handbook of Monetary Economics* (Vol. 2), Amsterdam: North Holland, pp. 1155-1184.

Hartmann, P. (2015), "Real estate markets and macro-prudential policy in Europe", *Journal of Money, Credit and Banking* 47(1), pp. 69-80.

Hartmann, P., K. Hubrich and M. Kremer (2013), "Introducing systemic financial instability into macroeconomics: how to meet the challenge?", *ECB Research Bulletin* No. 13, pp. 2-9, Frankfurt am Main (https://www.ecb.europa.eu/pub/pdf/other/researchbulletin19en.pdf).

Henry, J. and C. Kok (eds) (2013), "A macro stress testing framework for assessing systemic risks in the banking sector", ECB Occasional Paper No. 152, October, Frankfurt am Main (https://www.ecb.europa.eu/pub/pdf/scpops/ecbocp152.pdf).

Kowalik, M. (2011), Countercyclical capital regulation: should bank regulators use rules or discretion?", *Federal Reserve Bank of Kansas City Economic Review*, Second Quarter, pp. 59-80 (https://www.kansascityfed.org/publicat/econrev/pdf/11q2Kowalik.pdf).

Schularick, M. and A. Taylor (2012), "Credit booms gone bust: monetary policy, leverage cycles, and financial crises, 1870-2008", *American Economic Review* 102(2), pp. 1029-1061.

7 Making sense of the comprehensive assessment

Viral V. Acharya and Sascha Steffen
New York University and CEPR; University of Mannheim and ZEW

Motivation

In an earlier paper (Achary and Steffen 2014), we estimated capital shortfalls of European banks that are going to be part of the Single Supervisory Mechanism (SSM) using 'benchmark' stress tests. We documented that the comprehensive assessment might reveal a substantial lack of capital in many peripheral and core European banks. The ECB finalised its assessment of the largest banks in the Eurozone before it commenced their regulatory oversight in November 2014. It has now disclosed its own assessment of the solvency of the banking sector.

How do our benchmark capital shortfalls compare to the regulatory shortfall estimates?

Sample

The ECB included 130 banks in the comprehensive assessment. Of these, it will eventually supervise 120 banks directly. This set of banks includes 39 publicly listed financial institutions for which supervisory data, as well as our benchmark stress test data, are available.[1] We use balance sheet data from SNL Financial as of 31 December 2013, which is also the starting point of the comprehensive assessment.

1 The Bank of Cyprus was privatised in 2014. Banco Espirito Santo, a Portuguese lender, failed in August 2014; while the 'bad bank' will be wound down, the viable part part of the bank has been transferred into a new entity which was not included in the stress test exercise due to time constraints.

Table 1 Descriptive statistics

Country	C Tier 1	Equity/assets	Market-to-book	Market cap	Assets	Banks
France	10.86%	4.24%	0.68	127,696	4,543,804	3
Germany	12.95%	3.83%	0.61	50,570	2,204,035	3
Italy	10.61%	6.49%	0.61	83,000	2,190,872	11
Spain	11.65%	7.22%	1.00	146,082	2,080,440	5
Belgium	16.31%	4.00%	1.18	17,305	461,622	2
Greece	12.45%	8.27%	0.95	26,945	354,223	4
Ireland	13.30%	7.22%	4.91	68,303	287,468	3
Austria	11.64%	7.24%	0.72	11,453	221,022	2
Portugal	15.00%	4.48%	0.91	4,978	124,707	2
Malta	10.67%	7.70%	1.58	1,557	12,979	2
Slovakia	15.93%	11.94%	0.70	964	11,556	1
Cyprus	7.34%	6.25%	0.57	229	6,384	1
Total	11.68%	5.30%	0.84	539,083	12,499,112	39

Notes: This table reports descriptive statistics of the publicly listed banks included in the comprehensive assessment conducted by the ECB in 2014. C Tier 1 is the Core Tier 1 ratio, and is Core Tier 1 capital divided by risk-weighted assets (RWAs). Equity/assets is book equity over total assets. Market-to-Book is market value over book value of equity. Market cap is the market value of equity measured in million euros. Assets are total assets and measured in millions of euros. Banks are the number of public banks in each country. All data are as of 31 December 2013 and aggregated at the country level. Ratios are weighted using total assets.

Table 1 shows that these banks have €12.5 trillion in total assets and a market capitalisation of €539 billion. Table 1 also provides an overview of the mean regulatory capital ratio core equity Tier 1 (C Tier 1) as well as equity/asset and market-to-book ratios. The mean C Tier 1 capital ratio is 11.68%, the mean equity/asset ratio is 5.3%, and the market-to-book ratio is 0.84 and well below 1. Markets are substantially discounting banks' assets, and Cyprus, Italy and Germany lead the table with the banks that show the lowest market-to-book ratios.

Methodology

1. **Benchmark stress test results ('SRISK' or 'Capital Shortfall in a Systemic Crisis')**. We assume a systemic financial crisis with a global stock market decline of 40%. SRISK 5.5% VLAB is our measure for a bank's capital shortfall in this scenario, assuming a 5.5% prudential capital ratio with losses estimated using the VLAB methodology to estimate the downside risk of bank stock returns.2 While this scenario and the resulting SRISK measure use market data and market equity (instead of book equity) in determining leverage, the approach is conceptually similar to that of the EU stress tests, which is to estimate losses in a stress scenario and determine the capital shortfall between a prudential capital requirement and the remaining equity after losses.
2. **Supervisory stress test results**. We use following outcomes from the comprehensive assessment to relate to our benchmark stress tests:
   a. **Capital shortfall:** Capital shortfall of banks to a threshold of 5.5% Common Equity Tier 1 (CET1) in the adverse scenario in millions of euros.

2 This capital shortfall measure has been implemented based on Acharya at al. (2012) and Brownlees and Engle (2013) and. The data are provided by New York University's VLAB (http://vlab.stern.nyu.edu/welcome/risk/). The theoretical motivation for the measure can be found in Acharya et al. (2016). SRISK has been documented to be a comprehensive measure that includes losses due to both a bank's investments in assets and its exposure to fragile liabilities, which in the current European context relate, respectively, to holdings of peripheral sovereign bonds and (short-term) funding risk such as US money market fund withdrawals and other wholesale investors (Figure 5).

b. **Three-year cumulative impairment losses on financial and non-financial assets in the banking book ('loan losses'):** The cumulative impairment losses are measured in the adverse scenario.

c. **Three-year cumulative losses from the stress in the trading book ('trading losses'):** The cumulative impairment losses are measured in the adverse scenario.

d. **Three-year cumulative total losses:** Total losses are the sum of loan losses and trading losses.

We analyse and compare the benchmark regulatory capital shortfalls along two dimensions: the absolute size of the shortfalls, and the rank correlation of banks that incur shortfalls.

The calculation of capital shortfalls considers the losses banks incur in the banking and trading book. The ECB then calculates shortfalls using a regulatory capital ratio (the CET1 ratio). This ratio incorporates risk-weighted assets in the denominator. Moreover, the numerator is CET1 capital introduced by the Basel III framework and implemented in the EU in the capital requirements regulation and directive (CRR/CRD IV). These choices are problematic for two reasons:

1. **The use of risk weights** is questionable, as they are based on internal models for banks using the basic or advanced internal ratings-based modelling approach (IRB banks). Even in the standardised approach, risk weights do not necessarily reflect the true risk of the banks' assets (for example, sovereign debt still has a zero risk weight).
2. **Common Tier 1** is a 'new' measure of regulatory capital and incorporates a substantial number of transitional arrangements until it is fully implemented, i.e. a number of regulatory deductions from capital are going to be phased-in over time. Subtracting goodwill and other intangible assets is one example; the treatment of deferred tax assets (DTAs) is another example. Recognising these items, however, can be decided by the national competent authorities and thus gives them

considerable discretion. This discretion was heavily used in the comprehensive assessment as reported by the ECB.[3]

We thus compare the benchmark stress test results both to the capital shortfalls as calculated by the EBA/ECB and to the actual losses in the banking and trading book under the adverse scenario that form the basis of these shortfalls. The advantage of using these losses is that they are unaffected by risk weights or regulatory discretion.

Major results

Comprehensive assessment outcomes

- The regulatory capital shortfall as estimated by the ECB is €19.8 billion. Public banks thus account for more than 80% of the total capital shortfall reported by the ECB (€24.6 billion) (Table 2).
- Losses in the banking book (loan losses) and in the trading book (trading losses) are large and amount to €275 billion and €37 billion, respectively (Table 2).

Comparison of our benchmark capital shortfalls with comprehensive assessment shortfalls

- Our benchmark capital shortfall estimates amount to €450 billion for the 39 publicly listed banks. The countries with the largest expected shortfalls in a systemic crisis are France (€189 billion), Germany (€102 billion) and Italy (€76 billion). Malta and Slovakia (whose banking systems are among the smallest in the Eurozone) have no capital shortfalls under our benchmark estimates (Table 2).

3 The removal of the prudential filter on unrealised gains or losses on sovereign exposures held in the available-for-sale (AFS) portfolio is a notable exception. EBA-defined harmonised rules require a transitional phase-in of gains or losses (2014: 20%; 2015: 40%; 2016: 60%) (ECB, 2014). The ECB recognises that "there is a need to improve the consistency of capital and in particular the treatment of the deductions and the related quality of CET1 capital. This will be an issue for the SSM to address as a matter of priority".

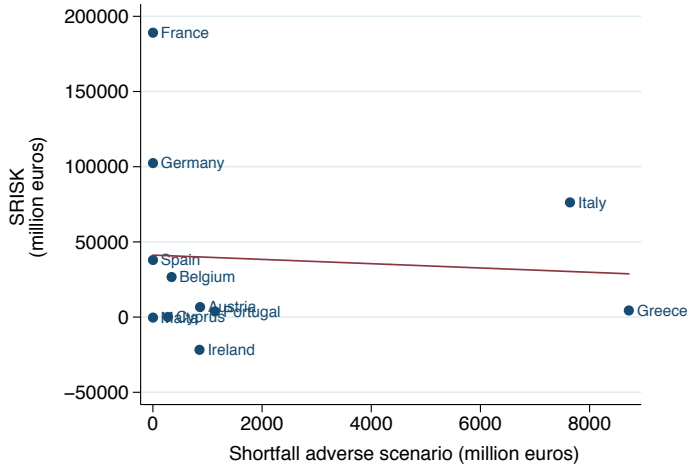
Table 2 Shortfall estimates

Country	SRISK 5.5% VLAB	Supervisory stress test results				
		Shortfall 5.5% CET1	Loan losses	Trading losses	Total losses	Banks
France	189,042	0	64,718	13,692	78,410	3
Germany	102,406	0	16,364	8,222	24,586	3
Italy	76,287	7,640	79,196	5,920	85,116	11
Spain	37,914	0	63,243	5,082	68,325	5
Belgium	26,616	339	5,766	1,877	7,642	2
Austria	6,677	865	8,694	626	9,320	2
Greece	4,360	8,721	21,836	1,142	22,978	4
Portugal	3,821	1,137	4,189	347	4,536	2
Ireland	3,053	855	9,785	487	10,272	3
Cyprus	167	277	529	11	540	1
Malta	0	0	200	22	222	2
Slovakia	0	0	0	0	0	1
Total	450,343	19,834	274,520	37,428	311,948	39

Notes: This table reports capital shortfalls under our benchmark stress test as well as based on the supervisory stress test as disclosed by the ECB. SRISK 5.5% VLAB is calculated assuming a 5.5% prudential capital ratio (which is the measure available on the NYU Stern Volatility Lab website) as of 31 December 2013. Shortfall 5.5% CET1 is the shortfall to the 5.5% common equity Tier 1 capital ratio in the adverse scenario. Loan losses are three-year cumulative impairment losses on financial and non-financial assets in the banking book. Trading losses arethree-year cumulative losses from the stress in the trading book. Total losses is the sum of loan losses and trading losses. Losses are incurred in the adverse scenario. Banks are the number of public banks in each country.

- The size of the regulatory capital shortfalls is less than 5% of the estimates using our benchmark stress test (Table 2).
- While the 5 largest banking systems (measured by total assets of banks in our sample), i.e. France, Germany, Italy, Spain and Belgium, have an estimated capital shortfall of €432 billion using our benchmark stress test, they have less than €8 billion shortfall in the adverse scenario of the regulatory assessment (Table 2).
- Capital shortfalls estimated under our benchmark stress tests are weakly but in fact *negatively* correlated with the supervisory shortfalls (Figure 1). The rank correlations reported in Table 3 support this negative association.

