

INTERNATIONAL MONETARY FUND

Macrofinancial Stress Testing—Principles and Practices¹

Prepared by the Monetary and Capital Markets Department

Approved by José Viñals

August 22, 2012

| Contents | Page |
|--|------|
| Glossary..... | 3 |
| Executive Summary..... | 4 |
| I. Introduction..... | 6 |
| II. Stress Testing: A Primer..... | 8 |
| A. What Is Stress Testing? | 8 |
| B. The Typology of Stress Tests | 11 |
| C. Stress Testing Models..... | 15 |
| III. Lessons from the Global Financial Crisis and the European Sovereign Debt Crisis..... | 18 |
| IV. “Best Practice” Principles and Actual Stress Testing Practices..... | 19 |
| Principle No. 1: Define Appropriately the Institutional Perimeter for the Tests | 20 |
| Principle No. 2: Identify All Relevant Channels of Risk Propagation..... | 25 |
| Principle No. 3: Include all Material Risks and Buffers | 29 |
| Principle No. 4: Make Use of the Investors’ Viewpoint in the Design of Stress Tests..... | 33 |
| Principle No. 5: Focus on Tail Risks..... | 37 |
| Principle No. 6: When Communicating Stress Test Results, Speak Smarter, Not Just Louder | 42 |
| Principle No. 7: Beware of the “Black Swan” | 44 |
| V. Conclusions and Operational Implications..... | 46 |
| References..... | 64 |

¹ This paper was prepared by Hiroko Oura and Liliana Schumacher, with contributions from Jorge Chan-Lau, Dimitri Demekas, Dale Gray, Heiko Hesse, Andreas Jobst, Emanuel Kopp, Sònia Muñoz, Li Lian Ong, Christine Sampic, Christian Schmieder, and Rodolfo Wehrhahn. The paper draws on a survey of stress testing practices at central banks and national supervisory authorities designed by Li Lian Ong, Hiroko Oura, and Liliana Schumacher and summarized by Ryan Scuzzarella.

Appendices

| | |
|---|----|
| I. Existing Supervisory Guidelines for Stress Testing by Banks..... | 50 |
| II. Stress Testing of Nonbank Financial Institutions: Insurance Companies and Financial Market Infrastructures..... | 52 |
| III. Recent Prominent Crisis Management Stress Tests..... | 56 |
| IV. Frontiers of Stress Testing..... | 61 |

Tables

| | |
|--|----|
| 1. Typology of Stress Tests..... | 12 |
| 2. A Comparison of Bottom-Up and Top-Down Stress Tests | 14 |
| 3. Comparing Balance Sheet-Based and Market-Price Based Approaches | 17 |
| 4. Practical Implications of "Best Practice" Principles for Stress Testers | 48 |

Figures

| | |
|--|----|
| 1. Transmission Channels of Shocks Between the Real Economy and the Financial Sector | 26 |
| 2. Institutional Setup of EBA Stress Tests..... | 59 |
| 3. Sovereign and Bank Interactions | 61 |

Boxes

| | |
|--|----|
| 1. Balance Sheet-Based and Market Priced-Based Approaches | 16 |
| 2. The Use of Network Models in Financial Stability Analysis and Stress Testing..... | 22 |
| 3. Integrating Liquidity and Solvency Risks and Bank Reactions in Stress Tests | 28 |
| 4. The Projection of Pre-Impairment Income Under Adverse Scenarios | 31 |
| 5. Market-Based Hurdle Rates..... | 36 |
| 6. The Importance of Risk Interdependence in System-Wide Stress Testing..... | 40 |

GLOSSARY

| | |
|-------|--|
| AFS | Available for sale |
| AIRB | Advanced internal ratings-based approach |
| BCBS | Basel Committee for Banking Supervision |
| BIS | Bank of International Settlements |
| BU | Bottom-up stress tests |
| CAR | Capital Adequacy Ratio |
| CCAR | Comprehensive Capital Assessment Review |
| CCP | Central Counterparty Clearing Houses |
| CDO | Collateralized debt obligation |
| CEBS | Committee of European Banking Supervisors |
| CRD | Capital Requirements Directive of the European Union |
| EBA | European Banking Authority |
| EDF | Expected default frequency |
| EVT | Extreme value theory |
| FMI | Financial market infrastructures |
| FSAP | Financial Sector Assessment Program |
| FSSA | Financial System Stability Assessments |
| GEV | Generalized extreme value theory |
| GFSR | Global Financial Stability Report |
| HTM | Held to maturity |
| ICAAP | Internal Capital Adequacy Assessment Process |
| ICSD | International Central Securities Depositories |
| IRB | Internal ratings-based approach |
| LGD | Loss given default |
| MKMV | Moody's KMV database |
| MTM | Mark to market |
| NPL | Nonperforming loan |
| OFR | Office of Financial Research, U.S. Treasury |
| PD | Probability of default |
| PIT | Point in Time |
| RWA | Risk-weighted asset |
| SCAP | Supervisory Capital Assessment Program |
| SIFI | Systemically important financial institution |
| SPV | Special Purpose Vehicle |
| SREP | Supervisory Review and Evaluation Process |
| TD | Top-down stress tests |
| TTC | Through the Cycle |
| VaR | Value at risk |

EXECUTIVE SUMMARY

The recent financial crisis drew unprecedented attention to the stress testing of financial institutions. On one hand, stress tests were criticized for having missed many of the vulnerabilities that led to the crisis. On the other, after the onset of the crisis, they were given a new role as crisis management tools to guide bank recapitalization and help restore confidence. This spurred an intense debate on the models, underlying assumptions, and uses of stress tests. Current stress testing practices, however, are not based on a systematic and comprehensive set of principles but have emerged from trial-and-error and often reflect constraints in human, technical, and data capabilities.

The Fund is uniquely placed to contribute to this debate. It has amassed vast practical experience with stress tests through more than a decade of Financial Sector Assessment Programs (FSAPs) in its member countries. Fund staff also play a major role in developing advanced stress testing models and cooperating with experts in supervisory agencies and central banks, including through the Expert Forum on Advanced Stress Testing Techniques.

This paper proposes seven “best practice” principles for stress testing. These should not be seen as a “general theory” of stress testing or a complete stress testing manual. They are practical guidelines derived from experience, especially that of the recent crisis. They are meant to provide practitioners with operational guidance on how to tailor stress tests to specific circumstances, including the degree of financial sector development, business models, and the macroeconomic environment in which financial institutions operate. As well, they are meant to empower an informed observer to evaluate critically the stress tests undertaken by country authorities, FSAP teams, or others. Although these principles focus mainly on stress tests conducted for macrofinancial surveillance purposes—the key interest for the Fund, central banks, and macroprudential authorities—much of the discussion also applies to stress tests undertaken for other purposes, such as microprudential oversight and institution-specific risk assessment. The principles are:

- ❖ *Define appropriately the institutional perimeter for the tests.*
- ❖ *Identify all relevant channels of risk propagation.*
- ❖ *Include all material risks and buffers.*
- ❖ *Make use of the investors’ viewpoint in the design of stress tests.*
- ❖ *Focus on tail risks.*
- ❖ *When communicating stress test results, speak smarter, not just louder.*
- ❖ *Beware of the “black swan.”*

These principles underscore that the success of stress tests cannot be reduced to the choice of a few parameters but should be seen in a broader context. To be sure, certain aspects of stress test design are crucial. The choice of risk scenarios, in terms of both the coverage of all relevant risk factors and their severity; the design of the tests so that they cover all important transmission channels and include realistic assumptions about buffers; and the choice of hurdle rates are key for the reliability of stress test results. 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. Lastly, the principles also emphasize that, regardless of technical refinements and improvements, stress test will always remain hypothetical statements and cannot, by themselves, predict or prevent crises. The key operational implications of these principles are summarized in Table 4 on page 48.

Adherence to these principles in practice is uneven. A survey of central banks and supervisory authorities in 23 countries and stress tests in FSAPs shows that despite major improvements since the crisis, practices still fall short of these principles. Shortcomings are particularly notable in three areas: identifying the channels of risk propagation, using the investors' viewpoint, and focusing on tail risks. They reflect both gaps in the analytical toolkit and weaknesses in implementation. Appreciating the implications of these gaps is crucial for the proper interpretation of the stress test results. And needless to say, closing these gaps is the key priority for the community of stress testing practitioners today.

I. INTRODUCTION

1. **The recent financial crisis placed a spotlight on stress testing of financial institutions, notably banks.** The experience highlighted the usefulness of stress tests as a diagnostic tool, but also revealed weaknesses in stress tests undertaken prior to the crisis by the banks themselves, supervisory authorities, and the IMF, all of whom to a greater or lesser extent failed to capture the risks that eventually materialized. At the same time, the crisis underscored the potential of stress tests in restoring market confidence in the financial system, as demonstrated by the successful Supervisory Capital Assessment Program ([SCAP](#)) exercise undertaken by the U.S. authorities in 2009. Stress testing, once an arcane subject, has become almost a household name.

2. **As a result of the attention and lessons learnt from the crisis, the approaches, underlying assumptions, and uses of stress tests are being scrutinized and actively debated.** The large and often confusing menu of choices in each of these areas has given rise to questions about the interpretation of the results and their comparability. And the ongoing financial crisis has made the communication of stress test results an increasingly sensitive issue for both supervisors and financial institutions struggling to balance the call for greater transparency with the need to avoid alarming markets and creating self-fulfilling prophecies.

3. **The Fund is well placed to contribute to this debate, given the experience accumulated for over a decade.** Stress testing of financial systems has been a key component of the Financial Sector Assessment Program (FSAP) launched in 1999 and, more recently, part of the analytical tools used in the Global Financial Stability Reports (GFSRs). As a result, the Fund has amassed a significant practical experience in applying stress testing in a wide range of countries. Fund staff have also played an instrumental role in developing and disseminating several advanced stress testing models and cooperate closely with technical experts in supervisory agencies and central banks in testing and implementing new techniques, including through the Expert Forum on Advanced Stress Testing Techniques. And the Fund is providing technical assistance and training to member countries interested in building or expanding stress testing capabilities.

4. **This paper discusses current practices of stress testing and proposes operational “best practice” principles for their design and implementation.** These practical guidelines are derived from the years of the IMF’s own experience in developing and using stress testing tools, including the constant internal review and evaluation of FSAP stress tests. They are not meant to be a “general theory” of stress testing or provide a comprehensive step-by-step stress testing manual. Instead, the key goal of the paper is to set realistic expectations about the use of stress tests, and explore how the design and implementation of stress tests can be improved to ensure that they remain useful tools in identifying financial sector vulnerabilities.

5. **The paper fills a major gap in this debate.** To staff’s knowledge, this is the first paper that puts forth specific, operational principles for system-wide stress tests. The Bank of International Settlements (BIS) and Committee of European Banking Supervisors (CEBS)—the

predecessor of the European Banking Authority (EBA)—have also proposed principles for stress testing, but those were mainly directed to banks performing stress tests as part of their risk management functions (Appendix I). In a recent paper, [Greenlaw et al. \(2012\)](#) proposed principles for stress tests that focus on risks that could have system-wide and economy-wide implications, but their principles are more conceptual, aimed at shifting the thinking about the purpose and goals of stress tests away from their microprudential focus on individual institutions toward systemic risk. These considerations are also echoed in the first [Annual Report](#) of the Office of Financial Research (OFR) at the U.S. Treasury. Although this paper concentrates mainly on stress tests conducted for macrofinancial surveillance purposes—the key interest for the Fund, central banks, and macroprudential authorities—much of the discussion also applies to stress tests undertaken for other purposes, such as microprudential oversight and institution-specific risk assessment.

6. **In addition, this paper provides the basis for a more systematic approach to stress testing in FSAPs.** The proposed principles establish a yardstick against which individual stress testing exercises can be evaluated, as well as an agenda for improvements in the stress testing toolkit of the Fund. These elements will provide important input into the next review of the FSAP, tentatively scheduled for 2014.

7. **The rest of the paper is organized as follows.** Section II presents a brief introduction to the basic concepts and tools of stress testing. Section III discusses the lessons of the recent global financial crisis and European sovereign debt crisis for stress testers. Section IV presents the seven principles and examines how closely actual stress testing practice corresponds to them; the latter is based not only on the Fund's own extensive experience but also that of its member countries, on the basis of a survey undertaken for this purpose.² And the final section summarizes the key conclusions and practical implications of these principles for stress testing practitioners.