Figure 1 SRISK versus the shortfall adverse scenario

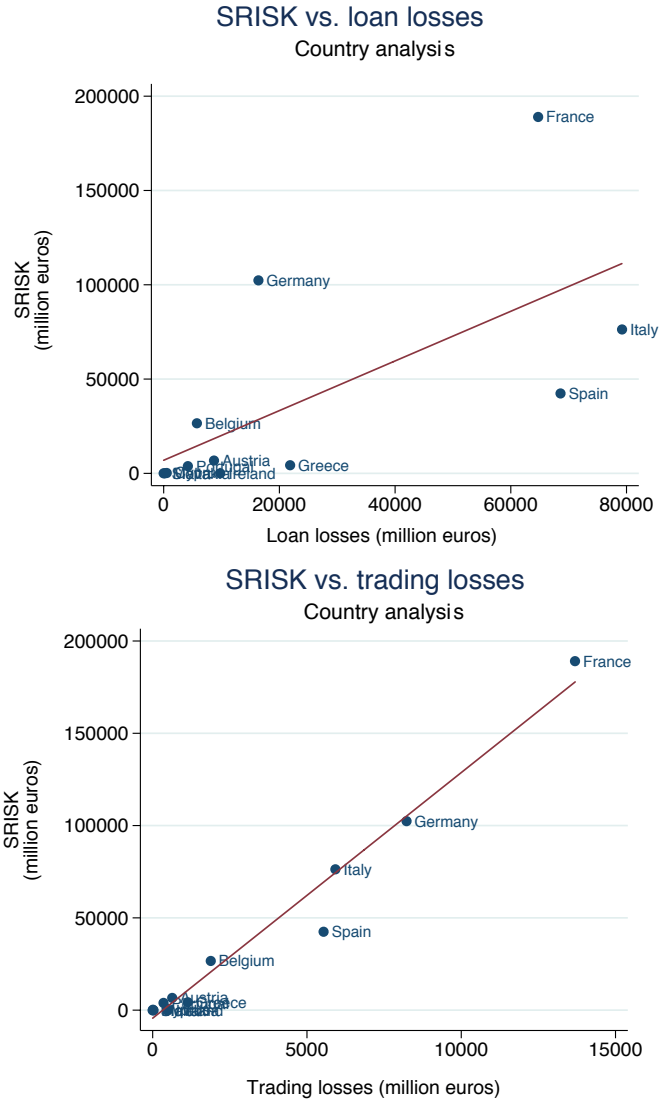


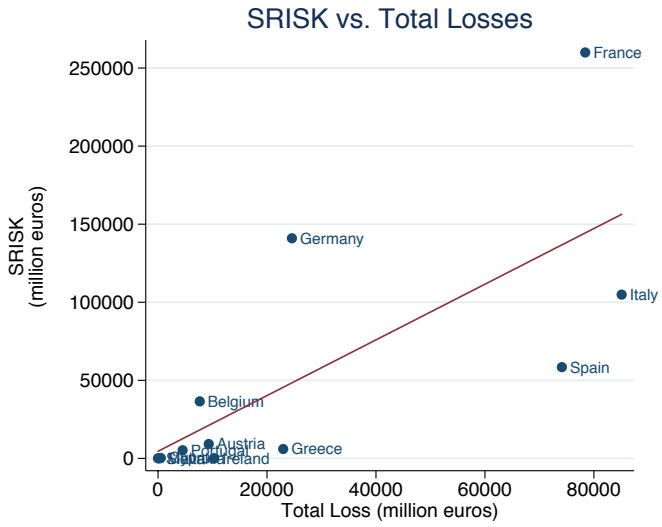
Notes: This figure plots SRISK as of 31 December 2013 against shortfall in the adverse scenario. Shortfall estimates are in millions of euros and aggregated over all public banks within each country.

Comparison of our benchmark capital shortfalls with comprehensive assessment losses

- The capital shortfalls estimated under our benchmark stress tests are highly correlated with the actual losses under the adverse scenario both in the banking book (rank correlation of 0.761) and the trading book (rank correlation of 0.937) (Figure 2 and Table 3).

Figure 2 SRISK versus loan losses





Notes: This figure plots SRISK as of 31 December 2013 against three-year cumulative loan losses, trading losses and total losses (loan losses + trading losses) in the adverse scenario. SRISK and loss estimates are in millions of euros and aggregated over all public banks within each country.

Table 3 Rank correlations

	SRISK 5.5% VLAB
Shortfall 5.5% CET 1	-0.058
Loan losses	0.761**
Trading losses	0.937**
Total losses	0.827**

Notes: Reports rank correlations of regulatory stress test capital shortfalls and losses with SRISK 5% VLAB. SRISK 5.5% VLAB is calculated assuming a 5.5% prudential capital ratio (which is the measure available on the NYU Stern Volatility Lab website) as of 31 December 2013. Shortfall 5.5% CET 1 is the shortfall to the 5.5% common equity Tier 1 capital ratio in the adverse scenario. Loan Losses are three-year cumulative impairment losses on financial and non-financial assets in the banking book. Trading losses are three-year cumulative losses from the stress in the trading book. Total losses is the sum of loan losses and Trading losses. Losses are incurred in the adverse scenario. ** indicates significance at the 1% level.

Implications

In Acharya and Steffen (2014), we provide a number of benchmark stress testing models to estimate capital shortfalls during a systemic crisis. The analyses suggest possible capital shortfalls between €80 billion and more than €700 billion depending on the respective model. The regulatory capital shortfall disclosed by the ECB on 26 October 2014 revealed a capital shortfall under an adverse scenario of €24.6 billion, of which €19.8 billion can be attributed to publicly listed banks.

The negative correlation between our benchmark estimates and the regulatory capital shortfall, but a positive correlation between our benchmark estimates and regulatory estimates of losses, suggests that regulatory stress test outcomes are potentially heavily affected by a) discretion of national regulators in measuring what is 'capital', and, especially, b) the use of risk-weighted assets in calculating the prudential capital requirement.

This highlights the importance of using multiple benchmark leverage ratios, such as the market-based approach we employ and simple leverage ratio (which is not affected by regulatory risk weights).

Moreover, the differences between the shortfalls we estimated in Acharya and Steffen (2014) and the ECB's estimates appear to be driven by the large banks in large countries such as France and Germany. No capital shortfall was identified for these banks during the comprehensive assessment. This is possibly due to the fact that systemic risk and feedback effects from the financial sector in the real sector, which are captured in the market data, have been completely ignored in regulatory assessment (Steffen 2014).

References

Acharya, V. and S. Steffen (2014), "Falling Short of Expectations – Stress Testing the Eurozone Banking System", Working Paper, NYU Stern School of Business.

Acharya, V., R. Engle and M. Richardson (2012), "Capital Shortfall: A New Approach to Ranking and Regulating Systemic Risks", *American Economic Review Papers & Proceedings* 102(3), pp. 59–64.

Acharya, V. V., L. Pedersen, T. Philippon, and M. Richardson (2016), "Measuring Systemic Risk", *Review of Financial Studies*, forthcoming.

Brownlees, C. and R. Engle (2013), "Volatility, Correlation and Tails for Systemic Risk Measurement", Working Paper, NYU Stern School of Business.

ECB (2014), *Aggregate Report on the Comprehensive Assessment*, Frankfurt am Main.

Steffen, S. (2014) "Robustness, validity and significance of the ECB's asset quality review and stress test exercise", Report requested by the European Parliament's Economic and Monetary Affairs Committee, October.

Appendix 1

Notes: This table reports descriptive statistics of the publicly listed banks included in the comprehensive assessment conducted by the European Central Bank (ECB) in 2014. C Tier 1 is the core Tier 1 ratio and is core Tier 1 capital divided by risk-weighted assets (RWAs). Equity/assets is book equity over total assets. Market-to-book is market value over book value of equity. Market cap is the market value of equity measured in millions of euros. Assets are total assets and measured in millions of euros. All data are as of 31 December 2013.

Bank	Country	Ticker	Assets	Market cap	C Tier 1	Equity/assets	Market-to-book
Erste Group Bank	Austria	EBS	200,118	10,922	11.44%	7.39%	0.74
Österreichische Volksbanken	Austria	VBPS	20,904	530	13.56%	5.84%	0.43
Dexia	Belgium	DEXB	222,936	78	21.24%	1.78%	0.02
KBC Group	Belgium	KBC	238,686	17,227	11.70%	6.08%	1.19
Hellenic Bank	Cyprus	HB	6,384	229	7.34%	6.25%	0.57
Crédit Agricole SA	France	ACA	1,519,089	23,316	9.96%	3.15%	0.49
BNP Paribas	France	BNP	1,810,522	70,611	11.73%	5.02%	0.78
Société Générale	France	GLE	1,214,193	33,769	10.68%	4.44%	0.63
Deutsche Bank	Germany	DBK	1,611,400	35,466	12.83%	3.41%	0.65
Commerzbank	Germany	CBK	549,654	13,375	13.06%	4.90%	0.50
Aareal Bank	Germany	ARL	42,981	1,729	15.93%	5.70%	0.71
Eurobank Ergasias	Greece	EUROB	77,586	3,029	10.43%	5.83%	0.67
Piraeus Bank	Greece	TPEIR	92,010	7,770	13.88%	9.28%	0.91
National Bank of Greece	Greece	ETE	110,930	9,242	10.28%	7.10%	1.17
Alpha Bank	Greece	ALPHA	73,697	6,905	16.06%	11.35%	0.83
Bank of Ireland	Ireland	BKIR	132,133	8,170	12.23%	5.97%	1.04
Permanent TSB Group Hldgs Plc	Ireland	IPM	37,601	1,646	13.11%	6.34%	0.69
Allied Irish Banks	Ireland	ALBK	117,734	58,487	14.56%	8.91%	5.57

Bank	Country	Ticker	Assets	Market cap	C Tier 1	Equity/assets	Market-to-book
UniCredit	Italy	UCG	827,538	31,267	10.57%	6.05%	0.62
Intesa Sanpaolo	Italy	ISP	624,179	29,269	11.33%	7.22%	0.65
Banca Monte dei Paschi	Italy	BMPS	198,461	2,056	10.65%	3.11%	0.33
Banco Popolare	Italy	BP	126,043	2,467	9.69%	6.76%	0.29
UBI Banca	Italy	UBI	124,242	4,466	12.60%	9.00%	0.40
Banca popolare dell'Emilia	Italy	BPE	61,758	2,317	8.56%	7.63%	0.49
Banca Popolare di Milano	Italy	PMI	49,353	1,458	7.21%	7.39%	0.40
Banca Carige	Italy	CRG	42,156	974	5.09%	3.90%	0.59
Mediobanca	Italy	MB	72,841	5,495	11.75%	9.54%	0.79
Banca Popolare di Sondrio	Italy	BPSO	32,770	1,295	7.89%	6.14%	0.64
Credito Emiliano	Italy	CE	31,531	1,939	9.94%	6.84%	0.90
Bank of Valletta	Malta	BOV	7,258	797	11.67%	7.95%	1.38
HSBC Bank Malta	Malta	HSB	5,722	759	9.39%	7.39%	1.80
Millennium BCP	Portugal	BCP	82,007	3,285	14.19%	3.99%	1.00
Banco BPI	Portugal	BPI	42,700	1,693	16.54%	5.40%	0.73
VUB banka	Slovakia	VUB	11,556	964	15.93%	11.94%	0.70
Banco Santander	Spain	SAN	1,115,637	73,826	11.71%	7.16%	0.92
BBVA	Spain	BBVA	599,517	51,866	11.59%	7.48%	1.16
Banco de Sabadell	Spain	SAB	163,442	7,590	11.96%	6.37%	0.73
Banco Popular Español	Spain	POP	146,709	8,327	10.69%	7.92%	0.72
Bankinter	Spain	BKT	55,136	4,474	12.85%	6.17%	1.31

Appendix 2

Notes: This table reports capital shortfalls under our benchmark stress test as well as based on the supervisory stress test as disclosed by the ECB. SRISK 5.5% VLAB is calculated assuming a 5.5% prudential capital ratio (which is the measure available on the NYU Stern Volatility Lab website) as of 31 December 2013. Shortfall 5.5% CET 1 is the shortfall to the 5.5% common equity Tier 1 capital ratio in the adverse scenario. Loan lare three-year cumulative impairment losses on financial and non-financial assets in the banking book. Trading losses are three-year cumulative losses from the stress in the trading book. Total losses is the sum of loan losses and Trading Losses. Losses are incurred in the adverse scenario.

Bank	Country	Ticker	SRISK 5.5% VLAB	Shortfall 5.5% CET 1	Loan losses	Trading losses	Total losses
Erste Group Bank AG	Austria	EBS	5,932	0	7,719	569	8,288
Österreichische Volksbanken	Austria	VBPS	745	865	975	57	1,032
Dexia SA	Belgium	DEXB	21,354	339	1,111	524	1,635
KBC Group NV	Belgium	KBC	5,262	0	4,654	1,353	6,007
Hellenic Bank Public Company Limited	Cyprus	HB	167	277	529	11	540
Crédit Agricole SA	France	ACA	81,523	0	25,138	2,339	27,477
BNP Paribas SA	France	BNP	58,034	0	25,228	6,788	32,016
Société Générale	France	GLE	49,485	0	14,353	4,564	18,917
Deutsche Bank AG	Germany	DBK	76,598	0	9,411	5,312	14,723
Commerzbank AG	Germany	CBK	24,246	0	6,622	2,868	9,490
Aareal Bank AG	Germany	ARL	1,562	0	331	42	373
Eurobank Ergasias SA	Greece	EUROB	2,471	4,628	5,291	189	5,479
Piraeus Bank SA	Greece	TPEIR	1,146	660	4,202	228	4,430
National Bank of Greece SA	Greece	ETE	597	3,433	7,314	518	7,832
Alpha Bank AE	Greece	ALPHA	145	0	5,029	207	5,236
Bank of Ireland	Ireland	BIR	2,161	0	4,289	157	4,446
Permanent TSB Group Holdings Plc	Ireland	IL0	892	855	1,300	4	1,304
Allied Irish Banks, Plc	Ireland	AIB	0	0	4,196	326	4,522

Bank	Country	Ticker	SRISK 5.5% VLAB	Shortfall 5.5% CET 1	Loan losses	Trading losses	Total losses
UniCredit SpA	Italy	UCG	30,361	0	25,199	2,096	27,295
Intesa Sanpaolo SpA	Italy	ISP	18,698	0	21,147	1,452	22,599
Banca Monte dei Paschi di Siena SpA	Italy	BMPS	9,865	4,250	8,699	583	9,282
Banco Popolare Societˆ Cooperativa	Italy	BP	5,528	427	5,483	407	5,889
Unione di Banche Italiane SCpA	Italy	UBI	3,881	0	7,106	134	7,240
Banca popolare dell'Emilia Romagna SC	Italy	BPE	1,881	128	2,666	162	2,828
Banca Popolare di Milano Scarl	Italy	PMI	1,845	684	1,706	147	1,853
Banca Carige SpA - Cassa di Risparmio di Genova e Imperia	Italy	CRG	1,725	1,835	1,922	66	1,989
Mediobanca - Banca di Credito Finanziario SpA	Italy	MB	1,028	0	2,943	605	3,547
Banca Popolare di Sondrio SCpA	Italy	BPSO	1,020	318	1,767	209	1,976
Credito Emiliano SpA	Italy	CE	455	0	558	61	619
HSBC Bank Malta Plc	Malta	HSB	0	0	98	1	99
Bank of Valletta Plc	Malta	BOV	0	0	200	22	222
Banco Comercial Portugus SA	Portugal	BCP	2,701	1,137	3,149	337	3,486
Banco BPI SA	Portugal	BPI	1,120	0	1,040	10	1,050
VUB banka	Slovakia	1VUB02AE	0	0	293	32	325
Banco Santander SA	Spain	SAN	23,832	0	36,661	2,758	39,420
Banco Bilbao Vizcaya Argentaria, SA	Spain	BBVA	5,611	0	15,880	2,079	17,959
Banco de Sabadell, SA	Spain	SAB	4,334	0	3,927	92	4,018
Banco Popular Espa–ol SA	Spain	POP	3,690	0	5,194	107	5,300
Bankinter SA	Spain	BKT	448	0	1,581	46	1,627

8 Stress testing in wartime and in peacetime

Til Schuermann[1]
Oliver Wyman and Wharton Financial Institutions Center

Introduction

Stress testing as a bank supervisory tool is rapidly spreading across the globe. From Australia, the US, Canada and the UK, to Scandinavia and continental Europe, stress testing is becoming a familiar sight to banks and bank supervisors alike. The tool that helped draw a line under the recent financial crisis by providing clarity on the health of the banking system is migrating into ordinary and everyday risk management as well as regulation and supervision. It worked so well in the crisis – in wartime – why not keep using it in peacetime?

In this chapter, I consider the merits and demerits of porting stress testing from the theatre of crisis fighting to peacetime bank supervision. The success of stress testing, starting with the US in 2009 with the Supervisory Capital Assessment Program (SCAP), has resulted in new laws (such has the Dodd-Frank Act in the US) and in regular supervisory stress testing programmes on both sides of the Atlantic. System-wide stress testing leverages the only informational advantage that the supervisor has over the banks: the ability to compare exposures, vulnerabilities, models, and resilience to shocks across firms. In every other way, the supervised are informationally advantaged vis à vis the supervisors. This horizontal perspective is one of the great strengths of a stress testing programme at any time.

1 I would like to thank Tim Colyer, Doug Elliott, Ugur Koyluoglu, James Mackintosh and Martin Scheicher for very helpful comments and suggestions, and Ron Anderson for his encouragement. All remaining errors are mine, of course.

Yet, effective as it was in helping end the financial crisis, the objectives and goals are different in peacetime. In wartime, a central objective is to provide credible insight into the health of bank balance sheets in an effort to stabilise the banking system and thus minimise damage to the real economy. Banks are notoriously opaque even at the best of times (Morgan, 2002). Since existing metrics such as risk-weighted assets (RWAs) and regulatory capital requirements were not really informative – banks were well capitalised by those metrics – a new approach was needed.

In the depth of the crisis, the credibility problem in banks was matched by the lack of credibility in their supervisors. The metrics by which supervisors judge bank solvency – those involving RWAs – were not believable. For that reason, significant and unprecedented disclosure of the stress testing process and details of results was needed so that the market could, effectively, check the supervisors' maths.

It is hard to restore confidence if there is no solution to a credibly revealed problem. The presence of a credible financial backstop is critical to allow authorities to comfortably reveal the true depth of the capital hole in the banking system. For example, of the 19 banks in the SCAP, ten needed capital. One of these – GMAC – needed to be nationalised and thus drew on the government backstop funded by TARP;[2] it was the only one. The other nine were able to fill their hole through a combination of fresh capital raises from the market, many of which were oversubscribed, and through retained earnings.

So what are necessary elements for a successful wartime stress test? First, the proposed scenario needs to be severe enough to probe the vulnerabilities of the banks and the banking system. Scenario design touches on what to stress, i.e. which risk factors to focus on (for instance, housing prices, unemployment or equity prices) and how much to stress them.

Second, the scenario needs to have a correspondingly severe translation to capital impacts via increased losses and reduced profitability. A harsh scenario that results in only a modest capital impact is hard to believe and may even exacerbate the credibility problem. Third, the disclosure regime has to be sufficiently detailed to allow verification

2 https://www.treasury.gov/initiatives/financial-stability/TARP-Programs/bank-investment-programs/Pages/default.aspx

of process and results. And fourth, there needs to be a credible capital backstop by the government in case banks who need to raise capital can't do so on their own in a timely manner (for the US case, six months following the release of SCAP results). Indeed, without a credible backstop, supervisors may be reluctant to propose a harsh enough scenario and/or provide sufficiently conservative translations of that scenario to the loss, profitability and ultimately capital impact outcomes. The perceived credibility of a backstop underscores the close tie between bank and sovereign risk.

We know much less about what a successful peacetime stress testing programme should look like. If wartime stress testing is about revealing the capital hole and filling it – i.e. getting capital into the banks – then peacetime must be a state where the hole, credibly sized, has been (or is being) filled, and credibility in both the banks and their supervisor(s) has been restored. At this stage, the relevant question is whether the banks, either individually (microprudential) or collectively (macroprudential), are sufficiently resilient to withstand real economy and financial shocks, given their strategic business objectives and plans. In other words, it is no longer about just surviving (that has already been demonstrated by the recently completed wartime stress test) but rather about ensuring that banks have the capacity to keep lending and to provide other key financial services.

The Federal Reserve has taken this a step further. The CCAR programme, by virtue of being a capital planning exercise, requires banks to demonstrate that they have sufficient financial resources (capital) to support their business plans even under severely adverse conditions. If so, banks can afford to pursue those strategies, and they may be aggressive, or capital could actually be returned to shareholders in the form of increased dividends or share repurchases. If wartime stress testing is about getting capital into the banking system, peacetime is about deciding whether to let it out.

Resilience can be attained and evaluated in two ways: quantitatively and qualitatively. The quantitative question is simple enough: is there enough capital to support the risks taken on by the bank? But bank supervisors traditionally spend their time evaluating whether banks' practices are safe and sound, in other words, bank supervisors regularly conduct qualitative assessments of banks. Are risk and capital/liquidity management practices at banks up to supervisory expectations – are they 'good enough'? Stress testing can provide new insights into this old question, as is elaborated below.

Anatomy of a stress testing program

Let's consider a three-component framework of stress testing to help parse this peacetime problem: scenario, modelling/forecasting machine and results. For each component, there are a set of questions to answer.

Scenario design

Who should design the scenario? How severe should it be, and severe for whom? What shape should it take (e.g. a sharp decline followed by gradual recovery, or a gradual decline and no recovery) and how long should the projection horizon be? How often should one run such a stress test?

A practical issue plaguing supervisors and bank risk managers alike is how to effectively span the state space. Which are the relevant risk factors to focus on, and how many do you need to sufficiently cover the exposures or vulnerabilities of the banks? For example, the SCAP in 2009 made do with just three risk factors, all domestic: GDP growth, the unemployment rate and a residential house price index. By 2014, the US stress scenario state space had grown to 16 domestic variables and three variables (GDP, inflation and FX) across each of four non-US regions (the UK, the Eurozone, Japan, and developing Asia) for a total of 28 variables.

The 2014 stress test in Europe had a far bigger challenge as it needed to cover 28 countries, of which 18 were part of the Eurozone, where the stress test was part of a wider 'comprehensive assessment'.[3] About ten risk factors were specified for each country, in addition to FX between euro and non-euro EU countries, plus two (GDP and inflation) for each of 20 regions comprising the rest of the world.

Banks with significant capital markets activities (sales and trading, investment banking) regularly cover a much larger set of financial risk factors. In the US, the six largest

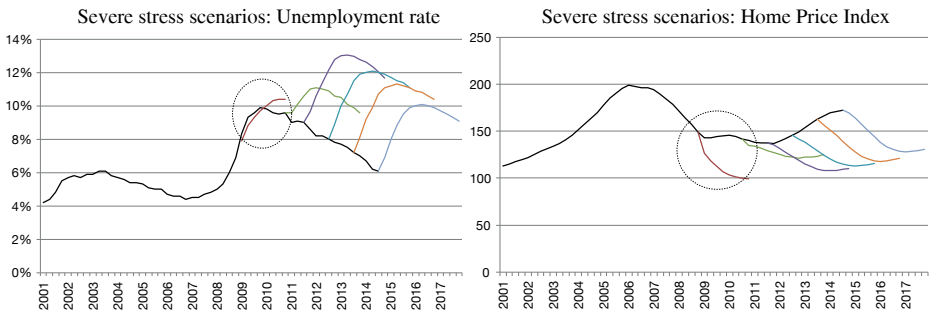
3 The comprehensive assessment comprised two elements: 1) an asset quality review (AQR) to enhance the transparency of bank exposures, including the adequacy of asset and collateral valuation and related provisions and 2) the stress test itself, performed in close cooperation with the European Banking Authority (EBA). Lithuania joined the Eurozone on 1 January 2015.

banks are required to conduct a global market shock (GMS) on their trading book. For CCAR-2015, the Fed specified about 24,000 parameters across about 20 categories such as equities, FX, rates, energy and commodities, securitised products, credit correlation and so on. The 2014 exercise in Europe had approximately 950 parameters plus about 580 sovereign haircuts (by country and maturity).

When considering the state space numbering in the dozen or so (per country) plus trading shocks for capital markets banks numbering in the thousands, the complexity of the scenario design problem becomes daunting. The challenges of building a model that can generate a set of coherent scenarios at such high dimensions are formidable![4]

Assuming one can agree on the state space and on the severity, expressed perhaps as a likelihood of occurrence (say one in 100), with multiple risk factors there is a continuum of equi-probable scenarios to choose from. Here we need some information about the banks' vulnerabilities. The SCAP provides an interesting example. Within a month of the start of the exercise, the realised unemployment rate already exceeded the projected rate under the stress scenario (Figure 1, left panel). However, US banks were most vulnerable to a decline in housing, and that risk factor was indeed appropriately stressed (Figure 1, right panel).

Figure 1 Federal Reserve severely adverse scenarios, 2009-2015



Note: Black line depicts realisations, and coloured lines denote stress scenario paths across different stress tests.

4 Kapinos and Mitnik (2015) have proposed a dimension reduction approach to top-down stress testing.

The chosen supervisory stress scenario will not be equally stressful for all banks. For instance, banks with modest loan exposures, particularly to mortgages, would be little harmed by stresses to house prices. Unless all banks have similar business models and exposures, it is hard to design a scenario that challenges the capital position of all banks.

To solve this problem, supervisors have to either design a myriad of scenarios to address a range of business models and exposure profiles, or they can ask banks to design their own scenarios in such a way as to probe their specific vulnerabilities. In this way, one would expect the scenario of, say, an internationally active trust and custody bank to be rather different from that of a regional commercial bank with a purely domestic business focus.

This approach of requiring both supervisory and bank designed scenarios is followed by the Federal Reserve's CCAR program. In Europe, bank-designed scenarios are executed separately from the supervisory tests under Basel's ICAAP.[5] By asking banks to design bespoke stress scenarios of at least equal severity to the generic supervisory scenario and run them side by side, supervisors are able to glean rich qualitative information about a bank's ability to effectively identify its risks and vulnerabilities, which is foundational to effective risk management. It is hard to overstate the value of this information to supervisors given their difficult task of assessing whether a bank's practices are safe and sound.