² The survey was conducted in November 2011 and covered (i) the broad use and definition of stress tests; (ii) banking sector stress tests (process and organization of stress tests; framework for solvency tests, including risk/scenario selections, macrofinancial linkages, determining capital adequacy; framework for liquidity stress tests; communication strategy), and (iii) issues regarding the use of stress tests and their application to the nonbank financial sector. A total of 26 central banks and supervisory authorities from 23 different countries responded (in some country cases, the responses were jointly submitted by more than one agency). Among the 23 countries, seven are emerging market economies and 16 are developed economies; 13 are European, 6 are Asian, and 4 are Western Hemisphere countries. A separate background paper provides details of the survey results.

II. STRESS TESTING: A PRIMER

A. What Is Stress Testing?

8. **Stress testing is a technique that measures the vulnerability of a portfolio, an institution, or an entire financial system under different hypothetical events or scenarios.** It is a quantitative “what if” exercise, estimating what would happen to capital, profits, cash flows, etc. of individual financial firms or the system as a whole if certain risks were to materialize.

9. **A complete stress testing exercise is more than just a numerical calculation of the impact of possible shocks.** It involves choices on the coverage of institutions, risks, and scenarios; the application of a quantitative framework to link various shock scenarios to solvency and liquidity measures; a strategy for the communication of the results; and follow-up measures, if warranted. In this paper, the term “stress test” is used to indicate this whole process.

10. **Stress tests typically evaluate two aspects of financial institutions’ performance: solvency and liquidity.** As most stress tests so far focus on the banking sector, this is the main focus in the rest of the paper as well. Emerging stress testing technologies for nonbank sectors, notably insurance and financial market infrastructures, are nevertheless becoming increasingly important, and are reviewed in Appendix II.

Solvency tests

- **An institution is solvent when the value of its assets is larger than its debt, i.e., there is a certain amount of positive equity capital.** The value of both assets and liabilities depends on future cash-flows, which are uncertain and depend on economic and financial conditions going forward. For an institution to be solvent as a going concern, it would need to maintain a minimum of positive equity capital so that it can absorb potential losses in the event of a shock. Higher amounts of capital than this minimum might be needed to ensure continued access to market funding at a reasonable cost.
- **A solvency test assesses whether the firm has sufficient capital to remain solvent in a hypothetically challenging environment by estimating profit, losses and valuation changes.** The main risk factors are potential losses from borrowers’ default (credit risk), and losses from securities due to changes in market prices such as interest rates, exchange rates and equity prices (market risk). A stress test may examine the impact of one source of risks (single factor tests) or multiple sources of risks (multiple factor tests). Risk factors could be combined in an ad hoc manner (combined shock test) or generated more coherently using a macroeconomic framework (macro scenario tests).
- **Solvency tests may cover varying segments of balance sheet.** A test for credit risk may cover total loans (including interbank lending) or loans to certain segments (such as corporate, mortgages, or credit card loans). Market risk is assessed for securities held for trading and in the available-for-sales (AFS) accounts but debt securities in held-to-maturity (HTM) account may be excluded, because these securities are supposed to be

paid in full at maturity as long as the issuer does not default.³ Off-balance sheet exposures, including contingent claims and securitization exposures, could be affected by both market and credit (including securities downgrade risks) risks and could potentially have highly non-linear responses to stress. Bank solvency tests usually do not adjust the value of liabilities for interest rates change (market risk), as the majority of bank liabilities are short-term deposits and money market instruments. On the other hand, stress testing for insurance companies and pension funds should take into account such adjustments, as their liabilities have long maturity and their present value depends on interest rates.

| <i>Simplified Bank Balance Sheet</i> | | | |
|---|---|---|---------------------|
| Assets | | Liabilities | |
| Cash and cash equivalent | | Central bank loans | |
| Money market assets | Interbank lending | Money market liabilities | Interbank borrowing |
| | Repos | | Repos |
| | CDs, CPs | | CDs, CPs |
| Securities | Held for trading | Customer deposits (financial institutions, public sector, corporate, and household) | |
| | Available for sales (AFS) equities and debt | Long-term borrowing | |
| | Held to maturity (HTM) securities | Debt instruments | |
| Loans to customers (financial institutions, public sector, corporate, and household) | | Derivatives | |
| Derivatives | | Other borrowings | |
| Other assets | | Equity Capital | |

| Off-Balance sheet items |
|---|
| <ul style="list-style-type: none"> • Derivatives • Contingent claims (credit lines, guarantees, (implicit) guarantee to special purpose vehicles (SPV)) • Securitization, re-securitization exposures |

- **Estimating solvency ratios in macro scenario stress tests requires estimating macrofinancial models.** A macrofinancial model estimates the empirical relationship between key risk parameters (non-performing loan (NPL) ratio, probability of default (PD),⁴ loss given default (LGD), credit rating, etc.⁵) and relevant macroeconomic

³ Values of securities in trading and AFS accounts are mostly assessed at market values (mark-to-market valuation). Losses and gains from trading securities are accounted for in the profit and loss statement. Unrealized capital gains and losses from AFS securities affect regulatory capital to varying degrees, depending on national regulation and accounting rules. Basel II does not refer specifically to the treatment of unrealized AFS losses, leaving it to national authorities to set their own rules. Under Basel III, all unrealized gains and losses from AFS securities directly affect CoreTier 1 capital. Securities in HTM accounts, on the other hand, are usually valued at book value, unless there are persistent and substantial unrealized losses. The precise valuation practices differ across jurisdictions.

⁴ Point-in-time (PIT) or through-the-cycle (TTC) default probabilities may be used, but the latter tend to dampen portfolio risk (Rosch, 2008).

variables, such as GDP, unemployment, exchange rate, and interest rate. This requires the use of econometric models—as well as considerable judgment—and the Fund provides technical assistance in this area. Macro scenario stress tests cover several years (typically one to three in the case of country supervisory authorities or central banks, and often longer in FSAPs), as credit risks materialize gradually in economic downturns. Therefore, pre-impairment profits (profits before loan and security losses) also need to be projected, since retained earnings in the test horizon would affect capital. This requires making assumptions about bank behavior (such as dividend payout policies and deleveraging, in case of adverse shocks), which introduces significant degrees of freedom—and complexity—to the exercise.

- **Solvency is measured by various capital ratios, typically following regulatory requirements.** Standard choices for banks are the ratio of statutory/core Tier 1/Tier1 capital to risk-weighted assets (RWA); leverage ratios (capital to assets); losses in percent of capital; or capital shortfalls (the amount of capital needed to maintain a certain capital ratio). Individual institutions or the system as a whole are said to “pass” or “fail” the test if the target capital ratio is above a pre-determined threshold or “hurdle rate.” Hurdle rates are often set at the current minimum regulatory requirement, but they could be set at different values if circumstances warrant, for example when new regulations (e.g., Basel III) are expected to be introduced or to maintain a certain level of market funding cost.⁶ The choice of the hurdle rate is a critical factor in stress testing exercises, especially when the results of the tests are directly linked to capital planning directed by supervisors.

Liquidity tests

- **A liquidity stress test examines whether financial institutions have enough cash inflows and liquid assets to withstand cash outflows in a stress scenario.** Financial institutions may encounter sudden cash outflows, for instance, because of:
 - **Sudden distress with their funding.** Financial intermediaries, particularly banks, have, by the nature of their business, a maturity mismatch in their balance sheet. If a large amount of deposits is withdrawn suddenly or funding markets (such as repos and commercial paper) freeze, the bank might face a liquidity shortage even if it is otherwise solvent.
 - **Interlinkages between asset market liquidity and funding liquidity.** Financial institutions active in taking market positions may face a sudden liquidity need when asset markets become volatile and collateral needs and margin calls increase. For

⁵ In a system under Basel I or Basel II standardized approach, only NPL ratio and loan classification are available. In a system under Basel II internal ratings-based (IRB) or advanced internal ratings-based (AIRB) approaches, banks should also maintain PD/LDG or credit rating data.

⁶ The Pillar 2 of Basel II framework also provides supervisors power to request above minimum requirement capital ratios in line with banks’ risk profiles.

example, those trading in highly leveraged derivatives markets may be facing a liquidity shortage if during the life of the transaction it becomes out of the money (even if at expiration this result is reversed) or if the assets are downgraded. Even when these positions are taken by legally separate special purpose vehicles (SPVs), some institutions may be forced to support these SPVs for reputational reasons, effectively internalizing the liquidity shortage.

- **Financial institutions encounter a liquidity shortage when they cannot generate sufficient cash in response to a shock.** If a bank has enough liquid assets, it can generate sufficient cash either by selling the assets or by using repos without making large losses. However, if its assets are mostly non-marketable loans or if the market value of collateral assets declines substantially below the book value (haircut), the bank will be short of liquidity. There are several possible hurdle rates that may be used in a liquidity stress test, such as the number of days the institution can tolerate a liquidity shock before a negative cash flow emerges; net cash flow position; and stressed liquidity ratios.
- **Liquidity and solvency stress events are often closely related and hard to disentangle.** In the event of funding distress, a liquidity shortage may turn into a solvency problem if assets cannot be sold or can be sold only at loss-making prices (“fire sales”). Funding cost increases in a liquidity stress event is a factor that could translate into solvency stress.

B. The Typology of Stress Tests

11. **There are four types of stress tests based on their ultimate objective** (presented in more detail in Table 1):

- **Stress testing as an internal risk management tool.** Financial institutions use stress testing to measure and manage the risks with their investments. One of the early adopters was J.P. Morgan in the mid-1990s, which used value-at-risk (VaR) to measure market risk. However, early stress testing had limited coverage of risk factors and exposures and little integration with the overall risk management and business and capital planning at firms.
- **Microprudential/supervisory stress testing.** The Basel II framework requires banks to conduct stress tests for market risk and, in some cases, credit risk as a part of minimum capital regulation (Pillar 1). Additional tests can be required in the context of Pillar 2 that provides supervisors powers to order management actions by banks if deemed necessary. [BCBS’s survey \(2012\)](#) indicates that supervisory stress tests are increasingly utilized to set capital requirements for specific banks, determine explicit capital buffers, or limit capital distributions by banks. The liquidity ratios in the context of Basel III and insurance regulation in Europe (Solvency II) utilize stress testing as an integral part of the regulatory framework.

Table 1. Typology of Stress Tests

| Features | Macroprudential (Surveillance) | Microprudential (Supervisory) | Crisis management | Internal risk management |
|---|--|--|--|---|
| Main objective | Unveil the sources of systemic risk and vulnerability in the context of surveillance and regular system-wide monitoring. | Assess the health of an individual institution, inform supervision of the institution. | Input for bank recapitalization and business restructuring plans. | Manage risks from existing portfolio, input for business planning. |
| Organized by | Central banks, macroprudential authorities, IMF. | Supervisor (microprudential authority). | Macro and/or microprudential authorities. | Financial institutions. |
| Coverage of institutions | All, or as many as possible institutions, especially systemically important institutions. | Supervised individual institutions (tests for different banks could take place at different times). | Varies, but it should include all distressed and near-distressed institutions. | Individual institution. |
| Frequency | Typically annual or semiannual for country authorities, or in the context of FSAPs. | Individual institutions are tested as needed. Increasing number of supervisors conduct regular stress tests (with common assumptions). | As needed. | High (daily or weekly) for market risks, lower for enterprise-wide exercises. |
| Nature of shocks | Systemic and common shocks across institutions. Shocks tend to be extreme. | Often idiosyncratic; common macro assumptions are sometimes made for horizontal or thematic review across institutions. | Ongoing systemic stress (baseline) or relatively mild shocks, mainly focusing on solvency risks. | Idiosyncratic or systemic (those that matter for the particular institution). |
| Capacity to incorporate systemic risks | Through macro and market-level shocks and additional system-wide features (e.g., network effects). | Through macro and market-level shocks. | Through macro and market-level shocks. | Through macro and market-level shocks. |
| Likelihood of assumed shocks | Low. | Low. | High. | Varies. |
| Assessment criteria (hurdle rates) | Current or prospective regulatory requirements or alternative thresholds, if appropriate. | Current or prospective regulatory requirements or alternative thresholds, if appropriate. | Current or prospective regulatory requirements or alternative thresholds, if appropriate. | Internal risk tolerance indicators and regulatory requirements. |
| Key output metric | Aggregate indicators for the system, and their dispersion. | Individual institution indicators. | Individual institution indicators. | Individual institution indicators. |
| Follow-up measures after tests | Typically no follow up for individual institutions, but often used as the basis for discussion of potential macroprudential or system-wide measures. | Institutions with weak results are often required to explain and take management actions if deemed necessary by supervisors. | “Failing” institutions are often required to take major management action, such as recapitalization, possibly with government support. | May or may not require management action. |
| Publication | Often. | Rarely. | Varies. | No. |
| Examples | FSAP, GFSR, Financial Stability Reports. | CCAR (US), tests required by Basel framework, CEBS/EBA. | SCAP (US), CEBS/EBA (2010, 2011), exercises in some IMF program countries (e.g., Greece, Ireland). | RiskMetrics (J.P. Morgan’s VaR model). |