A stress test effectively functions as a highly bespoke set of risk weights (Acharya et al. 2014). Ordinary regulatory risk weighting is largely invariant to the state of the world, current or expected. Moreover, risk weights are applied only to the left side of the balance sheet (assets) instead of the full set of bank financials (balance sheet and income statement) that impact the capital position of the bank.[6] When the risks change, so should the assessment (or weighting) of risks.

Stress scenarios can be countercyclical by leaning against the wind, much like monetary policy. Indeed, the Bank of England is explicitly using the stress scenario to help calibrate the Basel III countercyclical capital buffer of banks. The Federal

5 ICAAP: internal capital adequacy assessment process; see, for instance, Bank of England (2015).

6 To be sure, both ordinary risk weighting as well as stress testing considers off balance sheet exposures.

Reserve's scenario design regime does not go so far, but it does anchor the scenarios in the unemployment rate: it needs to increase by at least four percentage points but have a peak rate of no less than 10%.[7] For example, the unemployment rate in November 2015 is 5%, so the expectation for the next US stress test scenario is a doubling to 10%. This explicitly system-wide approach is a key feature of what is meant by 'macroprudential' (Demekas, 2015).

Regulatory capital regimes such as Basel I, II and III have typically been calibrated to a one-year horizon. Stress test horizons to date are typically two years (SCAP) to three years (EBA), and up to five years (Bank of England); the Fed's CCAR programme is for nine quarters. A horizon of two to three years roughly matches the asset duration of a typical commercial bank. An investment bank where the balance sheet is dominated by trading assets is much shorter duration, while insurance assets are much longer duration.

There is another possibly significant distinction between the US and European approaches. The scenario provided by the Fed and the PRA is published in quarterly time steps, and the EBA in annual time steps.[8] The projections of the bank outcomes – losses, profitability and, of course, capital – is done in quarterly time steps under the Fed regime, and annual time steps on the other side of the Atlantic. Given how quickly banks can fall into insolvency following a shock and the uneven timing of loss realisations across asset classes, as well as profitability dynamics, the more granular or higher frequency approach is likely better able to uncover vulnerabilities.

Since CCAR-2014, the Fed has also required the eight global systemically important banks (GSIBs) to conduct a counterparty default scenario as part of the overall stress test exercise.[9] This is an interesting addition to the test, as the stress scenario on its own may not be harsh enough to result in a default of a bank's largest counterparties, but such a default would clearly be quite stressful to the bank and yield valuable insights to

7 http://www.federalreserve.gov/newsevents/press/bcreg/20131107a.htm.

8 Federal Reserve: http://www.federalreserve.gov/newsevents/press/bcreg/2015-macro-scenario-tables.xlsx; EBA: https://www.eba.europa.eu/documents/10180/669262/2014-04-29_ESRB_Adverse_macroeconomic_scenario_-_specification_and_results_finall_version.pdf

9 G7 sovereigns and designated clearing counterparties are excluded.

the supervisor. Moreover, as part of the exercise banks are required to report their top 20 or so counterparties to the supervisor, who can use this data to stitch together at least part of the network of interdependencies across the largest financial institutions, a good example of gaining macroprudential insights from microprudential data.

Models and projections

With the scenario design settled, the next task is to translate that macro scenario into the micro outcomes of interest: losses, profitability and capital impact. Should these models be built by both banks and supervisors? How much leeway should banks have in building these models, in mapping the scenario to capital impact of their bank? How can supervisors effectively evaluate the reasonableness of the banks' projections?

All crisis stress testing involved supervisory models to some degree, and the US has gone furthest in building out and relying on such models for peacetime stress testing. An independent assessment with supervisory models using bank data was viewed as critical to the success of the SCAP (Hirtle et al. 2009). The stark difference in results from the initial European exercises in 2010 and 2011 and subsequent country-specific stress tests are at least in part attributable to the rigorous use of independent models in the latter. Two notable examples are Ireland in 2011 and Spain in 2012. The Irish stress test revealed a capital need for the four largest Irish banks of €24 billion only months after the CEBS-2010 EU-wide exercise found none of the four Irish banks needing any capital (see Central Bank of Ireland 2011). None of the 25 Spanish banks that participated in the EBA-2011 exercise was required to raise capital following the test, yet a subsequent exercise in 2012 revealed a capital need of €57bn across seven of the then 14 participating banking entities.[10] Frankly it seems hard to imagine how a supervisor could effectively evaluate the impact of scenarios on bank balance sheets (and income statements) without building their own models.

Acharya et al. (2014) point out that the early European stress tests suffered less from mild scenarios or from loss impacts that were too benign. Rather, they allowed

10 There had been a number of bank mergers in 2011 which reduced the number of entities without reducing the coverage of the banking system; see Oliver Wyman (2012).

profitability projections through the stress scenario that were perhaps overly optimistic. For many banks, profitability more than covered projected losses, implying no capital consumption at all. To illustrate this, Table 1 compares two US stress tests (2009 and 2015) and two European exercises (2011 and 2014). The average coverage ratio in the SCAP was 61%, and only 16% of the banks were projected to cover their losses with profits. This rose to 80% and 23%, respectively, for the peacetime stress test in 2015. The EBA exercise in 2011, at a time of acute sovereign risk concerns in Europe, allowed 42% of the tested banks to more than cover their projected losses with profits. By contrast, the 2014 exercise was markedly harsher by this metric: the average loss coverage ratio was 44%, and just 11% of the banks were able to cover the projected losses with profits.[11]

Table 1 Loss coverage ratios across stress tests

	Average coverage ratio	% of banks with coverage ratio >1
SCAP 2009	61%	16%
CCAR 2015	80%	23%
EBA 2011	75%	42%
CA 2014	44%	11%

Note: Coverage ratio = pre-provision profits over total losses.

How much freedom should banks have in modelling the impact of stress scenarios? US regulators have provided relatively few constraints in modelling approaches and parameters, while European regulators, perhaps concerned with banks' overly optimistic projections, have constrained some of the loss and profitability projection parameterisation. US supervisors have similar concerns, but with a richly built-out supervisory modelling machinery it is easier to confront any undue optimism more directly. There is no better way to understand a modelling challenge than to build it yourself.

The supervisory value of allowing for a richer diversity of modelling approaches could be quite high. Supervisors may learn from the range of modelling practice to both

11 To be sure, the maturity of modeling profitability under stress is much less than for loss models (Duane et al. 2014), which have been the longtime focus of bank risk managers.

improve their own modelling as well as guide best practices as they emerge and evolve in the industry. Moreover, some freedom of development mitigates against models monoculture, a risk that certainly exists in any model intensive risk management and supervisory process (Hirtle and Lehnert 2014). However, it is a resource-intensive proposition. The supervisory models and corresponding modelling expertise need to exist in the first place.

Modelling energy by the supervisor can be focused more on the microprudential problem – forming an independent view of bank-level results at a fairly granular level – or more on the macroprudential problem – understanding the system-wide effects involving spillovers into the rest of the financial system and the real economy. The Bank of England (2015) has made it clear that, for the medium term at least, their modelling energy will be focused on the latter. While the Fed has not been explicit about its view, judging by the enormous effort devoted to modelling granular bank-level financials, it seems the focus is more on the former.

If the peacetime objective is to ensure that the banking system, individually (for the largest, most systemically important banks) and perhaps collectively, is resilient to shocks, then the more microprudential focus of the Fed seems appropriate. If, however, the objective is to understand the resilience of the financial and real economic ecosystem to shocks, then the more macroprudential orientation of the Bank of England seems desirable. To be sure, the more an economy's credit formation and financial intermediation is dominated by banks, the less system-wide insights are lost by just focusing, in detail, on banks. The US, however, is one of the least bank-dependent industrialised economies.

Use and disclosure of results

The most creative and stringent scenario translated to capital impact outcomes with robust and conservative models is not enough; the results need to be properly communicated to the public. How much disclosure? And who should disclose?

Credibility of results is critical for effective crisis response. To quote Ong and Pazarbasioglu (2013, p. 1), "Credibility is the bedrock of any crisis stress test". A generous disclosure regime is needed for this to happen, both in terms of process and

outcomes/results. Both banks and supervisors need to overcome market scepticism, and transparency is the only way to regain lost confidence. The stress test has to produce genuinely new information (e.g. new insights into bank balance sheet resilience), and this information has to be credible.[12]

A lot of information on its own is not enough. The EBA in 2011 provided very rich disclosures, down to well-organized Excel files containing data and results,[13] but the projected capital need was too small to be credible (€2.5bn for the 90 participating banks). Examples of the right mix of transparency and credible results are the SCAP (2009), Ireland (2011), Spain (2012), and the 2014 comprehensive assessment by the EBA and ECB.[14]

Once credibility is re-established and the financial system is on the mend, more careful thought may need to be given to the peacetime disclosure regime. What is the purpose, and what information should be conveyed, and who should convey it? Goldstein and Sapra (2013) describe well the benefits and costs of stress test disclosures, pointing out, among other things, that the signal from the regulator, now highly credible, may drown out the signal from the bank (via its own disclosure), both hampering market discipline and inducing strategic behaviour in bank disclosures. My own view is that in a crisis, the supervisor ought to be very transparent and have rich disclosure. However, in peacetime a more modest disclosure regime may be desirable. Meanwhile, banks should disclose generously at all times to promote monitoring and market discipline.

Other than to disclose results, what decisions are or can be made in peacetime? In a crisis, the main decisions hinge on the capital shortfall projected by the stress test. That shortfall may be so acute as to question the viability of the bank forcing more drastic action (e.g. nationalisation or liquidation). A bank can either raise capital on its own

12 Gorton and Tallman (2015) make an interesting argument that crisis response before the Fed came into being involved suppressing information about individual banks but providing it about the system as a whole.

13 http://www.eba.europa.eu/documents/10180/15935/2011+EU-wide+stress+test+disclosure+templates+-+revised.xls/cc3c791a-af03-4c7b-a48f-97f671fd6bf0

14 SCAP: http://www.federalreserve.gov/newsevents/press/bcreg/20090507a.htm; Ireland (2011): http://www.centralbank.ie/regulation/industry-sectors/credit-institutions/Pages/FinancialMeasuresProgramme.aspx; Spain (2011): http://www.bde.es/f/webbde/GAP/Secciones/SalaPrensa/InformacionInteres/ReestructuracionSectorFinanciero/Ficheros/en/informe_oliverwymane.pdf; EBA/ECB (2014): http://www.eba.europa.eu/-/eba-publishes-2014-eu-wide-stress-test-results; https://www.bankingsupervision.europa.eu/banking/comprehensive/2014/html/index.en.html.

or make use of capital backstops and other programmes (e.g. a government-sponsored 'bad bank') to meet the revealed capital need. Banks that need capital are typically required to submit a capital plan that lays out how, and by when, the bank plans to fill the capital shortfall.

In peacetime, what supervisory action is contemplated? There may be further capital shortfalls arrived at through quantitative approaches developed in the crisis (and carried forward). Such shortfalls or stress test failures should be rare, however, if the wartime stress test was sufficiently severe, but they may occur from time to time.

What, then, of the wealth of qualitative information gathered in the course of a stress test, especially a peacetime stress test performed without the time pressure and exigency that comes with financial crisis management? The Federal Reserve, with its CCAR programme, has made use of this qualitative information to prevent firms from distributing capital back to their shareholders. Such 'failure' based on qualitative assessment is in addition to possible failure for quantitative reasons (outright projected capital shortfall). This has not (yet) been implemented by other supervisors. Note that the US has a mixed model, where banks subject to the DFAST (typically smaller banks) but not the CCAR programme are not exposed to the risk of such qualitative failures; see Hirtle and Lehnert (2014) for a discussion.

The bridge between wartime and peacetime is the capital plan. Recall that revealed capital shortfalls from a crisis stress test result in a capital plan, developed by the bank but approved by the regulator, which maps out how the bank will close the identified capital gap. The CCAR programme, ostensibly, is a capital planning exercise. Banks must demonstrate that their business/strategic plan for the coming nine quarters would survive stressful scenarios, one (or more) designed by the supervisor and one (or more) of their own design.

Other applications of stress testing

Stress testing is finding wider uses and applications. With increasing emphasis on recovery and resolution planning (RRP), the stress testing framework and machinery can be readily used to stress the bank more and more to the point of needing to take

drastic action, such as asset sales or business disposition to raise funds (recovery mode), or all the way to a terminal shock. Large US banks are making use of their CCAR stress testing machinery to satisfy RRP requirements.

Increasingly, the stress testing and projection capabilities are being used by banks for ordinary business decisions like budgeting and performance assessment. To satisfy stress testing requirements, the bank needs to generate dynamic projections of their balance sheet and income statements, conditional on a realisation of the economy and financial markets in the form of a two to five-year path. One of those realisations is a baseline scenario, the most likely path the economy will take. A natural application is to the budgeting process, and increasingly CCAR banks are discarding their old, more heuristic budgeting approach and making use of the baseline forecast from their CCAR process.

Essentially, the stress testing modelling machinery allows the bank to understand how much of its performance is driven by macro risk factors – unemployment, house prices, interest rates – which are largely beyond management control, and therefore what extra push is needed to achieve desired performance goals above and beyond what the economy and the market can deliver organically. The stress test machine can help a bank be much more articulate about the drivers of its business plan and thus be able to separate performance into 'alpha' and 'beta'.

Concluding thoughts

Stress testing as the dominant supervisory tool is no panacea. The pre-crisis supervisory regime of the GSEs in the US is a spectacular failure of using stress testing to set required capital levels (Frame et al. 2015). That regime involved two sets of interest rate shocks and was left unchanged after its implementation in 2002. This experience, however, points to one of the real values of stress testing, properly done, both from a micro- and macroprudential practice: risks are time varying, and so the risk assessment or risk weighting should also be time varying (Acharya et al. 2014).

Stress testing is not new; it is a well-worn component in the risk manager's toolkit. But comprehensive and dynamic stress testing with explicit conditioning on observable

macroeconomic and financial risk factors requiring projections of firm financials – balance sheet and income statement – is recent. It has provided banks, and supervisors, with new insights into their vulnerabilities. It has pushed the culture of senior management to be more imaginative in identifying risks and using stress testing (and reverse stress testing) as a way of evaluating forward looking actions. When aggregated across many banks, the whole provides more information than just the sum of the parts.

This discussion has focused just on capital stress testing, yet banks typically run short of liquidity long before they run short of capital. Much less has been written about liquidity stress testing, and to date little guidance has emerged from the regulatory and supervisory community about what a good liquidity stress testing process ought to look like.

As we leave the financial crisis behind, stress testing in peacetime is rapidly becoming the supervisory tool of choice – for quantitative assessment of resilience and capital adequacy; for qualitative insights into the risk and capital management practices of banks. The concreteness of the tested scenario – unemployment going to, say, 10%; house prices declining by 20%; equity market volatility doubling – is one of the real virtues of this risk management approach. It is tangible and lends itself to clear understanding by senior management, boards of directors, supervisors and the public. Stress scenarios are a tangible expression of risk appetite, one of the hardest, yet most important preference parameters in risk management, whether by the bank or by the supervisor. Given the inherent opacity of bank balance sheets, such clarity is especially welcome.

References

Acharya, V., R. Engle and D. Pierret (2014), "Testing Macroprudential Stress Tests: The Risk of Regulatory Risk Weights," NBER Working Paper No. 18968, Cambridge, MA.

Bank of England (2015), "The Bank of England's approach to stress testing the UK banking system" (http://www.bankofengland.co.uk/financialstability/Documents/stresstesting/2015/approach.pdf).

Central Bank of Ireland (2011), *The Financial Measures Programme Report*, Dublin.

Demekas, D.G. (2015), "Designing effective macroprudential stress tests: progress so far and the way forward", IMF Working Paper No. WP/15/146, Washington, DC.

Duane, M.P., T. Schuermann and P. Reynolds (2014), "Stress Testing Bank Profitability", *Journal of Risk Management in Financial Institutions* 7(1), pp. 72-84.

Frame, W.S., K. Gerardi and P. Willen (2015), "The Failure of Supervisory Stress Testing: Fannie Mae, Freddie Mac, and OFHEO", Federal Reserve Bank of Boston Working Paper No. 15-4 (http://ssrn.com/abstract=2675697).

Goldstein, I. and H. Sapra, 2013, "Should Banks' Stress Test Results be Disclosed? An Analysis of the Costs and Benefits", *Foundations and Trends in Finance*, vol. 8, pp. 1-54.

Gorton, G. and E. Tallman (2015), "How Did Pre-Fed Banking Panics End?" Available at http://faculty.som.yale.edu/garygorton/documents/GortonTallmanPaper_June_9_2015.pdf

Hirtle, B.J., T. Schuermann and K. Stiroh (2009), "Macroprudential Supervision of Financial Institutions: Lessons from the SCAP", Federal Reserve Bank of New York Staff Report No. 409.

Hirtle, B.J. and A. Lehnert (2014), "Supervisory Stress Tests", Federal Reserve Bank of New York Staff Report No. 696.

Kapinos, P.S. and O.A. Mitnik (2015), "A Top-down Approach to Stress-testing Banks", *Journal of Financial Services Research*, October.

Morgan, D.P. (2002), "Rating banks: risk and uncertainty in an opaque industry", *American Economic Review 92*(4), pp. 874-888.

Oliver Wyman (2012), *Asset Quality Review and Bottom-up Stress Test Exercise*, Madrid (http://www.bde.es/f/webbde/SSICOM/20120928/informe_ow280912e.pdf).

Ong, L.L. and C. Pazarbasioglu (2013), "Credibility and Crisis Stress Testing", IMF Working Paper WP/13/178, Washington, DC.

9 In praise of stress tests

Charles A. E. Goodhart
London School of Economics

Introduction

Bank regulation is necessary, but difficult to do well. It is necessary because the social costs of bank failures exceed the private costs, i.e. there are externalities. It has proven difficult to do such regulation effectively because there are a number of generic problems with the conduct of bank regulation. In this chapter, I will deal specifically with two such failings.

First, regulation has historically and, even still currently, been based mainly on ratio controls, initially mainly cash and liquidity ratios, more recently on capital ratios. But there was, until recently, never any graduated ladder of sanctions as the ratio levels held declined towards the acceptable minimum. Instead, ratio control operated as a cliff-edge device. Fall below it, and either the authorities or market forces (reputational) would require either speedy repair or closure. But once above, by whatever small buffer, the bank conformed. Moreover, given that the required ratio was set above that which a bank would voluntarily choose to keep on its own account (if not so, why have a regulatory ratio?), the effective buffer above the required minimum was generally kept at fairly minimal levels (Milne and Whalley 2001).

Second, whenever a bank gets into difficulties, it would be preferable if it could take steps on its own accord to recover its function as a going concern, rather than proceed into resolution. Banks are often too important, too large and too central as utilities – just like other utilities such as electricity, water and gas – to close down and liquidate. Indeed, both bailout and bail-in mechanisms are means of keeping banks operating as going concerns, under different management. But there are inherent problems both with bailout and bail-in techniques (Goodhart and Avgouleas 2014). It would be

usually preferable if a bank, running into difficulties, would put its own house in order, rather than having to go through an external support mechanism. The EU banking directive is called the Bank Recovery and Resolution Directive. But far more effort and attention has been given to 'resolution' than to 'recovery'. Nor can we expect banks to be keen to initiate recovery programmes on their own. To do so is tantamount to a confession that the previous strategy had not worked, and would, very likely, involve a major management reshuffle. Bank CEOs, such as Fuld and Goodwin, are no more likely to call time on themselves than turkeys are to vote for Christmas.

It is the thesis of this chapter that the regular conduct of stress tests will help to mitigate these generic shortcomings. Anyhow, in the next section I rehearse some of the problems of ratio controls. In the subsequent section I review some of the difficulties of getting banks to initiate recovery programs. Then in the final section, I briefly discuss some of the remaining problems of stress testing and conclude.

Some problems with ratio controls

One of the most astute macroeconomists ever was D.H. Robertson. In his customary, gently humorous, way, he saw several of the problems with such ratio controls in his early textbook on *Money*:

> *"If a proportion fixed by custom is arbitrary and misleading, a proportion fixed by law seems at first sight to be positively mischievous. An iron ration which you must not touch even in the throes of starvation is something of a mockery. Against such criticism it may be urged (though not too loudly) that in finance as in war rules are made to be broken on occasion, and that their object is not to ensure that certain things shall never be done, but that they shall not be done without good reason."* (Robertson 1922, p. 57)

Then again:

> *"But it is evident at once that the fixed fiduciary system is free from the defect which we noted, in connection with bank money, in the proportional system — namely, that it resembles the procedure of a certain municipality which tried to*

guard against a shortage of cabs by ordaining that there should always be at least one cab on the ranks." (p. 62)

Robertson also appreciated the point that, if the reserve ratio was required to be held (at all times) and the penalty cost in terms of interest foregone was sizeable, then the spare margin that the banks would voluntarily hold, above that required, would most likely be minimal. Again, he wrote:

"A legal arrangement of this kind is open to the objection that human nature being what it is, the law is sometimes held to encourage what it does not expressly forbid, and a bank may therefore be tempted to keep its proportion of reserves very near the bed-rock legal minimum. Any unexpected demand for common money may then present the bank with the alternative of infringing the law, or declaring itself insolvent while its reserves are still far from exhausted." (p. 56)

As a result, the use of required ratios, that could not be infringed without severe reputational damage, was accompanied by relatively small holdings of surplus capital (or liquid assets). So the potential buffering function of ratio controls was low, perhaps lower even than would have occurred without any such regulations. What such controls did do was to provide some extra support for bank creditors in the face of severely adverse conditions, and so *may* have reduced creditors' propensity to run.

These problems with ratio controls have persisted, despite having been spotted many decades ago. Thus, when he was Chairman of the Basel Committee on Banking Supervision at the time of introducing the Basel I Capital Adequacy Requirements (CARs), Peter Cooke raised the issue:

"On the level of capital, Cooke raised the question whether the number was to be 'a minimum*, or target or standard'. 'If the former, it will have to be set at a very low level. If the latter, it can be higher but will require a transition period for some countries and/or some banks in some countries'. Here I believe that the analysis can be faulted.* First, there was little appreciation that I found in the papers that if the regulators set a 'target' of X per cent, that banks, the market and ratings agencies would come to treat it as a reputational minimum. Second, if it was to be*

a target or standard, it implies that it is conceivable, indeed fully acceptable under some circumstances, that banks might fall below such a level. But there was no discussion at all of how national supervisors should react, and what sanctions they might apply, if the CAR standard was breached, but it would have been difficult or impossible for an international committee such as the BCBS to go into this latter subject.

*[*Peter Cooke has responded to me, in private correspondence, that members of the Committee, in his recollection, were aware of both these points. On the first, it was realised that the target/standard would come to be regarded as a reputational minimum (although the degree to which this occurred was not anticipated). On the second, members of the Committee recognized that the standard could be breeched in particular circumstances, but it would be up to individual national supervisors to decide on the action which should then be taken.]"* (Goodhart 2011, pp. 177-178).

The main problem was, and remains, that central bank officials felt that they had no locus for imposing sanctions on commercial banks (Goodhart 2011, Chapter 14). They could advise on the appropriate level of ratios, leaving it largely to market reputational factors to enforce their maintenance, and, if the required ratios were too egregiously transgressed, they had the statutory right to withdraw the banking licence, and hence shut down the bank. But this led to a cliff-edge system, with banks largely free to do as they chose so long as they remained (often just) on the right side of the requirements, but with most of them maintaining small spare buffers above the cliff-edge minimum. So the system as a whole was fragile in the face of any sizeable common shock.

The alternative, which many academic commentators would have preferred, would have been a ladder of sanctions, increasing in severity as the bank held less capital/liquidity, but probably with a lower absolute minimum level below which a bank would be shut. There are, however, now signs that the authorities have started to move in this direction. In particular, the Basel III capital adequacy requirement (CAR), as promulgated by the BCBS and Financial Stability Board (FSB), contains a 'conservation' range wherein a bank's core equity tier one ratio (CET1) lies between 7% and 4.5% of risk-weighted

assets (RWAs), with 4.5% being the absolute minimum level. Within this conservation range, a bank would not be able to make certain payments, e.g. on dividends, equity buy-backs, acquisitions, and so on. So, it *would* face sanctions, short of closure.

Perhaps even more important, however, the application of regular stress tests does, in principle, allow the regulatory authorities to test whether banks' buffers would be strong enough to resist seriously adverse conditions, and to require strengthening additions to those buffers when they are assessed to be insufficient. There are, however, two shibboleths of political correctness in banking that complicate the issue. The first is the need for transparency. So, any bank that is deemed to fail the stress test must be named, and is consequently shamed, and will, therefore, have markets turned against it. This causes the whole process to become distorted – banks will try to game the exercise by setting their resources at levels that will just satisfy the authorities' presumed requirements; and the authorities will try to set their initial shock assumptions at levels that will just find a preordained set of banks failing the test, not too many, nor too few. If the exercise was done behind closed doors, without fear of leaks, the stress tests could be done more honestly. As it stands, an outsider, like myself, reckons that there will be a modicum of stage management about the whole exercise. Insiders will protest that everything is clean and straightforward, but we will remain sceptical.