- **Macroprudential/surveillance stress testing.** Over the past two decades, many country authorities have started using stress test exercises to analyze system-wide risks, in addition to institution-specific risks. The results are often reported in their Financial Stability Reports. The IMF also regularly includes stress testing in FSAPs since the program's inception in 1999. A few country authorities indicated in the survey that the FSAP stress tests were the first such exercises undertaken in their countries.
- **Crisis management stress testing.** Stress tests have also been used, especially after the recent crisis, to assess whether key financial institutions need to be recapitalized or not, possibly with public support. In particular, the U.S. [SCAP](#) exercise and the exercises organized by the CEBS/EBA in [2010](#) and [2011](#) attracted attention because banks were required to recapitalize based on the test results, and the detailed methodology and individual banks' results were published. In recent IMF programs with banking sector distress (including Ireland, Greece, and Portugal), estimating bank recapitalization needs through stress tests was an important component. As this use of stress tests as a crisis management tool is relatively new, Appendix III presents the key features, similarities with, and differences from other types of stress tests, using three recent prominent examples of this type of stress tests (the U.S. SCAP, the CEBS/EBA tests in 2010 and 2011, and the EBA Capital Assessment exercise of 2012) as illustrations.

12. **The risk coverage and methodologies have evolved over time, as the use of stress tests was broadened.** Financial institutions are now expected to manage enterprise-wide risks, which cover broad ranges of exposures and risk factors in an integrated manner, crossing over the internal segmentations of various business lines. Similarly, macroprudential stress tests evolved from single factor tests to macro scenario tests.

13. **Depending on the objective of stress testing, follow up managerial or supervisory actions may be taken.** Macroprudential tests, including in FSAPs, typically do not prescribe bank-specific action, although they could lead to macroprudential policy recommendations. Supervisory stress tests are increasingly used to guide supervisory action, ranging from improving data collection, targeting examinations, and closer monitoring, to requiring bank management actions, such as raising additional capital, reducing certain exposures, capping dividends, and updating individual institutions' resolution plans. Follow-up is almost certain in crisis management stress tests, which are expressly designed to estimate capital shortfalls.

14. **Stress testing in FSAPs and by the national authorities may be conducted either as top-down or bottom-up exercises or both.** Top-down (TD) exercises are defined as those conducted by the national authorities or Fund staff (typically in FSAPs) using bank-by-bank data and applying a consistent methodology and assumptions. Bottom-up (BU) exercises are carried out by individual financial institutions using their own internal data and models, often under

common assumptions. Some supervisory tests include bank-specific risks⁷ (Table 2), including reverse stress tests based on shocks that could render a specific institution insolvent. Liquidity tests are often conducted as bottom-up exercises because they require granular data and depend on banks' liquidity strategies, and banks are given more flexibility regarding detailed assumptions compared to solvency tests. FSAPs almost always include a TD test, frequently supplemented by a BU test. Many national authorities use both TD and BU and emphasize the importance of running their own TD tests in order to effectively validate BU results.

Table 2. A Comparison of Bottom-Up and Top-Down Stress Tests

| Feature | Bottom-Up (BU) | Top-Down (TD) |
|------------|--|--|
| Strengths | <ul style="list-style-type: none"> • Reflects granular data and covers exposures and risk-mitigating tools more comprehensively, including those that are hard to cover in TD tests (such as risks from complex structured products, hedging strategies, and counterparty risks). • Utilizes advanced internal models of financial institutions, which could potentially yield better results. • May reveal risks that could otherwise be missed. • Provides insights in the risk management capacity and culture of a particular institution. • Application of severe common shocks may encourage individual institutions to prepare for tail events than they might otherwise not be prepared to contemplate. | <ul style="list-style-type: none"> • Ensures uniformity in methodology and consistency of assumptions across institutions. • Ensures full understanding of the details and limitations of the model used. • Provides an effective tool for the supervisory authority or FSAP team to validate BU tests. • Once a core framework is in place, implementation is relatively resource-effective. • Can be implemented in systems where institutions have limited risk management capacities. |
| Weaknesses | <ul style="list-style-type: none"> • Implementation tends to be resource intensive and depends on the cooperation of individual institutions. • Results may be influenced by institution-specific assumptions, data, and models that hamper comparisons across institutions. | <ul style="list-style-type: none"> • Estimates might not be precise due to data limitations. • Standardization may come at the cost of not reflecting each institution's strategic and managerial decisions. |

15. **Communication practices differ across the four types of stress testing.** Close communication between banks and supervisors, between supervisory agencies within a country, or between FSAP teams and country authorities is required for effective stress tests. Public communication of stress test results, on the other hand, is not common, although recently this has been changing, especially for crisis management stress tests.⁸ Macroprudential/surveillance stress test results are typically reported in financial stability reports—or Financial System Stability

⁷ There is some variation in BU practices among country authorities, including BU tests being conducted with institution-specific assumptions, or implemented by the supervisory authority using bank-by-bank data.

⁸ In the U.S., the Dodd-Frank Act requires communication of [Comprehensive Capital Analysis and Review \(CCAR\)](#) by individual banks and by the Federal Reserve Board.

Assessments (FSSAs), in the case of FSAPs—but in varying degrees of aggregation, usually without identifying individual institutions. In general, dissemination of stress test results is controversial. Several country authorities have voiced concerns that public dissemination might create unrealistic expectations, lead to misinterpretations in the mass media, and potentially detract from the value of the stress tests as a supervisory tool as banks focus too much on the media impact. For FSAPs, the publication of FSSAs is voluntary, but the majority of countries—including most of the jurisdictions with systemically important financial sectors—publish the document. The accompanying stress testing Technical Notes are published much less frequently.

C. Stress Testing Models

16. **Stress tests use a wide variety of analytical models that fall broadly into two categories.** There is a large and often bewildering choice of models relating stress factors to solvency or liquidity of individual institutions or financial systems, varying widely in terms of complexity and data requirements. At the risk of some oversimplification, they can be classified into two broad families: models predicated on a detailed analysis of balance sheets of individual institutions (sometimes called “fundamental” approaches); and models based on summary default measures for individual portfolios, institutions, or entire systems embedded in market prices, such as stocks, bonds, and derivatives.

17. **Both approaches have strengths and weaknesses and should be seen as complements rather than substitutes.** Balance sheet-based approaches can identify the source of individual vulnerabilities in the balance sheet. They are thus more informative, and can be applied to emerging and low income countries where stock markets are thin or illiquid. But they are backward-looking, data intensive, hard to update frequently, and not well-suited for capturing interdependence (portfolio) and contagion effects across institutions. Market price-based approaches, on the other hand, are more flexible, can easily incorporate portfolio effects and risk factors as perceived by the market, and can be updated as frequently as desired. But they make it difficult to disentangle the precise source of vulnerabilities, are sensitive to short-term swings in market perceptions that may have little to do with fundamentals, and cannot be applied to countries or entities with limited or no market price data. Box 1 presents an overview of the two families of models, and Table 3 a detailed comparison of their operational aspects.

18. **Model choices are somewhat constrained in emerging or developing economies with limited data or weak accounting and regulatory system, but some options are nonetheless available.** At the same time, these countries tend to have simpler financial and economic systems, and key economic risks and vulnerabilities are relatively straightforward to identify. Simple balance sheet-based tests (as those illustrated in [Čihák, 2007](#)) with single or multi-factor shocks can be implemented in most countries using basic supervisory data. In cases where supervisory data are patchy or unreliable, or the magnitude of uncertainty too large to draw strong conclusions from the tests, the priority should be to improve supervision and data availability rather than develop stress testing approaches. Developing a macro-financial linkage model for macro scenario tests, on the other hand, can be a challenge in some emerging and

developing countries, given limitations in the quality or availability of long time series data. One option in such cases is to utilize cross-country data.⁹

Box 1. Balance Sheet-Based and Market Price-Based Approaches*

There are two broad types of approaches to estimate the extent of bank distress: balance-sheet based and market price-based approaches. Balance sheet-based models spell out all (on- and off-balance sheet) positions in a bank portfolio and the risks to which these are exposed. Market price-based models are based on summary bank default measures embedded in asset prices (such as bank stocks, bonds, and derivatives). These measures are extracted from market prices by solving for the default probability implicit in them, using standard pricing formulas. The IMF often uses the Contingent Claims Approach (CCA), a market price-based approach built on standard option pricing model (Gray, 2008). When used for stress tests, market price-based methodologies need to project the market-based default measures for the period covered by the test.

Both approaches have advantages and disadvantages. Balance sheet-based models are more informative but have been blamed for being backward-looking, data intensive, and hard to update given typical lags in the release of the relevant information. Market price-based measures are forward-looking, can be easily updated, and in many cases already incorporate portfolio effects (for example, the parametric approach of Credit Risk+ (Avesani et al., 2008), and the non-parametric CIMDO copula function (Segoviano, 2006)) that need to be explicitly modeled in balance sheet-based models. However, they also reflect the impact of market swings that may be unrelated to fundamentals.

While data limitations may dictate the use of one or the other type of model in a specific set of circumstances, these two methodologies do not convey the same type of information and they should be considered complementary rather than substitutes. Balance sheet-based models are typically used by financial stability/supervisory authorities since they need disaggregated information on the sources of vulnerabilities in order to adopt risk-mitigating measures. Market price-based approaches can be used when the interest lies on understanding the market assessment of bank solvency under stress.

Because of their detailed nature, balance sheet-based models face severe limitations in capturing all risks in an integrated manner. For this reason, in practice, some forms of risk have received overwhelming attention (such as default risk of private counterparties in the banking book) at the expense of others (such as sovereign default risk, downgrading risk of sovereign and private counterparties, counterparty risks of derivatives, and liquidity risks, including funding costs). Capturing the intrinsic dependencies across different types of risk is a major challenge for balance sheet-based models. For example, counterparty risk is inherently dependent on the evolution of market risk; and the global crisis has shown that systemic liquidity risk cannot be assessed without considering the solvency profile of institutions. Methodologies aimed at the valuation of all positions held by banks and the assessment of all risks in an integrated manner are being developed (see, for instance, Barnhill et al., 2002, and Barnhill and Schumacher, 2011) but are data-intensive and yet to be refined.

A common challenge for both types of models is finding a way to stress test individual institutions to system-wide risks. This requires assessing joint entity default probabilities and measures of default dependence in order to produce estimates of systemic loss distribution. In the context of the balance sheet-based models, this extension can be undertaken by superimposing a network of claims to keep track of default effects of one institution onto the others. Market price-based models, on the other hand, typically treat the banking system as a portfolio of banks and derive a distribution of systemic losses using portfolio analysis techniques similar to those used for individual bank portfolios.

* Prepared by Liliana Schumacher.

⁹ For instance, [Annex 1.6 in the April 2010 GFSR](#) shows that an NPL projection model using data from Latin America and emerging Asia performs reasonably well in predicting loan quality in Central and Eastern European countries, where time series are relatively short.