Two economists, Fernando Alvarez and Gadi Barlevy, have recently written a (highly mathematical) paper on "Mandatory Disclosure and Financial Contagion" (Alvarez and Barlevy 2015), coming to a similar conclusion:

> *"Our model does not imply disclosure is always desirable, even in the presence of contagion. To the contrary, in our benchmark model not only is disclosure sometimes undesirable, but it may be optimal to force banks to keep information hidden. This is because secrecy can sustain socially beneficial risk-sharing between banks. The notion that opacity is desirable for sustaining insurance dates back to Hirshleifer (1971), and has been recently applied to explain the tendency towards secrecy in the banking sector by Goldstein and Leitner (2013), Faria-e Castro, Martinez, and Philippon (2015), Dang et al. (2014). As in these papers, our benchmark model implies mandatory disclosure cannot improve welfare in normal times, in contrast to the view advocated in Bernanke (2013)."* (p. 2)

The second shibboleth is that there must be no further bailouts using 'taxpayer funds'. Given that the public failure of a stress test makes a weak bank even more fragile by turning markets against it, the most efficient way of rescuing the situation is for an immediate public sector injection of capital into such failing bank(s), as was done with the use of TARP funds in the first US stress test of 2009, bringing capital requirements back into line without any need for further deleveraging. The terms on which such injections would be made should be sufficiently onerous so as to provide a clear incentive to the bank(s) involved to replace government funding with private sector funding as soon as reasonably possible. Indeed, if it was not sufficiently onerous, in the EU it would transgress the prohibition on state aid. Moreover, if the bank failing the stress test could commit to raising the required additional capital within a short space of time, it should be allowed to do so, rather than accept public funding. While a bank poised to fail a stress test might seek to raise new equity shortly before the results of a stress test were published, the market would draw its own cynical conclusions.

The combination of these two shibboleths – transparency of outcome and no public sector financing – leaves the stress test process in a more complex and delicate condition than would ideally be desirable. If banks go into the stress tests in a reasonably strong condition, as has been the case in the US since 2010, then the shortcomings of the weaker banks can generally be fixed by a temporary prohibition on pay-outs for dividends, buy-backs, excessive remuneration, and so on, until retained earnings can fill the perceived hole. But if the banks are so weak as to need additional external finance, as may still be the case in some instances in Europe, then there is a problem. Weak banks would be further weakened by public test failure, and there is no clear way out of this problem. Raising further equity at a time when market values would have been trashed by the exercise would be enormously unpopular with both existing shareholders and management; merger with a stronger bank may raise concerns about competition and oligopoly; deleveraging will have adverse macroeconomic effects (see Orphanides 2014).

Perhaps some further thought still needs to be applied to the question of how stress tests could best be arranged so as not only to identify, but also to encourage a resultant strengthening of the position of the weaker banks in the system. This should be a situation where public funding for equity injections could be used efficiently and

profitably to the benefit of taxpayers, but the blanket opposition to any further bailouts prevents the adoption of sensible policy.

Recovery

In my view (see Goodhart and Avgouleas 2014), the bail-in proposals for resolving systemic banks (GSIBs) are seriously flawed. The resolution process will be triggered too late to prevent large-scale losses. Under the TLAC and MREL proposals, such losses will fall immediately, and in a concentrated fashion, on a much smaller group of creditors. If institutional investors, such as pension funds and insurance companies, are to be prevented from putting their clients at risk by investing heavily in such bail-inable debt, the residual buyers (e.g. hedge funds and sovereign wealth funds) may be thin on the ground and flighty. The potentiality for contagion and procyclicality is clear.

But this train left the station some time ago; there is no prospect of reversal. The need, therefore, is to shore up the system so that the likelihood of a GSIB entering the resolution process is reduced as far as possible. There is such a possibility, in the guise of the recovery stage of the bank recovery and resolution procedure, as in the EU's Bank Recovery and Resolution Directive (BRRD). All systemic financial intermediaries (GSIFIs), especially such banks (GSIBs), are now required to write living wills, whereby they indicate how they would handle a condition of significant danger (recovery) and then how, if such fragility developed further into failure (from going concern to gone concern), the resulting resolution could be facilitated. The public has been informed that the regulatory authorities have found the initial attempts by the GSIBs in the US to be unsatisfactory. Presumably, some mutually acceptable programme for each GSIB's recovery and resolution planning will eventually be obtained, though, for confidentiality reasons, the details will presumably be kept secret.

So, I assume that all GSIBs and GSIFIs will, in due course, develop a recovery programme acceptable to their regulator – the home regulator when the GSIB has a single point of entry (SPOE), the college of regulators for GSIBs with multiple points of entry (MPOE). The problem, therefore, is not so much *what* (recovery consists of) but *when* (it is triggered). At present, there is no clarity on this. Indeed, the EBA

appears to be inviting banks to select a trigger for initiating recovery for themselves. But, for rather obvious reasons, no bank's CEO is going to start such a process himself.

In the absence of any such effective trigger, Miguel Segoviano and I wrote a paper proposing that the trigger should depend on the probability of default, as estimated from observable market data (e.g. of bank equity valuation and volatility), using various alternative formulae (Goodhart and Segoviano 2015). There is reason to expect, however, that the use of market data in this way will not prove acceptable either to the regulators or to the regulated. The regulators will not want to give up discretion, and the regulated will argue, sometimes justifiably, that such market valuations are distorted, manipulated and inaccurate, especially during panics and crises and under the influence of short sellers. Moreover, linking the initiation of recovery programmes to market valuations could lead to certain 'cliff-edge' problems.

Be that as it may, much more emphasis should be placed on the determination of, and threshold for, the recovery stage. There are several reasons why this needs to be done. First, bankers, if left to themselves, are likely to enter the recovery stage voluntarily far too late. A concern about reputation, should the news leak (reputational stigma), and the likelihood that top management will be overly self-confident in their ability to keep going successfully (think of Fuld and Goodwin), will combine to make management reluctant to call time on themselves.

Second, the authorities are pressing banks to hold significant quantities of contingent convertible bonds (CoCos), but seem to want to keep discretion over the occasion and timing of them being triggered. Such discretion makes their pricing much more problematic. If this trigger were related instead to a well-defined principle for initiating recovery, this problem would be lessened. But it seems unlikely that the authorities will be prepared to cede their discretionary powers in favour of any quasi-automatic trigger mechanism.

But if the authorities want to use discretion in order to initiate a bank's recovery programme, how can they get sufficiently up-to-date and reasonably accurate data to do so? Once again, the answer would seem to be that this could be obtained courtesy of the annual stress tests. These stress tests *should* provide the authorities with an early warning signal of which banks were flirting perilously close to the danger area, should

a severe adverse shock occur. Put another way, if a bank collapsed in year t having sailed easily through the prior stress test in year t-1, there would have been something amiss with that test. Of course, the chosen scenario for the stress test in any year may diverge considerably from the shock that actually occurs to weaken the bank, but doing a new stress test each year, with changing parameters, should give the authorities an increasingly rounded picture of each bank's strengths and weaknesses.

Moreover, the accounting/statistical basis of a stress test, i.e. mark-to-crisis, is exactly that which regulators should want to use, *not* the misleading mark-to-market measure (see Caccioli et al. 2012).

Some remaining problems

Of course, much can go wrong with stress tests. They cannot, or cannot easily, take account of second-round – and subsequent round – interacting, amplifying effects. The particular set of stresses imagined for the purpose of the test may be far removed from those that actually occur. In part this could be because those setting the tests do not want to consider the possibility that their own policies could fail (e.g. a break-up of the Eurozone). They are time and resource consuming, and so can only be done occasionally, usually once a year. There are a range of other problems – some technical – with both the design and conduct of stress tests (e.g. that the tests have focussed primarily on risk-weighted assets rather than on the more encompassing, and better predictive, simple leverage ratios), which it is not my purpose to pursue here

Instead, the conclusion here is that, fallible as they may be, the conduct of annual stress tests gives the regulatory authorities their best available chance of dealing with fragile banks while there is still enough time to avert a, potentially contagious, failure. The key requirement is to have ready-made plans on the shelf in advance for how best to back-stop the weakest banks. This could be by some combination of forced retained earnings, forced raising of additional external equity, injection of public sector funding, or initiation of the recovery programme. How this might be done needs to be considered and reviewed before the exercise is completed.

Assuming that this latter can be done, then stress tests may become the most effective tool available to regulators. Thus Cecchetti (2015) writes:

"Stress tests may be the most powerful prudential tool we have at our disposal for safeguarding the resilience of the financial system. They take seriously the fact that, when a large common shock hits, there is no one to sell assets to or raise capital from. By ensuring that each individual institution can withstand significant stress, we ensure the system can, too. And, importantly, by adjusting the stress scenarios, prudential authorities can maintain resilience. At least in principle, they can both account for changes in the distribution of the shocks and ensure that the amplification potential of the propagation mechanism does not increase …

By changing the stress scenarios, prudential authorities are changing the level of capital that banks are required to hold. Passing tests with higher stresses necessarily requires more capital. And, the target is a given level of systemic resilience – resilience that requires both being able to withstand larger shocks and being able to mitigate the extent to which a given shock is transmitted to the economic and financial system more broadly.

I should note that some people would view this as simply a way of implementing a countercyclical buffer. That is, rather than rely mechanically [on] an indicator like credit growth, authorities would instead use stress testing as a way to calibrate the required amount of capital. There is clearly a sense in which the objectives are the same – maintaining systemic resilience – just the method of getting there is different. My sense is that stress testing is more flexible, faster, and less politically contentious than Basel III's countercyclical capital buffer.

We are still early in the process of developing prudential policy aimed at reducing the harmful impact of the asset price or lending busts that inevitably follow the booms. Over the years, through a combination of thought and experimentation, we can hope to develop a better articulated set of models that help us to understand what tools to bring to bear and when. But until we do, I believe that stress tests will be the most powerful tool we have in our effort to maintain systemic resilience."

While I strongly agree with Cecchetti's claim about the potential powers of the use of stress test, it does, by the same token, raise a question about how such additional powers may be made accountable and proportionate. At present, the relevant authorities can choose any scenario for the stress test that they think fit, apply proprietary and undisclosed modelling to assess the outcome, and use the results to jack up required capital (and liquidity) to whatever level they prefer without a by-your-leave from either the affected banks or the legislature. On the other hand, giving a right of appeal to the banks (and appeal to what body?) would both lengthen the whole process and make it much more expensive. Perhaps the relevant authorities should be required to report the conduct and outcome of each stress test to a select committee of the legislature, and be prepared to justify their actions resulting from such testing to that same committee. This process would preferably mostly be in public, but might have to be in private and confidential when referring to developments of a particular tested institution.

Be that as it may, stress testing is likely to remain such a central and powerful tool in the armoury of the regulatory authorities that its constitutional and legal setting deserves further thought.

Of course, the main problem that many expert commentators see with the *current* conduct of stress tests (e.g. Dowd 2015, Goldstein forthcoming 2016-17) is that these tests, especially the European ones, have not been rigorous and tough enough. But this is primarily, in my view, because the European authorities have not yet fully resolved the question, raised earlier, of how to provide back-stop funding to recapitalise the weaker banks. But when this hurdle is, we would hope, eventually overcome, the opposite problem of going too far may hail into sight.

References

Alvarez, F. and G. Barlevy (2015), "Mandatory Disclosure and Financial Contagion", NBER Working Paper 21328, Cambridge, MA.

Bernanke, B.S. (2013), "Monitoring the Financial System", speech at the 49th Annual Conference on Bank Structure and Competition, May (http://www.federalreserve.gov/newsevents/speech/bernanke20130510a.pdf).

Caccioli, F., J.-P. Bouchaud and J.-D. Farmer (2012), "A proposal for impact-adjusted valuation: Critrical leverage and execution risk" (http://arxiv.org/pdf/1204.0922.pdf).

Cecchetti, S.G. (2015), "On the Separation of Monetary and Prudential Policy: How much of the pre-crisis consensus remains?", CEPR Discussion Paper 10949, London.

Dang, T.V., G. Gorton, B. Holmstrom and G. Ordonez (2014), "Banks as Secret Keepers", NBER Working Paper 20255, Cambridge, MA.

Dowd, K. (2015), "Central Bank Stress Tests: Mad, bad and dangerous", *Cato Journal* 35(4), pp. 507-524.

Faria-e Castro, M., J. Martinez and T. Philippon (2015), "Runs versus Lemons: Information Disclosure and Fiscal Capacity", NYU working paper.

Goldstein, I. and Y. Leitner (2013), "Stress Tests and Information Disclosure", Federal Reserve Bank of Philadelphia Working Paper.

Goldstein, M. (2016/17), *Banking's Final Exam: Lessons from the US and European Union Bank Stress Tests*, Washington, DC: Peterson Institute.

Goodhart, C. (2011), *The Basel Committee on Banking Supervision: A History of the Early Years, 1974-1997*, Cambridge, UK: Cambridge University Press.

Goodhart, C. and E. Avgouleas (2014), "A Critical Evaluation of Bail-in as a Bank Recapitalisation Mechanism", in F. Allen, E. Carletti and J. Gray (eds), *Bearing the Losses from Bank and Sovereign Default in the Eurozone*, Philadelphia, PA: Wharton Financial Institutions Center, pp. 65-97.

Goodhart, C. and M. Segoviano (2015), "Optimal Bank Recovery", International Monetary Fund Working Paper WP/15/217, Washington, DC.

Hirshleifer, J. (1971), "The Private and Social Value of Information and the Reward to Inventive Activity", *American Economic Review* 61(4), pp. 561-74.

Milne, A. and A. Whalley (2001), "Bank Capital Regulation and Incentives for Risk-Taking", Cass Business School Research Paper (http://ssrn.com/abstract=299319).