Table 3. Comparing Balance Sheet-Based and Market-Price Based Approaches

| | Balance Sheet-Based | Market Price-Based |
|--------------------------------|---|---|
| Primary input data | <ul style="list-style-type: none"> Accounting data (balance sheet, profit and loss account, matrix of interbank exposures). | <ul style="list-style-type: none"> Financial market data (equity prices, bond yields, CDS spreads or equity option-based probability of distress of banks). |
| Secondary input data | <ul style="list-style-type: none"> Probability of distress (and loss-given-default) or NPL ratios/loan classification of borrowers (for credit risk); Market data (equity prices, exchange rates, interest rates, price volatility, term premium) to calibrate shocks. | <ul style="list-style-type: none"> Balance sheet data (combination of equity prices and accounting data obtain key input variable, such as the Expected Default Frequency by Moody's). |
| Type of test | <ul style="list-style-type: none"> Solvency, liquidity and network analyses. | <ul style="list-style-type: none"> To date, largely focused on solvency and its interdependence among key financial institutions. Initial attempts at testing for liquidity stress. |
| Frequency | <ul style="list-style-type: none"> Varies depending on the reporting cycle (quarterly, semi-annual, annual). | <ul style="list-style-type: none"> Daily or lower frequency. |
| Application | <ul style="list-style-type: none"> Most banks or financial systems (including emerging markets and low income countries) as long as financial reporting or supervisory data exists and is available. | <ul style="list-style-type: none"> Limited to market data-rich countries and institutions that are quoted on the market (it generally cannot cover mutuals, privately-held, or government-owned companies). Stand-alone analysis for subsidiaries may be difficult. |
| Link(s) to macro scenarios | <ul style="list-style-type: none"> Possible, by estimating additional macro-financial model(s), linking macro scenario variables and risk factors (PDs of borrowers, NPL ratios, etc.). | <ul style="list-style-type: none"> Possible, by estimating additional macro-financial model(s), linking macro scenario variables and risk factors (PDs of banks, volatility or leverage of banks, etc.). |
| Estimation of systemic effects | <ul style="list-style-type: none"> By considering common macro shocks across banks (e.g. GDP, inflation); and incorporating network effects (interbank exposures). | <ul style="list-style-type: none"> By considering common macro shocks across banks and incorporating interdependence (portfolio) effects among banks, which may be estimated using asset prices. |
| Output | <ul style="list-style-type: none"> Various capital ratios. Liquidity ratios. Capital shortfalls. The number/share of banks breaching minimum requirements. | <ul style="list-style-type: none"> Expected losses. Unexpected or tail losses. Contingent liabilities for the government. Probabilities of spillover among banks. |
| Strength | <ul style="list-style-type: none"> Pinpoints the type of risk that creates the vulnerability (e.g., credit losses from housing loans, market valuation losses from exposures to sovereigns, losses from currency mismatches). Possible to adjust for supervisory weakness (e.g., under-provisioning, forbearance etc.). | <ul style="list-style-type: none"> Less data-intensive than the accounting-based approach. Focuses on systemic risks/losses and tail events. Incorporates risk factors priced by the market. |
| Weakness | <ul style="list-style-type: none"> Data intensive (especially for network analysis). Quality of the analysis depends on the granularity and availability of the data. | <ul style="list-style-type: none"> The causes of different risks are difficult to disentangle ("black box"). Estimated vulnerability measures may be very volatile during periods when markets are under significant stress, and links with balance sheet fundamentals may be obscured. |
| Selected Examples | <ul style="list-style-type: none"> Stress tester 101 (Cihak, 2007). Next generation stress testing (Schmieder, Puh and Hasan, 2011). Network analysis (Espinoza-Vega and Sole, 2010). CreditRisk+ (Credit Suisse Financial Products). | <ul style="list-style-type: none"> Systemic CCA (Gray and Jobst, 2010). CoVaR (Adrian and Brunnermeier, 2008). Distress dependence (Segoviano and Goodhart, 2009). SES and MES (Acharya et al. 2010). |

19. **This is a very dynamic area, with model refinements and new methodological approaches constantly being developed.** As the discussion in some of the following sections illustrates, Fund staff are very active in this area, in close cooperation with advanced economy central banks and financial stability agencies and the academia. The Expert Forum on Advanced Stress Testing Techniques, organized by the IMF’s Monetary and Capital Markets Department (MCM) and a different cooperating central bank on an annual basis, has become one of the pre-eminent fora for exploring some of these new approaches among stress testing practitioners.

III. LESSONS FROM THE GLOBAL FINANCIAL CRISIS AND THE EUROPEAN SOVEREIGN DEBT CRISIS

20. **The global financial crisis and recent European sovereign debt crisis had a major impact on the way stress tests are conducted, as well as on how their results are used.** *First*, the experience highlighted some weaknesses of the pre-existing stress testing approaches. There is broad agreement that the majority of the stress testing exercises before the crisis by the financial industry, country authorities, as well as FSAPs, failed to detect key vulnerabilities (Haldane 2009a and 2009b). *Second*, the aftermath of the crises spurred the use of stress tests for crisis management purposes, notably in the U.S. and Europe, albeit with mixed results (Appendix III). *Third*, as mentioned earlier, the use of stress tests for crisis management engendered a move toward greater transparency in disseminating stress test results which, nonetheless, remains controversial.

21. **Why did pre-crisis stress tests fail to detect the vulnerabilities that eventually materialized?**

- **The institutional perimeter of stress tests was too narrow.** Stress tests did not, as a rule, cover what came to be known as “shadow banking” (e.g., money market funds and insurance companies writing credit insurance) that played a key role in originating or transmitting shocks.
- **Key shock transmission and propagation channels were not covered.** Interconnectedness among key financial institutions through cross-exposures propagated and amplified shocks (as in the case of Lehman Brothers). Second-round feedback effects between the financial sector and the real economy and/or sovereign risk were not incorporated in macrofinancial stress testing.
- **Some key risk factors were missed.** Shocks hit multiple markets and countries at the same time due to common stress or contagion from one to another, generating a systemic shock. A number of specific risk factors were also missed, including counterparty risk, basis risk, and contingent risks ([BCBS, 2009](#)).¹⁰

¹⁰ Counterparty risk is a type of credit risk (e.g., risk of default of a counterparty to a derivative transaction). Basis risk for hedging is the risk that hedging becomes imperfect because of the difference between the asset whose price is

- **Balance sheet valuations did not fully reflect economic value.** Stress tests based on regulatory and accounting norms overestimated the resilience of the financial system. Market pressures on sovereign debt, for instance, are often not fully reflected in regulatory capital, which does not value every security at market values. And hurdle rates in solvency tests reflecting regulatory minima proved insufficient when markets demanded larger capital from specific banks, resulting in some banks “passing” the stress test but facing severe distress shortly thereafter.
- **The shocks were not severe enough and did not examine truly tail events.** In some cases, the stress scenarios turned out to be too benign compared to the actual shocks ex post. In other cases (e.g., CDOs and other structured products), the data series were simply not long enough to provide the basis for constructing a stress scenario. Time-varying correlation and extreme market risks, which were the main contributors to tail risks, were not well reflected in stress tests (Rosch and Scheule, 2008).
- **Many shocks were not included in the tests, as they were considered “unthinkable.”** For instance, severe liquidity risks, involving the complete seizure of key funding markets and their linkages with solvency, were not included in stress testing exercises. The risk of sovereign default risk in advanced economies was similarly not covered. These scenarios were considered too extreme to be plausible.

22. **Efforts are being made to improve stress testing frameworks in light of this experience, but this is still an unfinished agenda.** Current practices have incorporated some lessons learned from the crisis, notably in terms of the types of scenarios and the severity of shocks, partly because shocks that materialized during the global financial crisis provide a good benchmark for tail risks. However, challenges remain with incorporating all relevant risk propagation channels, as well as feedback effects between the financial sector and the real economy. Appendix IV presents some of the emerging methodological approaches at the frontier of stress testing techniques that attempt to tackle these challenges.

IV. “BEST PRACTICE” PRINCIPLES AND ACTUAL STRESS TESTING PRACTICES

23. **Stress test design and implementation should ideally be based on “best practice” principles that incorporate the lessons from the recent crisis and are sufficiently operational.** Such principles should provide practitioners operational guidance on how to tailor stress testing to a wide variety of country and sector circumstances while maintaining minimum standards that can enhance comparability across different exercises, which is particularly important for FSAPs; protect against their most obvious pitfalls, especially in light of the lessons from the recent crisis; and interpret their results appropriately. They should also allow an

to be hedged and the asset underlying the derivative or because of a mismatch between expiration date of the futures and the actual selling date of the asset. Contingent risks could arise either from legally binding credit and liquidity lines or from reputational concerns related to, for example, off-balance sheet vehicles.

informed observer to evaluate critically the various stress testing exercises undertaken by country or supranational authorities or IMF FSAP teams. And they should take into account the fact that stress testing is ultimately as much of an art as a science.

24. **In reality, stress testing practices have thus far been unsystematic.** Practices in various institutions, including the Fund, often reflect existing constraints in human, technical, and data capabilities on the institutions undertaking stress testing. They have emerged out of trial-and-error or, in some cases, even as a matter of habit or convenience. They do not always reflect a systematic effort to build an approach to stress testing based on first principles.

25. **This paper proposes seven “best practice” principles for stress testing.** These principles are mainly focused on stress tests for macroprudential surveillance, although they are to a greater or lesser extent applicable to all types of stress tests. As such, they are related to, but do not overlap with the Principles for Sound Stress Testing Practices ([BCBS, 2009](#)), which focus on banks’ own stress testing practices (Appendix I). The remaining sections discuss their implications and evaluate to what extent actual practices correspond to these, based on the survey results.

- ❖ Define appropriately the institutional perimeter for the tests.
- ❖ Identify all relevant channels of risk propagation.
- ❖ Include all material risks and buffers.
- ❖ Make use of the investors’ viewpoint in the design of stress tests.
- ❖ Focus on tail risks.
- ❖ When communicating stress test results, speak smarter, not just louder.
- ❖ Beware of the “black swan.”

Principle No. 1: Define Appropriately the Institutional Perimeter for the Tests

26. **This principle targets the selection of the institutions to be included in the tests.** For system-wide tests, this involves a choice of which institutions to include and which to leave out. This choice requires an assessment of which banks are systemically important (i.e., capable of triggering or amplifying systemic risk).

27. **Size, substitutability, complexity, and interconnectedness are the criteria that are used to assess systemic importance.**¹¹ A bank’s distress or failure is more likely to cause damage to other banks, markets, or the economy if its activities comprise a large share of financial intermediation. The higher the network of contractual obligations in which an institution operates, the higher the likelihood that its failure will materially raise the likelihood of distress at other

¹¹ See IMF/BIS/FSB (2009), [Report on Guidance to Assess the Systemic Importance of Financial Institutions, Markets, and Instruments: initial considerations](#); and BCBS (2011) [Global systemically important banks: Assessment methodology and the additional loss absorbency requirement](#).

institutions. The systemic impact of a bank's distress or failure is expected to be negatively related to its degree of substitutability as a market participant or service provider (in the case where specific institutions provide critical services, such as a market infrastructure). The systemic impact of a bank's distress or failure is higher for more complex institutions, as the costs and time needed to resolve these are greater.

28. While size, degree of substitutability, and complexity are observable features, an assessment of interconnectedness requires the use of sophisticated network approaches.

Network approaches can be useful in identifying systemic institutions that should be covered by stress tests (Box 2). Financial institutions hold claims against each other, forming a network, which can be thought of as a matrix of bilateral claims. Recent network models provide rich dimensions of interconnectedness and can identify systemically important institutions or groups of financial institutions that are at the center of the network, going beyond simple metrics of cross-exposures. In addition, network models can also be used directly to measure the likelihood of multiple defaults by systemically important institutions, a key feature of systemic risk (stress testing individual institutions' solvency and simply aggregating the outcome would tend to underestimate systemic risk—see Principle 5).

29. Separately, stress testing requires identifying the appropriate perimeter of the activities of financial conglomerates that would be covered by the tests. Some of the largest financial institutions have a wide range of activities, often including cross-border activities or cross-industry activities that could cover banking, insurance, pension fund, investment fund, various financial SPVs, and even non-financial activities. The ultimate ownership structure and economic links may not be always clear. And although some of the SPVs may be legally independent, the crisis illustrated that they could still imply contingent liabilities for the “parent” companies if the latter choose to provide support beyond legal obligations for reputational reasons. Analysis of global banks in their home countries often tend to focus on their activities in that country, but cross-border exposures may be relevant in assessing inward and outward spillover effects. Analysis of global banks in host countries sometimes put potentially the most important risk—the health of the parent company—outside of the scope of the stress tests due to limited information on group-wide activities available to host country supervisors. The stress tester needs to use judgment as to whether these activities should be consolidated or segregated.