Orphanides, A. (2014), "Stress Tests as a Policy Tool: The European Experience During the Crisis", MIT Sloan School Working Paper 5122-14, Cambridge, MA.

Robertson, D. (1922), *Money*, London: Nisbet & Co.

10 Stress testing as a policy instrument: Some thoughts

Udaibir S. Das[1]
IMF

"The stress test was a decisive turning point. From then on, the U.S. banking system would strengthen steadily—and eventually the economy would follow."
—Ben Bernanke, Courage to Act, 2015

"Stress tests for banks are a predictable act of public theatre."
—Gillian Tett, Financial Times, 2015

"My idea is that ... in the future we should move to a yearly exercise, standard, simple, hopefully not so tense in terms of expectations around it."
—EBA Chairman Andrea Enria, 2014

Introduction

Stress testing of financial firms – especially banks – is nearly two decades old. Financial firms use it for portfolio risk management, in some cases to meet prudential and regulatory purposes. Traditionally, stress testing was developed by investment banks for individual portfolio applications, but increasingly the techniques have been used by authorities as supervisory and macroprudential tools, in particular after the Asian Crisis (1997) and the recent Global Crisis (2008).

1 This is the text of a presentation delivered at the London School of Economics, Systemic Risk Centre, "Conference on Stress Testing and Macroprudential Regulation: A Trans-Atlantic Assessment" (29-30 October 2015). The views expressed are those of the author and must in no manner be attributed to the IMF. I would like to thank Miguel Segoviano, Liliana Schumacher, Teng Teng Xu, and Christina Daniel and other IMF colleagues, for their inputs.

The IMF started to use stress testing in its 1999 Financial Sector Assessment Program (FSAP) after the Asian Crisis as a forward-looking way to assess risks and vulnerabilities in a financial system.[2] Before then, the typical vulnerability analysis was based on backward-looking indicators, such as Financial Soundness Indicators (FSIs), which were limited tools for the assessment of risks to financial stability. From 2000 to 2015, the FSAP concluded 300 risk assessments, including stress testing, in 170 jurisdictions across advanced economies, emerging economies and low-income countries under different institutional and data-availability conditions. The FSAP experience has helped highlight gaps in the literature and methodologies, raised awareness in the official sector on the use of stress tests as an input into policy making, and emphasised the continuing need for making further advances in stress testing techniques. The FSAP stress test work has been widely commented on. Overall, the comments have been balanced – they have been criticised by some, but supported by most.

Stress testing today is no longer a narrow technical exercise, but a key instrument underpinning financial policy decision making. The official sector responsible for the oversight and supervision of the financial system has been busy developing systems and capacity to implement regular stress test programmes at the individual firm level as well as the system-wide level. Apart from gaining an understanding of the robustness of financial balance sheets, the motivation underlying this trend is to help policymakers to better capture macro financial linkages between the financial system and the real sector (and vice versa). Most countries are still in the *early* to *intermediate* stages of implementation. Even those countries regarded as being in the *advanced* phase of implementation face many technical and operational challenges with respect to their

2 The FSAP replaced ad hoc series of assessments of the financial system. It has led to a distinct improvement in the IMF's ability to conduct financial sector surveillance and in understanding the key linkages between financial sector vulnerabilities and macroeconomic stability, prompted better discussions with authorities, and helped support policy and institutional reform.

stress testing programmes. Improvements are therefore ongoing at the country level, within international standard setting bodies, and at the IMF.[3,4]

Experience

The FSAP experience shows ample heterogeneity in stress testing approaches and practices across countries. In many cases, stress test objectives are not clearly defined and aligned with the use of stress test results for policymaking. The frequency of stress tests varies, although the current trend is to have annual regular exercises. Similarly, the approaches toward stress tests (whether top-down, bottom-up, or hybrid) are not always consistent or explained. Also, the scope and coverage of risks (solvency risk, market risk, and liquidity risk) are often uneven.

On technical aspects, several challenges exist. These relate to the difficulties in validating stress testing approaches; gaps in data collection and data quality assurance; concerns about how to choose the right level of severity and a relevant scenario; the omission of feedback effects in the estimation of risk parameters prevailing and their linkages to scenarios; coverage of all significant channels of risk transmission and feedback effects; and weak follow-up of stress test results with corresponding policy measures. The challenge today for the official sector is *how* to conduct stress tests and to *use* them appropriately. With a rise in public scrutiny and expectations, the importance of governance and integrity of stress testing has become as important as the focus on getting the modelling and stress test parameters right.

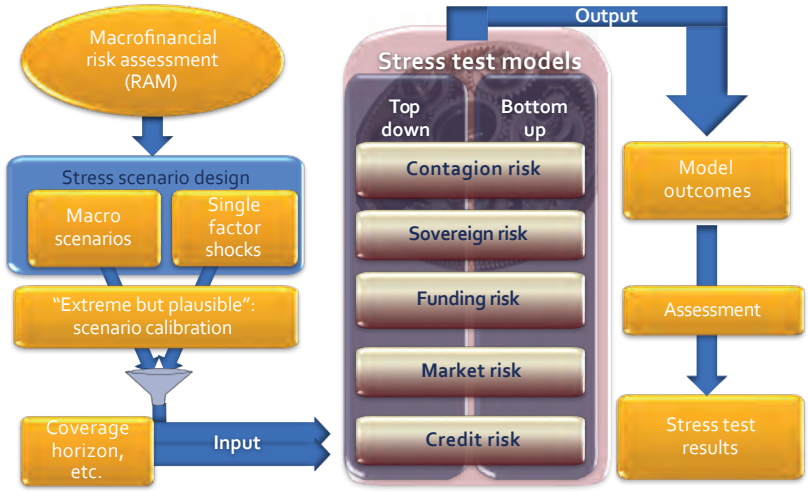
3 Bank supervisors, for example, are including stress tests as a regular feature in their standard supervisory processes. In an effort to improve banks' stress testing practices and provide authorities with proper guidance for their assessment, the Basel Committee of Banking Supervision (BCBS) developed a set of high-level principles for sound stress testing practices for banks, the Principles for sound stress testing practices and supervision, or 'Principles' (2009).

4 IMF staff are working on a project to define a set of good or desirable practices to guide authorities in the development of robust and effective authority-led "macroprudential stress test" (MaST) frameworks. The work will also try and delineate a set of benchmarks to assess the adequacy ofMaST frameworks for micro-and macroprudential purposes. These should help support more reliable and transparent design of stress testing frameworks and processes.

Stress testing

Stress tests permeate all major areas of IMF work, from surveillance to capacity development, to program work. In addition, the IMF offers technical assistance and cooperation on stress testing. However, the stress testing framework in FSAPs is the most prominent (see Figure 1). Stress testing in FSAPs always includes a top-down component, and often also a bottom-up element. Typically, the risks covered include credit, market, funding, sovereign and contagion risks. Results are assessed against regulatory ratios, which are used as hurdle rates.

Figure 1 Stress test workflow



The forward-looking approach that characterises stress testing starts globally at the IMF with the identification of risks at a global level. This represents the overarching risk assessment evaluation that guides all financial (multilateral and bilateral) surveillance. In addition to global risks, regional risks and country-specific risks are also identified for bilateral surveillance purposes. These three layers of risks would then feed into the scenario design of the stress test work as part of an FSAP. The stress test results in an FSAP could contribute to risks and vulnerability assessments in the IMF's Article IV consultations – the institution's periodic check up of their member countries' economies – and the country team's discussions with the authorities.

FSAP stress tests typically include the following exercises: (i) *solvency* resilience (using scenario and sensitivity analysis); (ii) *liquidity* resilience (using regulatory ratios and cash-flow approaches); and (iii) *contagion analysis* (impact of a shock transmitted from an individual firm to the rest of the financial system). On interconnectedness amongst institutions, FSAP reports have often commented on contagion risks.[5] However, mainly due to data limitations, most FSAP stress tests have used ad hoc approaches to the assessment of systemic risk and have tried to be 'roughly right rather than precisely wrong!' Methodologies have also varied. Most stress tests are based on balance sheet data. However, IMF staff has also developed approaches using market-based measures of distress, in particular for systemic risk assessment to cope with the lack of adequate accounting measures for systemic risk analysis.

The macrofinancial risk assessment in the risk assessment matrix (RAM) identifies the key external and internal risks in the economy and to the financial system.[6] These risk factors are incorporated into the overall macro scenarios (scenario analysis) and the single factor shocks (sensitivity analysis). The scenarios are calibrated based on the principle of extreme but plausible shocks, and serve as an input into stress testing models.

A typical stress testing exercise in an FSAP involves many decisions, some more visible than others, but all equally important for the reliability of the test results (see Figure 2). It begins with the coverage of financial institutions and activities. When stress tests cannot cover all financial institutions, the key is to cover those which are systemically

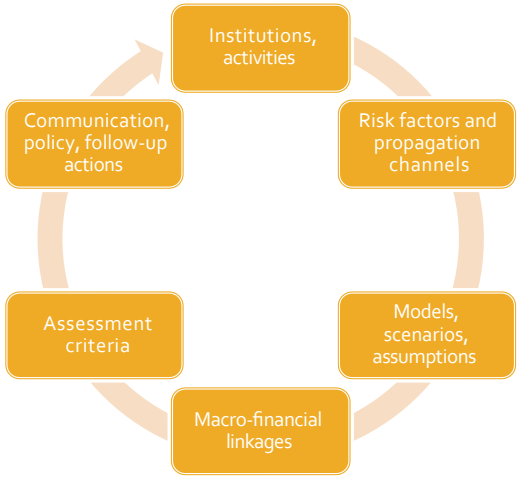
5 The first three countries where a quantitative attempt was made were Gabon, Luxembourg, and Sweden, as far back as in 2000. The Swedish and Luxembourg FSAP stress tests incorporated contagion analysis to estimate the impact of a bank's settlement failure on other banks' liquidity and capital.

6 The RAM is a structured framework for analyzing risks and potential impact at the IMF. It is also a standard feature of staff reports for Article IV consultations and the Financial System Stability Assessment (FSSA), the key publication for an FSAP. The RAM provides views on risks and levels of concern by identifying key vulnerabilities that could materially deviate from the WEO baseline path; the likelihood of realisation (low-medium-high); and the economic impact (low-medium-high). The RAMs in FSAPs focus on financial sector impact and macro feedback, and include a detailed description of transmission channels.

important. From the risk-assessment perspective, it is important to identify the key risk factors, internal and external, and the propagation channels.

Until recently, most of the FSAP effort was centred on banks. More recently, non-banks have been included and, at present, IMF staff undertakes their own stress tests (top-down) of insurance and asset management companies (in addition to bottom-up tests carried out by the institutions themselves) in countries where these institutions have systemic importance.

Figure 2 Stress tests: Key decisions



Constraints

Typically, stress tests have seven building blocks: (i) choice of the scenario and hurdle rates (to help define the degree of resilience that policymakers would like to enforce as a measure of financial stability); (ii) coverage of institutions and the scope of activities; (iii) risk factors, including models to assess impact and propagation channels (to understand macro financial linkages); (iv) assessment of data availability, its reliability, and granularity; (v) validating and interpretation of the results; (vi) communication across a range of stakeholders; and (vii) policy action supported by stress test results.

The interactions between these building blocks are often complex and country specific. The presumption is that the official sector 'knows its financial system' vis-á-vis major players, business models, transactions, and key counterparties. But this is not always the case. For a meaningful stress test, in-depth knowledge of the system (risks, business models, network mapping, and contagion channels) is essential. Structural aspects may have an impact on the stress scenario. Likewise, understanding key activities, markets, exposures, and counterparties ensures that all potential sources of risk in the stress test are factored in comprehensively.

Stress testing is often viewed as a 'black box'. The FSAPs have been placing great emphasis on published stress test results to explain the limitations implied by data and methodological approaches, and about scenario and severity choices. Greater 'health warnings' about the interpretation of results are also essential.

There is a considerable gap between the 'good practice' approaches to modelling shocks and those used in many other cases. For example, in contrast to good practice, some assessments have avoided analysing the consequences of politically sensitive shocks (e.g. sovereign defaults). The lack of high quality data on the financial system is often not emphasised sufficiently. In some countries, more caution is needed before using available statistical data at face value, either for stress testing or other analysis.

Despite the shortcomings, stress testing exercises have helped discipline policymaking. One can argue that the test severity (defined by the scenario choice and hurdle rates) implies a social choice in the social risk-return function. In this regard, stress tests raise a legitimate discussion on the trade-off between growth and financial stability (and the use of stress tests for this purpose) and important questions such as: Does the pursuance of financial stability (using stress tests as a barometer) inhibit growth by making relatively riskier projects too expensive to fund? Or instead, does it promote higher long-term growth by minimising market disruptions and the realisation of systemic risks over time? Most importantly, is stress testing a measure of financial stability? Is it a tool for societies to decide where in the social-return (GDP growth) frontier they want to be? If so, is stress testing a major tool to conduct an optimal choice of growth versus financial stability? And if so, who needs to make these decisions?