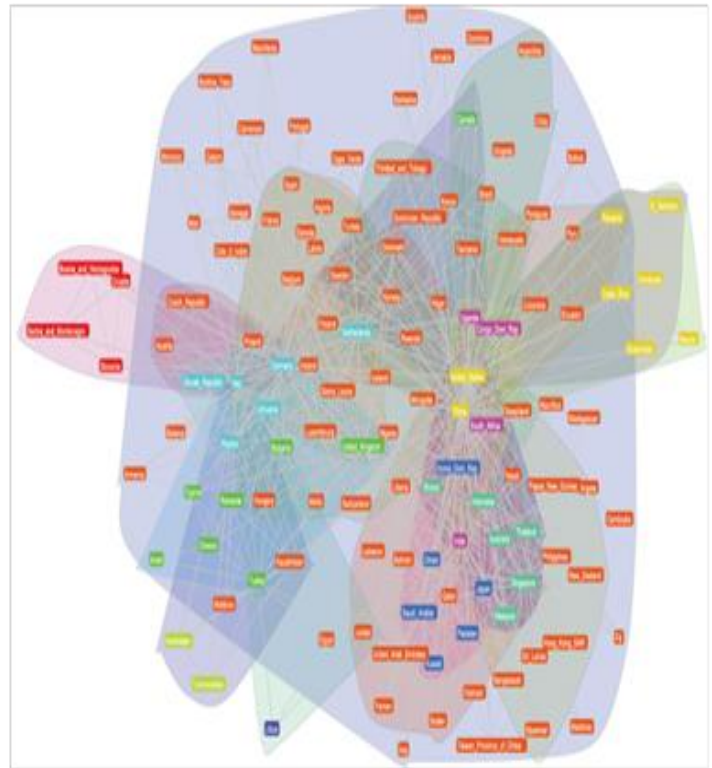
Box 2. The Use of Network Models in Financial Stability Analysis and Stress Testing*

Network models provide a tool to measure interconnectedness, i.e., linkages among financial institutions, systems, or entire countries, through claims held against each other or other channels. The importance of interconnectedness as a channel of contagion has been studied for some time (e.g., Allen and Gale, 2000), among others, explored how linkages among banks through direct exposures could be a source of contagion) and was spectacularly underscored by the failure of Lehman Brothers. Stress testing for FMIs also makes extensive use of network analysis. Nevertheless, it is important to keep in mind that the relationship between interconnectedness and financial stability is not simple and monotonic: interconnectedness may enhance or reduce financial stability, depending on the degree of cross-institution or cross-border integration and the precise pattern of cross-exposures or other linkages ([Čihák, Muñoz and Scuzzarella, 2011](#)).

Simple network models, measuring interconnectedness through cross exposures, have been used to add contagion channels to stress testing. The IMF’s introductory stress testing kit by [Čihák \(2007\)](#) includes a simple feature analyzing how a failure of a bank may affect other banks directly if it defaults on its borrowings. [Espinosa and Solé \(2010\)](#) and [Tressel \(2010\)](#) incorporate an additional channel: the failure of a bank may affect other banks indirectly because it stops it from lending to them, thus eliminating a source of liquidity in the system. Such contagion analysis has been part of [IMF/FSB’s Early Warning Exercises](#) (cross-border contagion using BIS cross-border bank exposures data) and stress testing in FSAPs. Network effects can be incorporated in stress tests in an ad hoc manner (selecting randomly the bank(s) that fail, or “trigger” banks) or integrated with the macro stress testing, where the institution(s) that fail the solvency or liquidity tests under a stress scenario become the trigger banks.

A recent crop of more advanced network models can analyze a richer dimension of the network structure. These models examine four measures of interconnectedness, depending on the structure of cross-exposure links, to identify key nodes (institutions, financial systems, or entire countries) in the network: (1) “in-degree” is the number of links that point to a particular node; (2) “closeness” is the inverse of the average distance from one node to others; (3) “betweenness” focuses on the shortest path between nodes; and (4) “prestige” assigns increasing scores to nodes that are connected to other high-scoring nodes. These models can, among other things, describe the disproportionately important role of financial centers in transmitting shocks around the world compared to minor financial systems; or identify banks that may be small but could play a critical role in connecting financial centers. Another technique known as “cluster analysis” separates the network into subgroups (“clusters”) of nodes that have closer connections to each other than with those outside of the cluster. It can help identify subgroups of nodes with close connections and “gatekeeper” institutions or systems that bridge across different clusters. This technique was used by the Fund to identify the 25 jurisdictions with the most systemically important financial sectors that are at the center of the global financial network ([IMF, 2010a](#)). A similar technique was applied by the Reserve Bank of India to identify core institutions, including banks and nonbanks.

Cluster Analysis and Global Trade and Financial Architecture

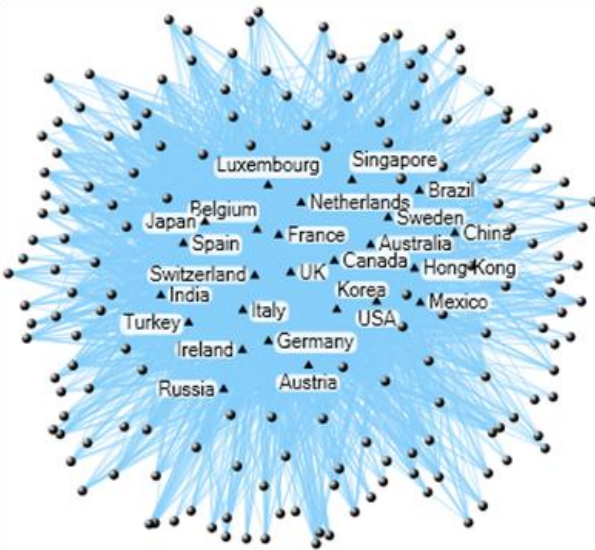


Source: International Monetary Fund (forthcoming), *Enhancing Surveillance: Interconnectedness and Clusters*. Each entry is a country. Links between countries represent trade and financial linkages. Cluster technique is applied to identify subgroups (clusters) of countries with particularly strong links to each other (colored shapes including several countries).

* Prepared by Sónia Muñoz and Hiroko Oura.

Box 2 (contd.). The Use of Network Models in Financial Stability Analysis and Stress Testing

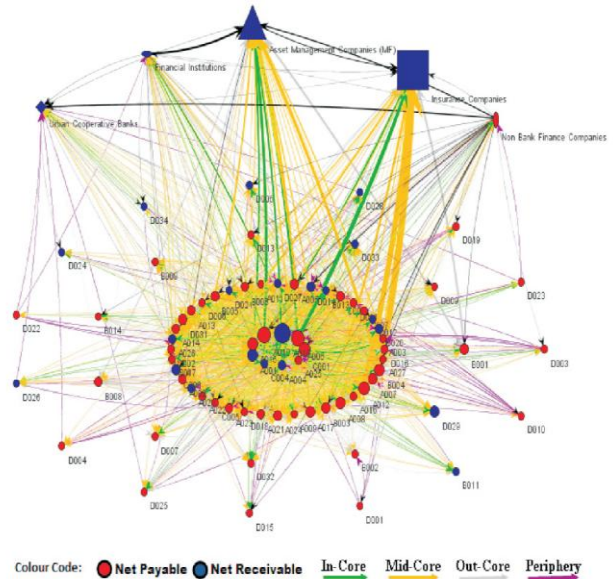
The Position of the 25 Jurisdictions with Systemically Important Financial Sectors in the Global Network



Source: [International Monetary Fund, 2010a](#).

Each round dot is a country and blue lines represent cross-exposures between countries. Triangles are the 25 most systemically important financial centers identified by applying cluster analysis to various measures of interconnectedness.

India: Network Structure of the Financial System



Source: [Reserve Bank of India Financial Stability Report, 2011](#).

Round dots are banks (red has net payable and blue has net receivable) and other shapes represent cooperative banks, asset management companies, insurance companies, and finance companies. The network has four distinct tiers, including the most connected banks (inner core circle of dots), mid core, outer core, and periphery.

These network models can help both model contagion directly and provide input to stress tests. For instance, the Reserve Bank of India's network model, based on the tiered network with a highly clustered central core, showed that a failure of a bank with large exposures to the insurance and mutual funds segments of the financial system could cause distress to ten other institutions, including three insurance companies. Moreover, by identifying systemically important institutions at the core of the financial system, this approach can help set the right perimeter for stress testing.

30. **Finally, defining the proper perimeter also calls for the inclusion of nonbank institutions that may trigger or propagate systemic risk, such as FMI and insurance companies.** The concern about FMIs, such as payments systems, central securities depositories, securities settlement systems, and central counterparties, is not their balance sheet per se but rather their safe and reliable functioning. Their systemic importance is given by the key role played by the services they provide.¹² Stress tests of insurance companies are also becoming more common. Although traditional insurance activities are not very likely to trigger systemic risk, insurance companies may engage into complex transactions and create systemic risk through their links with banks. In this regard, a network approach integrating all financial institutions would be less likely to leave out sources of potential systemic risk.

¹² For this reason, the CPSS-IOSCO standard requires an assessment of FMIs using stress testing.

31. **Applying this principle requires a good understanding of the system’s main features before undertaking the stress tests.** This includes who are the relevant market participants, how they operate, what their business models are, what type of transactions they perform and with whom, which are the areas of risk concentration and the likely channels of risk transmission. A formal mapping of this understanding into a network of claims and potential claims would facilitate the job, but other, more heuristic tools may also be used to get an understanding of the system, including market intelligence and conversations with market participants.

32. **How well do actual stress testing practices correspond to this principle?**

- Bank stress tests tend to be comprehensive, either covering all institutions in the system or focusing on systemically important institutions. FSAP stress tests typically cover at least seventy percent of the locally-incorporated commercial banks (including subsidiaries of foreign banks and state-owned banks) by assets. Country authorities’ exercises focus on private domestically-owned commercial banks, followed by foreign subsidiaries and state-owned banks. Based on the survey results, the coverage ranges from 60 to 100 percent of the system by assets (with the median being 85 percent and 16 banks). The number of banks in the sample varies from below 5 to over 1,000.
- In cases where only a subset of the system is included in stress tests, the methodology to establish the perimeter of the tests varies. The size of the balance sheet and interconnectedness are the key factors, followed by the size of local retail activities and legislative definitions. Formal network approaches are increasingly used to assess interconnectedness, although the more sophisticated models are still used in a minority of cases, even in advanced economies. Various indicators of interconnectedness, in addition to size, played critical roles for determining the 25 jurisdictions with systemically important financial sectors that are required to undergo mandatory stability assessments under the FSAP ([IMF, 2010a](#)). Network models were part of the [Chile](#) and [Finland](#) FSAP exercises in identifying important linkages across financial institutions.
- When it comes to including nonbank institutions in stress tests, practices are much more uneven. FSAPs always test the banking sector, and occasionally the insurance sector; other sectors are rarely tested. In cases where insurance is covered, efforts are made to align the economic shocks to those used for the banking stress tests. But since each segment of the financial system may react to a certain macroeconomic scenario differently,¹³ the main scenario needs to be complemented with sector-specific scenarios. Moreover, insurers can also be vulnerable to other types of events (for example, catastrophic risks, such as floods, earthquakes, and other natural disasters, are important for nonlife insurers; and epidemics or long-term changes in mortality are important for life insurers) but these are rarely covered in FSAPs, which tend to focus on economic risks.

¹³ For instance, banks and insurers often react to interest rate shocks differently as banks typically have positive duration gaps, losing from rising rates, but insurers have negative duration gaps, gaining from higher rates.

Among country authorities, about 40 percent of the respondents indicate they only test the banking sector and another 45 percent also test the insurance sector. Tests of pension funds and FMIs are undertaken on a much more ad hoc basis, if at all. The U.S. [SCAP](#) and [CCAR](#) include a life insurance company (Metlife) and the former auto loan arm of General Motors (Ally), in addition to investment and commercial banks, though these companies currently operate with bank holding company licenses.¹⁴ As the operator and overseer of the main payments system, some central banks use simulation tools to assess the impact of operational disruptions of the financial market infrastructure (FMI) itself or a major participant. Incidents are simulated in order to identify recovery times, critical participants, and contingency measures. Stress tests of a central counterparty clearing houses (CCP) take into account extreme but plausible market conditions, and are typically framed in terms of the number of participant defaults that a CCP can withstand.

Principle No. 2: Identify All Relevant Channels of Risk Propagation

33. In addition to network effects among financial intermediaries, there are other channels of shock propagation that relate financial intermediaries to each other and to other agents in the economy. Key examples of these propagation channels, illustrated in Figure 1, include:

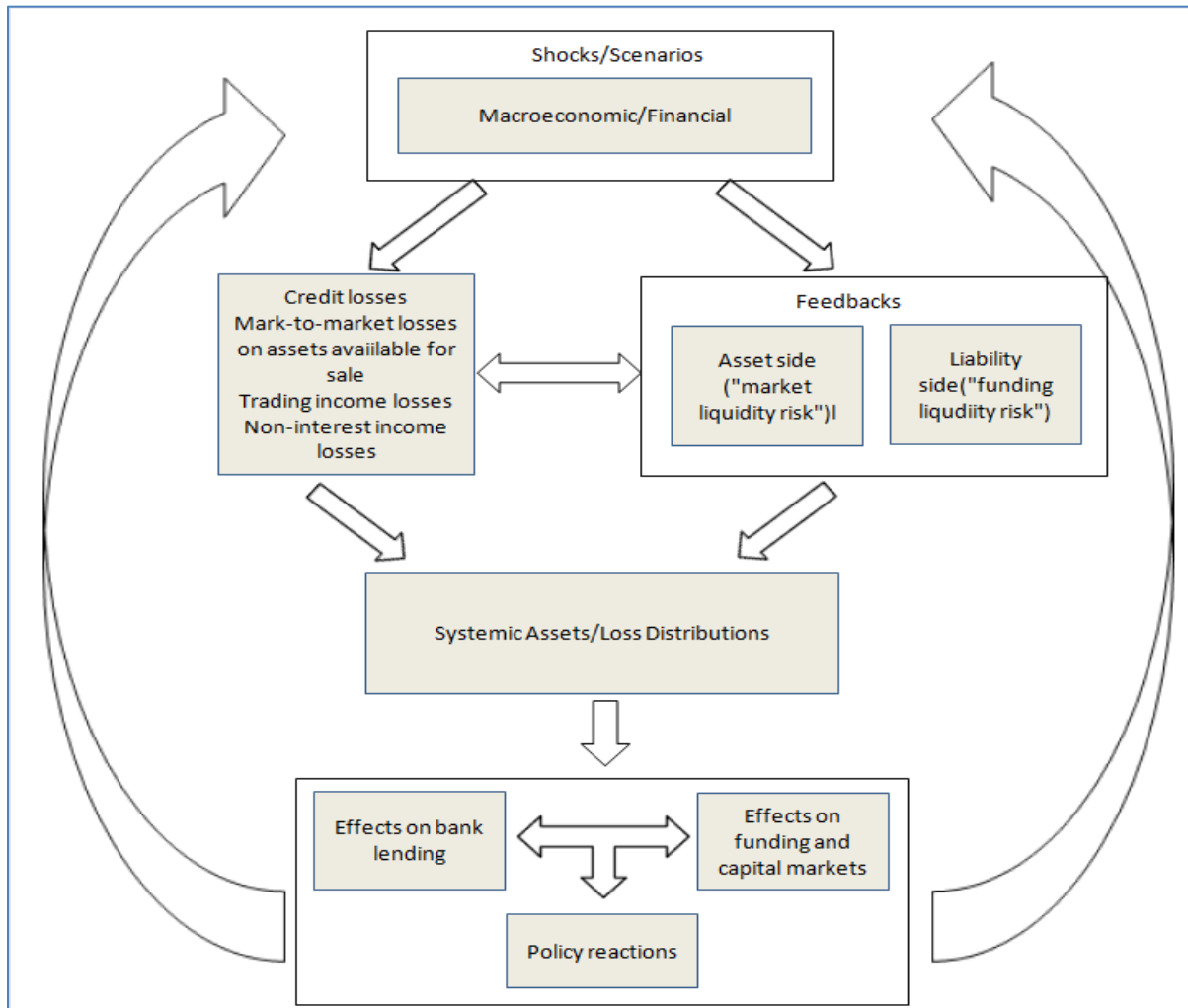
- The feedback between liquidity and solvency risks. This includes the (highly non-linear) relation between funding costs and risk perception by bank creditors; and the relation between asset sales motivated by banks' reactions to liquidity problems (fire sales) and further declines in asset values, that in turn aggravate solvency problems.
- The feedback from financial instability to the real economy. For example, bank reactions to financial stress (e.g., deleveraging, capital flight) can have effects on the real economy that in turn degrade bank asset values further (second-round effects).
- The bank-sovereign link. Traditionally, the key source of risk in the relation between banks and the sovereign was related to the size of fiscal contingent liabilities. However, the recent experience in Europe shows that the relationship between bank and sovereign risk is much more complex (discussed in more detail in Appendix IV).
- Policy feedbacks. Policy reactions (or lack of them) could have a significant impact on risk transmission and on the duration of a crisis.

34. Proper stress test design requires a careful examination of the transmission channels and an understanding of the range of possible responses of financial institutions and capital markets to different shocks. While progress is being made in all of these fronts, the pace is uneven. There are still gaps in our understanding of the interaction of the real economy with the

¹⁴ The Dodd-Frank Act allows the Federal Reserve Board to use its discretion to include banks and nonbanks in the stress tests.

financial sector—the macro-financial framework—and of the role of the plumbing of the financial architecture and business practices in amplifying and transmitting negative shocks.

Figure 1. Transmission Channels of Shocks Between the Real Economy and the Financial Sector



35. **The operational implementation of this principle remains a major challenge, especially when it comes to the feedback between the real and financial sectors.** As noted above, reliable stress tests require the calibration of propagation channels (with historical information or expert judgment) and their incorporation in stress test design and implementation in the face of incomplete information, including dealing with the tail risk arising from “unknown unknowns.” Progress is being made rapidly on models addressing some of the transmission channels highlighted above, with central banks and the IMF playing a leading role (Box 3). Although these models have enhanced our understanding of amplification mechanisms, such as those arising from funding shocks (e.g., the liquidity spiral of Brunnermeier and Pedersen, 2009), their incorporation in actual stress tests is a complex undertaking owing to the absence of a fully specified macro-financial model. Given this, some of these effects are sometimes integrated in an

ad hoc manner, by adding a second-round top-down stress test to bottom-up bank-by-bank stress tests to assess potential bank reactions to the first-round shock. For example, financial institutions could provide a qualitative discussion of responses they would adopt if they were to face such stress (e.g., portfolio allocation, deleveraging, or securing credit lines with central banks). Such responses could form a basis for improving the design of stress test exercise, e.g., designing the second-round shock by Financial Stability authorities.

36. Against this background, it is no surprise that stress testing practice generally falls short in this area.

- Most country authorities that responded to the survey, as well as almost all FSAPs, incorporate liquidity shocks in their stress tests, typically assuming deposit withdrawals and, in many cases, reductions in interbank positions and haircuts for liquid assets. In addition, a few also account for liquidity needs from off-balance positions and withdrawals of other types of wholesale funding. About half of the respondents consider domestic currency and foreign currency liquidity positions separately. However, most of these liquidity stress tests are implemented independent of solvency tests, excluding the possibility of experiencing a more severe run on liabilities when banks are likely to make substantial losses (illustrated by the case of Bear Sterns and the recent experience of some banks in the European periphery). Most large European banks compute their maximum risk tolerance (ECB, 2008) utilizing a stochastic approach, which aims at estimating their “liquidity-at-risk” (maximum liquidity gap within a certain time horizon and for a given confidence level) or their “liquidity value at risk” (maximum cost of liquidity under certain assumptions). These approaches do not adequately incorporate traditional credit risk or links to macroeconomic scenarios.
- Feedback effects from the financial sector to the macroeconomy or the complex bank-sovereign linkages are rarely incorporated in either FSAPs or country authorities’ stress tests. In some IMF programs, fiscal-financial linkages have been explicitly modeled by adding bank recapitalization needs estimated in a stress test to fiscal debt sustainability analysis as contingent liability to the government,¹⁵ or by relying on market-price based approaches (Appendix IV). But other types of feedback effects, such as the link from distressed and deleveraging banking sector to economic growth through credit crunch, are not captured. This field would require a new generation of macroeconomic models that include the financial sector and intermediation. Such research has been spurred by the crisis, in both policy institutions and academia, but it would take some time before sufficiently operational models are developed.¹⁶

¹⁵ See [IMF Staff Report for Greece, March 2012](#).

¹⁶ Two recent examples of the former include the GFSR (IMF [2011b](#)), which presented a DSGE model with a banking sector; and Bank of Japan (Ishikawa et al., [2012](#)), which developed an econometric model incorporating the interactions between the financial sector and macro economy. The IMF’s Research Department is also expanding its global DSGE—Global Integrated Monetary and Fiscal (GIMF) model—to explicitly include the financial sector.

Box 3. Integrating Liquidity and Solvency Risks and Bank Reactions in Stress Tests*

Banks have numerous ways to react to credit and funding shocks. High-quality capital and profits are usually the first line of defense, and retained earnings can help buffer banks' capital levels. Banks have an inherent capacity to generate liquid assets by using high-quality eligible securities as collateral for market or central bank funding if interbank markets freeze. As seen post-Lehman, fire sales of securities are also an option, but at a considerable cost in an environment of sharply declining asset prices. Deleveraging, especially targeted at assets with higher risk weights, is also a way to raise capital adequacy ratios by reducing risk-weighted assets (RWAs). In practice, banks have been using a combination of these, as well as other hybrid measures, ranging from debt-to-equity conversions to issuance of convertible bonds to optimizing risk-weighted assets, to react to shocks.

Incorporating banks' reactions to shocks is a critical input into the design of informative stress tests, especially over longer time horizons. This, however, requires modeling solvency and liquidity shocks in a coherent manner because *first*, when banks react to financial stress, the source of the shock (solvency or liquidity) is not always clear; and *second*, the measures banks take in reaction to these shocks have both capital and liquidity aspects that are not easy to disentangle.

A relatively simple (but somewhat ad hoc) way to integrate solvency and liquidity shocks is to conduct two-round stress tests, with a bottom-up (BU) first round and a top-down (TD) second round. If, for example, the majority of banks report in the BU first-round test asset sales of particular asset classes in response to the shock, the TD second-round test could impose haircuts on those assets; if banks report that they would discontinue reverse repos, the analysis would incorporate a reduction in repo roll-overs. The quantification of these haircuts or roll-over rates could be based on historical information, cross-country experience, or expert judgment.

Recently, a number of analytical approaches have attempted to integrate solvency and liquidity more systematically.

- [Schmieder et al. \(2012\)](#) simulate the increase in funding costs resulting from a change in solvency, indicated by a change in a bank's (implied) rating.
- The [Dutch Central Bank](#) developed a stress testing model that tries to endogenize market and funding liquidity risk by including feedback mechanisms that capture both behavioral and reputational effects. A number of central banks and bank supervisors have been successfully using this framework.
- The [Hong Kong Monetary Authority](#) sought explicitly to capture the link between default risk and deposit outflows. Their framework allows simulating the impact of mark-to-market losses on banks' solvency position leading to deposit outflows, asset fire sales by banks, and a consequent sharp increase in contingent liquidity risk.
- Barnhill & Schumacher (2011) developed a more general empirical model, incorporating the previous two approaches that attempt to be more comprehensive in terms of the source of the solvency shocks and compute the longer-term impact of funding shocks.
- Another attempt to integrate funding liquidity risks and solvency risk is the Risk Assessment Model for Systemic Institutions ([RAMSI](#)) developed by the Bank of England. The framework simulates banks' liquidity positions conditional on their capitalization under stress, and other relevant dimensions, such as a decrease in confidence among market participants under stress.

* Prepared by Heiko Hesse.

- On a positive note, the analysis of spillover effects across financial institutions are increasingly covered using network analysis or systemic risk measures using market-based approaches. The [Luxembourg](#) FSAP examined network effects including bank-by-bank cross-border exposures, including exposures to parents and subsidiaries in the same financial group. Among country authorities, stress testing models such as the [Bank of England](#)'s RAMSI model, for instance, add contagion effects using network models to macro scenario solvency tests. Market-based approaches (see Box 1) reflect interconnectedness across institutions in a reduced-form manner in a process of estimating systemic losses accounting for interdependence among institutions (e.g., the 2010 [U.S. FSAP](#)).