Stress test preconditions

A number of preconditions are important for the usefulness of stress testing in financial stability or macroprudential policymaking. The nine key elements are:

1. Broad support on the expected resilience of the financial system;
2. Top management support to monitor and stress the system;
3. Obtaining buy-in from all stakeholders;
4. Establishing a clear follow-up plan in line with the specific objective of the exercise;
5. Appropriately defining the institutional perimeter for the different stress tests;
6. Identifying all relevant channels of risk propagation and including all material risks and buffers;
7. Ensuring proper communication of stress test results;
8. Periodically keeping the stress test framework under top management review; and
9. Strengthening capacity and skilled resources.

The implementation of the principles would call for conservative estimates of the mapping of shocks onto losses and for severe scenarios as a means to overcome those limitations. Judgemental adjustments to statistical estimates would be required as well as thorough cross-checking of outputs from different models including by combining top-down and bottom-up approaches. Recognising the shortcomings of the stress test framework is a better basis for realistic results. Inter agency buy-in, as well as from the banks, is a precondition for the commitment of time and resources.

Systemic risk

A new wave of analytical, theoretical, and empirical approaches in stress testing is taking hold. The first wave lasted from the late 1990s to 2007, when the focus was on individual institutions. Today, the knowhow is moving towards (what I like to call) "systemic risk analysis" (SRA). Stress tests on individual institutions are becoming part of the SRA. In the wave underway, a 'piece-wise approach' that evaluates the vulnerability of the financial sector to single risk factors under various macroeconomic

stress scenarios is becoming common place.[7] In the period ahead, 'piece-wise' stress tests will be augmented by more regular and continuing macrofinancial (or macroprudential) stress tests, and topped up with a more formalised and regular SRA.

This is not as simple as it may sound. Several methodological challenges have to be confronted. Starting with a deeper and more thorough understanding of risk transmission channels, further research is needed on the use of macro stress testing techniques as an operational tool to help incorporate financial stability factors into macro policy decision making. SRA would also need to pay closer attention to the correlation of risks and risk measures over time and across all (or most) financial institutions and sectoral balance sheets, to the length of the time horizon used for simulations, and to the potential instability of all reduced-form parameter estimates because of feedback effects.

How quickly could we get to a stage when more integrated or systemic-level stress tests could be undertaken by combining the analysis of traditional stress tests with multiple risk factors into a single estimate of the probability distribution of aggregate losses that could materialise under any given stress scenario? Since SRA will not be about a pass or fail test, would systemic risk tolerance[8] levels be established by each jurisdiction and relate to other goals, such as growth path and growth speed?

FSAP macro stress tests, and some ongoing technical work by IMF staff, are looking into these issues, including how best to use them as an effective crisis management and resolution tool. There are also other benefits, stemming largely from the way tests of a systemic nature can discipline thinking about financial stability, use of macroprudential tools, and by providing a measure of resilience that supervisors and regulators can target.

An SRA approach could require the agency (or agencies) responsible for financial stability or system risk to define the state's risk appetite or risk tolerance level. In its

7 As Basel III is rolling out, several important issues remain to be addressed such as identifying the underlying causes of dependence among institutions, the limited time horizon generally used for the analysis and the potential instability of reduced-form parameter estimates because of feedback effects.

8 Risk tolerance could be loosely defined as the aggregate level risk(s) the state or public exchequer is willing to assume within its risk capacity to ensure that the financial system continues to function and intermediate.

absence, accountability of official action either by the central bank or regulator becomes more difficult to communicate. SRA will be its formal articulation of the willingness to use supervisory backstops and make policy interventions to effectively manage certain risks and avoid, or minimise, other risks in the pursuit of macrofinancial stability.

The risk tolerance level could also help link up the broader economic growth strategy with the level of tolerance for financial instability. Decision makers must be accountable for their actions, and financial firms will have to meet their prudential, and other, obligations honestly and in a manner that is consistent with the state's risk tolerance level.

Role of policymaker

Ultimately, systemic risk is a judgement call. Stress tests, while an essential tool of risk management, highlight the need for extensive judgement and challenge from the official sector and the financial firms. It is the role of the policymaker in making use of macro or macroprudnetial stress tests to which I will now turn.

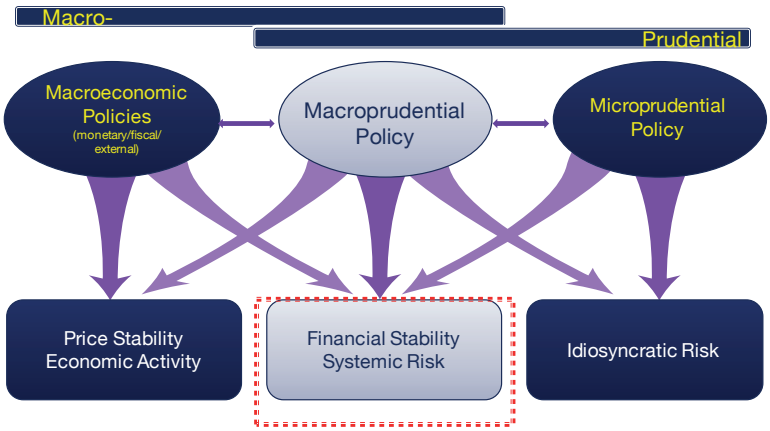
Ensuring consistency with business and capital plans goes in line with holding management accountable for the identification of risk exposures and for the escalation of breaches in risk limits; it is also aligned with the establishment of internal and external safeguards and assessment of the risk framework.

Moreover, it also requires policymakers to describe these risks and explain how they are being managed or mitigated. In addition, it assigns responsibility to the agency in charge of monitoring financial system risks.

Proper use of stress test results means that policymakers are required to extend their responsibilities beyond banks or financial firms and their shareholders and take full account of the interests of depositors, policyholders, and other recognised stakeholders.

Any stress test framework carries inherent limitations, but its benefits depend on the way it is structured and implemented. Stress tests need to be tailored to specific circumstances, including the degree of financial sector development, business models, and the macroeconomic environment in which banks operate and other goals which are also important to societies, besides financial stability (see Figure 3).

Figure 3 Consideration of policy interactions



The success of stress tests cannot be reduced to the choice of a few parameters, but should be set out in a broader context. Certain aspects of stress test design remain crucial. But ultimately, the effectiveness of stress tests does not depend on just a few parameters, but on the context within which they are conducted. This requires a clear ex ante understanding of the stress tests' objectives; knowledge of the key individual financial institutions in the system, their business models, principal sources of risk, and main channels of risk transmission; appropriate decisions on the tests' perimeter and coverage; the use of other complementary assessment tools; a communications strategy tailored to the circumstances and purpose of the tests; and a credible commitment to take the measures that may be required to address vulnerabilities uncovered by the tests.

Governance and integrity of the stress test framework matters a lot. This includes clarity in the objectives, allocation of responsibilities between the official agencies and the financial firms, and data used. A clear follow-up plan, tailored to the specific objective of the exercise, is essential.

In any follow-up, communication issues figure prominently. How much to communicate, in what form, and to whom, are perennial, exceedingly tough questions. The answers will again partly depend on the nature of the exercise and the context, including the broader communication strategy for financial stability policy. The communication strategy and the interaction between supervisors and banks should be designed so as not to undermine the effectiveness of banks' own stress testing strategies.

I think it is desirable that overarching policy measures taken based on stress tests be subject to public debate. It is also fully proper that the agency be asked to explain the background and rationale for the decisions.

However, it is clear that, at this point, there is a lot of variation in how different agencies worldwide envisage the use of stress testing, and this has also varied along the recent history of stress testing – from its use as an instrument to help markets separate good and bad banks, to its present use as a supervisory (US) or macroprudential (UK, Sweden) tool.

Concluding thoughts

First, we expect stress testing frameworks to continue to evolve. For the FSAP, we are constantly improving our approaches to identify key systemic risks to be used in stress testing, to better understand and model the propagation channels of shocks to balance sheets, and to ensure broad coverage of risks and financial institutions in the conduct of stress tests. For risks that are covered, we are also working to ensure size of the shocks are sufficiently large yet plausible, based on past crisis episodes and appropriate calibration in our scenario design models.

Second, the aftermath of the crisis saw stress tests being put to a new use as crisis management tools. In other words, it is no longer an analytical exercise asking 'what if' questions, but a tool to determine the possible need for bank recapitalisation and/or other macroprudential policy actions. While FSAP stress testing does not publish bank-by-bank results, the findings and recommendations focus on systemic risks and system-wide interventions.

Third, one of the key lessons from the crisis was the need to focus on system-wide or macroprudential risks that were not (and arguably could not be) captured by microprudential oversight. The growing literature on macroprudential policy includes an important component on how to identify and monitor systemic risk – an area where the IMF is making important contributions (see the Annex). Macroprudential stress tests are one important component of the toolkit for systemic risk monitoring. The stress testing exercise in FSAPs has also become more mindful of the policy interactions

between microprudential, macroprudential and macroeconomic policies, which could have an impact on the overall stability of the financial system and systemic risks.

As mentioned earlier, macroprudential stress tests are becoming one important component of the toolkit for systemic risk monitoring. It is essential that stress tests are embedded in the policy mix, to inform systemic vulnerabilities and to calibrate macroprudential policy tools.

Finally, going forward, institution-specific and sector-specific stress testing should move towards what described the SRA earlier. The idea is to take a more integrated perspective of risks and interactions, where we combine the traditional stress testing of individual institutions, with a systemic view on the interconnectedness in the financial system, including sovereign risk interactions and the macrofinancial linkages in the economy, as well as cross-border financial linkages.

Annex: Some quantitative methods for systemic risk analysis[9]

1. Contingent claims analysis (CCA) and the new Moody's CreditEdge Database

A basic CCA framework includes risk-adjusted balance sheets and default risk/ credit spreads for financial institutions, corporates, sovereigns, and households. This includes calibration of the CCA models as well as applications including modelling expected default frequencies, uses in stress testing, estimating capital needed to target specific default risk levels, and economy-wide risk transmission analysis.

2. Systemic contingent claims analysis

A systemic CCA framework measures systemic solvency risk based on market-implied expected losses of financial institutions, with practical applications for the financial sector risk management and the system-wide capital assessment in stress testing. The

9 This work is being led by Dale Gray (Items 1 and 2), Jorge Chan-Lau (Item 3), and Miguel Segoviano (Item 4), all IMF staff.

approach uses CCA to generate aggregate estimates of the joint default risk of multiple institutions as a conditional tail expectation using multivariate extreme value theory. In addition, the framework also helps quantify the individual contributions to systemic risk and contingent liabilities of the financial sector during times of stress.

3. Bottom-up default analysis (BuDA) tool for macrofinancial surveillance

BuDA is a simulation-based, bottom-up default analysis tool jointly developed by IMF staff with the Risk Management Institute at the National University of Singapore. Building on a rich and comprehensive probability of default database covering around 60,000 firms in more than 100 countries, it allows mapping the impact of macroeconomic scenarios on the probabilities of default of individual firms and sectors. There are two current model implementations. The first model incorporates a simple VAR-based macroeconomic transmission model and standardised baseline and adverse scenarios building on WEO baseline projections. The second model allows users to input their own macroeconomic scenarios, including the selection of the relevant variables.

4. Systemic risk and interconnectedness (SyRIN) framework: Integrating banks and non-banks

The SyRIN framework spans banks and non-banks, including the insurance, pension, mutual fund and hedge fund sectors. The framework produces complementary systemic risk measures that are useful to quantify: (i) the evolution of systemic risk; (ii) the institutions/sectors that contribute most to systemic risk; (iii) how vulnerable specific institutions/sectors are to distress in other institutions/sectors; and (iv) quantify systemic loss produced in the system by specific institutions, accounting for their relative size and interconnectedness with the system. SyRIN metrics can be used to define authorities' vulnerability tolerance to systemic losses and calibrate capital buffers. The SyRIN framework has two advantages that make it easily implementable. First, it can be developed in a modular manner; hence, risk of specific sectors can be added consistently as data becomes available. Second, it can be implemented with market-based and supervisory based data; therefore, it can be implemented in systems in which market data is not available or is unreliable.

Since the onset of the Global Crisis in 2007-08, stress testing has emerged as a major component of the supervisory toolkit. For most of the large global banks in the US and Europe, meeting the standards to pass their annual supervisory stress tests is the binding regulatory constraint. So the outcome of banks' stress tests is headline news these days. This has been a remarkable development for a tool which ten years ago was little known, apart from amongst a small fraternity of banks' risk modellers and their supervisory counterparts.

This volume collects the contributions of a group of policymakers, academics and specialists in stress test design, from both sides of the Atlantic, who discussed stress testing experiences and methodologies during a two-day conference at the London School of Economics.

It presents the current state of the art of stress testing for bank supervision and macroprudential regulation, and it serves as a useful guide to the frontier issues that the next generation of stress tests needs to address.

Centre for Economic Policy Research
33 Great Sutton Street
London EC1V 0DX
Tel: +44 (0)20 7183 8801
Email: cepr@cepr.org www.cepr.org

ISBN 978-1-907142-98-7

CEPR Press