Principle No. 3: Include all Material Risks and Buffers

37. **Capturing all risks is key to obtaining reliable stress test results.** Until the global financial crisis, stress tests typically focused on credit risk from customer loans and market risk from marketable securities. The crisis revealed that this coverage was inadequate and other sources of risk, such as sovereign, funding, systemic liquidity, and counterparty risks, should also be included in the stress tests to capture all potential sources of vulnerability. For globally active financial institutions, incorporating cross-border activities through cross-ownership, credit and market risk exposures, and funding (including parent-subsidiary funding and liquidity transfers) is important for both home and host country authorities to assess the full risk profiles of the financial institutions.

38. **Nevertheless, incorporating some sources of risk in stress tests still remains controversial.** Some risks (e.g., own sovereign) are so large and hard to hedge that financial institutions and—especially—entire systems are likely to be very vulnerable to them, and protecting against them (e.g., through additional capital) is bound to be so costly as to be infeasible. Similar issues arise for system-wide liquidity shocks. This has led many to question the usefulness of stress testing these types of risk. Other risks have simply fallen out of fashion: for example, in an environment of low and stable inflation and interest rates, interest rate risks, such as refinancing and reinvestment risks arising from mismatches, are often being neglected. Regardless of the validity of some of these arguments, the incorporation of all risks in a stress test remains an issue of paramount importance in order to obtain a complete picture and guide the search for risk-mitigating solutions. Not all potential risks need to be addressed with additional capital, while comprehensive and candid stress tests could help gauge the consequences of inaction or delay (e.g., in addressing sovereign problems).

39. **In addition to risk factors contributing to losses, the impact of buffers, such as pre-impairment income, should also be part of the stress tests.** Many of the recent macro scenario tests, notably in FSAPs, have a two-year or even longer time horizon. Over such test horizons, profits can have a non-trivial impact on the test results. For instance, many of the supervisory stress test before the crisis estimated potential losses that were more than absorbed by projected pre-impairment profits; and in the EBA exercises, credit and market risk impairment costs were

mostly absorbed by projected pre-impairment profits. Ignoring these profit buffers could thus exaggerate the impact of the shocks. At the same time, overly optimistic expectations about potential fees and additional interest income under a stress scenario can mask the impact of losses. Therefore, the importance of understanding and modeling properly business conditions under stress cannot be overemphasized. Separately, test results could differ depending on how the impact of stress scenarios on RWA is incorporated: as with income, changes in RWA can affect the stress test results, in some cases substantially.

40. **Modeling the link between the non-impaired components of income and macroeconomic stress is a challenging task.** It is particularly challenging for top-down exercises for a number of reasons, including the lack of granular information (what assets and liabilities are based on fixed or floating interest rates, maturities, and banks' hedging practices); the complexity of the sources of bank income; and the likely changes in bank behavior under stress as they strive to protect their balance sheet (for example, many banks attempt to increase earnings through higher fees and commissions when entering a downturn). Recent approaches used in FSAP stress tests to tackle these challenges are described in Box 4. Even in bottom-up tests, this is an area where banks have many degrees of freedom, partly because there is no widely accepted single model for parametrization for non-impaired profits. Careful examination of individual banks' models and comparison of assumptions and results across institutions with similar business models should thus be an integral part of the exercise.

41. **Implementing this principle in practice requires that the stress tester scrutinize bank portfolios and have a good understanding of risks and business conditions before designing the stress tests.** This would enable the stress tester to capture all relevant risks and incorporate buffers appropriately in the design of the exercise. It also requires taking as comprehensive an approach to risks as possible, even if some of them may be "too big to mitigate". Finally, this principle also provides an argument in favor of bottom-up over top-down tests to the extent that stress conducted by the banks themselves (with adequate supervisory scrutiny) may be more comprehensive and informative, since banks know better their portfolios and risks.

Box 4. The Projection of Pre-Impairment Income Under Adverse Scenarios*

Stress test results depend not only on shock-related capital losses but also on pre-impairment income (profit before losses from loans and security portfolio), as income earned in the test period contributes to capital. This factor is particularly important to consider in stress tests with long time horizons. In many cases, the size of pre-impairment income can be substantial compared to the shock-related losses. But projecting pre-impairment income is complex, and there is no widely agreed methodology to do so. Instead, stress testers employ a range of techniques.

The most straightforward way to project bank profit in line with macroeconomic conditions is to estimate the elasticity of pre-impairment income-to-capital to economic growth. Hardy and Schmieder (forthcoming) estimate this elasticity using Bankscope data for more than 16,000 banks, and find the value ranging from 1 to 1.5: in other words, a one percentage point change in real GDP growth reduces the pre-impairment income-to-capital ratio by 1 to 1.5 percentage points. The average pre-impairment income-to-capital ratio in both advanced and emerging economies is about 12 to 15 percent, and a moderate stress event (a 4 percentage point drop in annual GDP growth in advanced economies and a 6 percentage point drop in emerging economies) would reduce it by about half. However, for banks in advanced economies, the elasticity goes up to 4 under severe stress conditions, implying a strongly non-linear relation between macroeconomic conditions and bank profits.

When data permit, components of pre-impairment income may be projected separately. Key subcomponents for banks are net interest income, fee and commission income, and non-interest expenses (including salaries). Net interest income could be estimated in line with the interest rate assumption included in macro scenario and assumptions for interest rate pass-through to bank borrowers. Hardy and Schmieder (forthcoming) find that net interest income is more closely linked to macroeconomic conditions than other sources of income (such as fees and commissions), except for trading income under highly adverse conditions. [A GFSR study](#) also documented the close linkages between net interest margin and the slope of yield curve. Fees and commission income could be projected by examining their sources: if the majority is related to the sales and trading of securities, it is reasonable to project them in line with asset prices; if fees are mostly related to loan origination and credit cards, they should be linked to credit growth or employment.

Net interest income, in particular, could also be estimated incorporating interactions between solvency and liquidity stresses. When depositors and wholesale fund providers suspect a bank may incur substantial losses, they might withdraw their funding, and the bank would need to offer higher interest rates in order to maintain access to funding. These higher funding costs, if not sufficiently passed onto borrowers, could reduce interest margins and net interest income. But as recently seen in Europe, wholesale funding costs are much more elastic than deposits (since at least part of the depositors' funds are protected by deposit insurance) and harder to pass onto borrowers in weak economic conditions. [Schmieder et al. \(2012\)](#) and other IMF exercises have therefore constructed a method linking a bank's funding cost to its solvency distress by using structural models for credit spreads (similar to Moody's KMV). While the increase in funding costs has been found to be relatively contained for well-capitalized banks, it rises sharply once banks come closer to the regulatory minimum. Depending on the portion of the increase in funding costs banks can pass on, their income (i.e., net interest income) can shrink substantially.

* Prepared by Christian Schmieder and Hiroko Oura.

42. The experience of the recent crises has spurred improvements in this area among stress testing practitioners, but there are still important gaps.

- The global financial crisis and ongoing European sovereign debt crisis, as discussed in Section III, illustrated that many of the relevant risk factors were not sufficiently addressed in stress testing exercises. Since then, the list of included risks has expanded to incorporate the missed factors and recent economic and regulatory developments. Before the crisis, credit risk, asset prices (mainly stock and real estate), and liquidity and funding

risk were the top risk factors taken into account in stress tests,¹⁷ based on the survey responses. The situation was similar for stress tests conducted in the context of FSAPs. After the crisis, the attention paid to liquidity and funding risk has increased, and new sources of risks are covered, including sovereign risk, low profitability, regulation-related risks, and contagion risks arising from interconnectedness.

- Sovereign risks have recently become more relevant in both FSAP and national authorities' tests. EBA tests in 2010 and 2011 and FSAPs in advanced economies since 2011 have been covering sovereign risk. In most cases, these are modeled as market risks by applying haircuts on sovereign securities.¹⁸ However, the methodologies differ across exercises, especially regarding the coverage of positions (stressing exposures to own sovereign in HTM account in particular, discussed more under Principle 4), and the size of the shock (which could be too small if shocks are calibrated based on historical data, as discussed under Principles 5 and 6).
- Stress tests are expected to capture supervisory weaknesses in emerging and developing economies. FSAPs typically attempt to quantify the over-reporting of capital adequacy by adjusting reported figures for potential under-provisioning, weaknesses in loan classification and provisioning rules, collateral overvaluation, forbearance, and concentration risk from large exposures, among other factors. However, these adjustments are judgmental and often involve large estimation errors, and should therefore be cross-checked with expert judgment of the supervisory authority.
- Ensuring methodological consistency in projecting pre-impairment income and RWA without dismissing each institution's idiosyncratic characteristics is a major challenge. In bottom-up exercises, the supervisory authority or the FSAP team typically attempt to impose some harmonization by enforcing a few uniform assumptions on pre-impairment income (for example, interests pass-through). On the other hand, the impact of a shock on RWA can differ across countries and institutions partly reflecting differences in regulatory treatment.¹⁹

¹⁷ [Ong and Čihák \(2010\)](#) discuss how the pre-crisis stress tests on Iceland generated deceptively benign results by ignoring liquidity risk (deposit withdrawal). They illustrated that liquidity stress tests using detailed pre-crisis disclosure information on funding positions could have indicated a vulnerability that materialized later.

¹⁸ Sovereign risk is ultimately a specific kind of credit risk. However, "sovereign risks" for advanced economies typically mean sovereign market or spread risks, which are mark-to-market (unrealized) valuation losses upon changes in market prices of sovereign bonds rather than outright default risks ([BIS, 2011](#)).

¹⁹ A stress scenario would not only affect capital (numerator) but also RWA (denominator). Survey results indicate that practices in calculating the latter vary ([Le Lesle and Avramova, 2012](#)). Banks under the Basel II advanced IRB approach (AIRB) calculate their RWA using borrower- or loan-specific PD/LGD data following Basel formulae. A deterioration in PD/LGD leads to increases in RWA. On the other hand, banks under the standardized approach apply specific risk weights depending on the type of exposures, and deterioration in loan quality does not necessarily affect RWA.

- Lastly, some risks are still often neglected in system-wide stress tests, such the impact of downgrades (capital requirements for credit risk are based on default probabilities only and do not include capital buffers for downgrading risk) and interest rate risks in the banking book. Incorporating risks from cross-border exposures fully could be a challenge even if their importance is recognized in both home and host countries, due to (i) data issues as home-host collaboration becomes critical; and (ii) need to construct global macro scenario, which might be outside the scope of macroeconomic modeling team of country authorities. Some recent FSAPs managed these difficulties by relying on bottom-up exercises and utilizing the IMF's global macroeconomic framework to provide global macro scenarios.

Principle No. 4: Make Use of the Investors' Viewpoint in the Design of Stress Tests

43. **Market perceptions of solvency and asset values are of paramount importance.** In the past decade, a large number of global financial institutions started to rely more on uninsured short-term wholesale funding and less on insured deposits. During the recent crisis, investors in these bank funding instruments, concerned about asset values and uncertain about the holdings and valuation practices of the institutions, triggered confidence shocks that caused major bank distress. Delays in recognizing that the crisis was motivated by solvency concerns, as well as political difficulties in finding solutions to address these concerns, made the crisis longer and deeper. The bottom line is that in a world where a large fraction of financial institutions' liabilities are uninsured and sovereigns default, markets can effectively force a bank closure. This can be done by penalizing a bank with higher funding costs and, at the limit, depriving the bank from additional funding altogether. This market-imposed discipline will clearly affect bank performance through higher losses, need to deleverage, and possible further impairments triggered by second-round effects.

44. **The operational implication of this principle is that market views need to complement stress tests based on regulatory and accounting standards.** There are several ways to do this.

- **Adopting mark-to-market (MTM) methodologies to the valuation of all bank assets and liabilities under the baseline and adverse scenarios.** In some cases, this would mean ignoring the prevailing accounting approach used for securities held to maturity, namely amortized cost principle net of any impairment provision based on incurred loss. This approach assumes that these securities are not affected by market swings and consequently changes in their value do not have an impact on financial institutions' equity. While this may be true from an income perspective (to the extent that these securities would not be sold in an adverse scenario), it may not be true from a valuation perspective used by investors to assess the institutions' risk profile and funding costs. Given that funding costs are an important element in all stress tests, adopting a full mark-to-market approach would provide a useful benchmark.

- **Using economic rather than statutory capital as the basis for stress tests.** Economic capital is a concept that attempts to capture the bank’s underlying economic value, and may deviate substantially from statutory capital, calculated on the basis of the jurisdiction’s current regulatory and accounting standards. For instance, in the event of a substantial decline in asset prices, banks’ securities holdings could carry unrealized losses that are not fully reflected in regulatory or accounting capital.²⁰ Similarly, loan provisioning may become inadequate once substantial drops in collateral valuation are fully accounted for.
- **Using point-in-time (PIT) parameters to measure expected and unexpected losses.** PIT parameters, as opposed to regulatory approaches to capital measurement typically based on through-the-cycle (TTC) parameters, may be a better way to reflect investors’ assessments of measures of economic capital.
- **Stressing market risk appetite.** Market risk appetite can be explicitly stressed when designing stress test scenarios. For example, extreme adverse scenarios that represent a severe global crisis—say, of the magnitude of the 2008–09 crisis—could incorporate an increase in the market price of risk similar to that during 2008–09. This would have an important impact on the bank’s risk-adjusted balance sheet, leading to higher credit spreads/funding costs, higher potential losses to bank creditors, and lower equity values. If these effects are not accounted for, the distribution of bank losses under an adverse scenario may be underestimated and lead to an overly optimistic assessment of bank capital during stress.²¹
- **Imposing hurdle rates based on targeted funding costs.** Hurdle rates based on regulatory ratios reflect what regulators consider an adequate solvency ratio, but the markets’ assessment of a bank’s solvency may be different. And in a world where markets are able to impose discipline on banks, markets may demand—and banks would have an incentive to target—capital ratios that enable them to attain a certain risk rating and/or keep their funding costs under a certain ceiling. Using hurdle rates that reflect market views (in addition to the regulatory minima) in stress tests recognizes this simple but stark reality. Otherwise, if stress actually materializes, banks that “passed” the stress test may fail the market test—a situation that arose recently in Europe. Box 5 presents two approaches developed by Fund staff to calculate market-based hurdle rates. For

²⁰ This idea was implemented by [the EBA 2011 EU Capital Exercise](#), which imposed the MTM of banks’ all European Economic Area (EEA) sovereign debt exposures held to maturity or AFS, in order to calculate the banks’ actual capital ratios and impose a mandatory capitalization of any capital shortfall.

²¹ The market price of risk is the compensation that a risk-averse investor would require for getting into a risk position. Many stress tests methodologies use actual default probabilities (e.g., historical number of historical defaults in an industry or non-performing loans). But these approaches do not reflect the true price of risk, and also ignore its volatility over time. Several papers have developed methodologies to convert historical default probabilities into risk-adjusted default probabilities (also called risk-neutral default probabilities), for instance [Espinoza and Segoviano \(2011\)](#) among others. FSAPs to [Israel](#) and [Sweden](#) have included the market price of risk in stress tests.

supervisory authorities or central banks, which typically have better access to bank-specific and market information, the identification of the risk-return trade-off as perceived by the market should be simpler.

45. **Imposing hurdle rates that may be higher than regulatory minima is particularly important for macroprudential stress tests.** From a macroprudential standpoint, financial institutions have to be sufficiently capitalized not only to ensure their own viability in the event of a system-wide shock but also to prevent them from becoming transmission channels. This may well require higher capital than is necessary when a financial institution is considered on a stand-alone basis ([OFR, 2012](#)).

46. **This principle also has implications for the decision to disseminate or not stress test results.** Publishing stress test results can help remove asymmetric information during periods of uncertainty and restore market confidence. Even in the case of stress tests undertaken for surveillance purposes during non-crisis periods, communication of their results could create awareness of risks, promote more realistic risk pricing, and enhance market discipline during good times—which, in turn, could reduce the probability of future, sudden reversals of investors' mood. However, for the publication of results to yield these benefits, stress tests need to be candid assessments of risk, explicit about the coverage and limitations, and the announcement of their results needs to be accompanied by the measures that will convincingly redress any vulnerabilities unveiled by the stress tests, including but not necessary limited to capital injections.

47. **Actual stress testing practices generally fall short of this principle.**

- Market-based hurdle rates are generally not used by country authorities, and only rarely used in FSAPs ([U.S. FSAP 2010](#), [IMF 2011a](#)). However, following the crisis, funding costs are increasingly being modeled and incorporated explicitly in stress tests. While this is an important improvement, it is not enough. As shown in Box 5, the relationship between funding costs and bank solvency is non-linear: once funding costs start to rise, it becomes difficult to bring them down, access to funding can quickly dry up, and a liquidity shock may become a solvency shock. This suggests that there are some market-related solvency thresholds that can and should be used as hurdle rates.
- The assessment of sovereign risk is another area where market considerations are generally not incorporated in stress tests. Regulatory and accounting standards could understate the valuation risks for sovereign securities substantially when a significant share of these securities is held in HTM accounts,²² as is the case in many European countries. Some recent FSAPs have started including unrealized losses from securities in HTM accounts by applying mark-to-market valuations in response to yield changes, recognizing that it is the economic rather than the accounting valuation that matters for sustainability. But some national authorities or financial institutions have not embraced

²² See footnote 3 for the definition of AFS, HTM, and trading account and the treatment of securities in different types of accounts in supervisory framework.

Box 5. Market-Based Hurdle Rates*

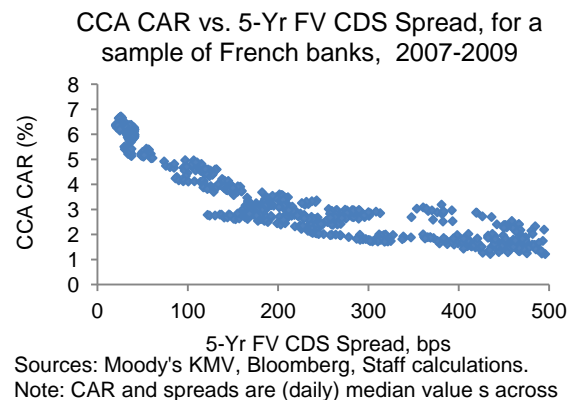
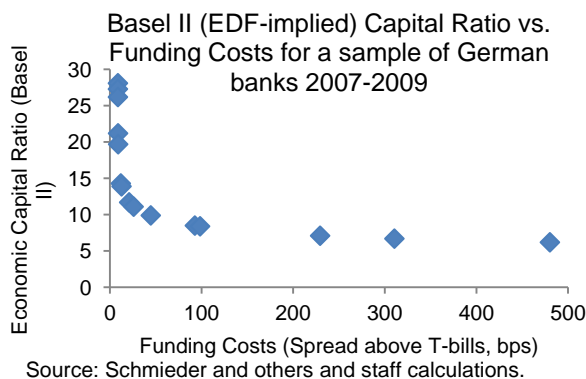
As a complement to regulatory hurdle rates, market-based hurdle rates should be used in stress tests. These would be the capital ratios that would be compatible with stable funding costs, which in many cases is the relevant “market test” for banks in a stress scenario.

This market-based hurdle rate can be calculated directly from the trade-off between banks’ capital ratios and their funding costs. For example, [Schmieder et al. \(2012\)](#) derive a non-linear relationship between solvency (measured as Basel II IRB-implied capitalization ratios consistent with banks’ default probabilities) and funding costs for a sample of German banks as follows:

- Funding costs (i.e., bank interest expenses measured as the excess spread above the government bond rate) are estimated using data provided by the ECB for an average German bank for 12 quarters during 2007–2009.
- German banks’ weighted average MKMV’s Expected Default Frequency (EDF) is then plotted against the funding costs for the same 12 quarters.
- The EDFs are translated into capitalization ratios using the IRB formula to establish a link between funding costs and capital adequacy ratios as shown in the left hand-side figure below (the mapping includes an additional capital cushion of 2.5 percentage points, in line with empirical evidence).

The results suggest that a regulatory capital ratio of about 15 percent would be a good benchmark. According to the authors’ calculations, this would be equivalent to a Basel III Core Tier 1 capital ratio of about 11 percent. Below this level, banks become very sensitive to funding costs as the relationship becomes highly non-linear: increases in funding costs—assuming that banks can pass only a fraction of that on to their customers—when capital ratios are below 10 percent can get out of control (the spread increases by about 50 bps when capital ratios go from 9.9 percent to 8.5 and by 184 bps when capital ratios are reduced further down to 7.1 percent). On the other hand, savings in funding costs above 15 percent appear small.

The Contingent Claims Analysis (CCA), using option pricing theory, recovers the banks’ probability of default implied in bank stock prices, the market value of bank assets (i.e., the value of assets adjusted by the banks’ default risk), and the market value of risky debt. The latter can also be thought of as the risk premium required by the bondholders to compensate for the expected losses on their claims. This information can be used to construct a CCA capital ratio (market capitalization to market value of assets ratio) that can be mapped onto the corresponding spread required by banks’ creditors. The right hand-side figure below presents these results for a sample of French banks. Again, the relationship is non-linear and provides a methodology to determine a market-based hurdle rate.



* Prepared by Liliana Schumacher.

this approach. The EBA, for instance, chose not to account for unrealized valuation losses from sovereign securities in HTM accounts in their [2010](#) and [2011](#) exercises, deciding instead simply to disclose the details of exposures, including those in HTM account. On the other hand, the Bank of Japan regularly tests interest rate shocks on all securities for its Financial System Report, including those in HTM accounts.

- Adjusting the initial capital for already realized shocks is often an integral part of supervisory and crisis management exercises. For instance, the [EBA's capital exercise of end-2011](#) required banks to mark-to-market all their sovereign bond holdings incorporating the large yield changes materialized in 2011. The stress test exercises conducted by Blackrock for Ireland and Greece for estimating recapitalization need included re-assessment of existing loans by going through re-underwriting process. However, surveillance stress tests often do not include these adjustments.
- On a positive note, recent FSAPs and EBA exercises are ensuring the use of point-in-time PDs.

Principle No. 5: Focus on Tail Risks

48. **The rule of thumb for stress tests has traditionally been to apply “extreme but plausible” shocks, but there is no systematic way to determine these.** Typically, the size of shocks is calibrated on past experience: “worst-in-a-decade” events, “one percent probability” tail events,²³ or an “x standard deviation” shock.²⁴ One obvious problem with this approach is that historical experience varies across systems and changes over time: a “worst-in-a-decade” scenario today would look very different than the same scenario in 2007. Another problem, particularly relevant for stress tests conducted before the crisis, is that the history of many financial products (for example, asset-backed securities) was too short to provide a reasonable amount of volatility, and the shocks calibrated on this experience were inevitably too mild. And obviously, an approach based on historical data would not work when considering an event that has never happened. [BCBS \(2004\)](#) has provided some quantitative guidance for determining tail risks for single factor shock stress tests.²⁵ However, there is no comparable guidance on macro scenario tests.

49. **This conundrum is more severe when stress tests take place in crisis or near-crisis situations.** In these cases, the financial institution or system is already experiencing significant distress, with substantial uncertainty domestically and possibly externally. Some supervisory authorities may thus be reluctant to use excessively negative tail scenarios on top of an already

²³ If a 50 percent decline in equity prices happened only once in the past 100 years, it is a one percent tail event.

²⁴ Assuming a normal distribution, a two standard deviation negative shock would have a 2.275 percent probability of occurrence, which can be thought of as approximately equivalent to a once-in-fifty-year event.

²⁵ [BCBS \(2004\)](#) suggests for G10 currencies either a ± 200 basis point parallel rate shock or 1st and 99th percentile of observed interest rate changes using a one-year holding period and a minimum five years of observations.

stressed baseline. Moreover, publishing the outcome of stress tests that incorporate extreme scenarios in these circumstances could trigger self-fulfilling crises. Political economy factors could also come into play when choosing scenarios, especially when the results are supposed to be used as input for determining bank closure or rescue, recapitalization needs, or regional/international rescue packages. On the other hand, as indicated by the contrasting experience of SCAP and EBA tests (Appendix III), compromising on the severity of scenarios could undermine the credibility of the exercise and prolong the crisis. These trade-offs are not easy to tackle. In principle, an effective stress test in a near-crisis situation should not compromise on the severity of scenarios, but instead mitigate the possible adverse market impact of the results by having credible support measures in place. If this is not feasible, it might be preferable not to conduct stress tests at all, but release critical exposure data and provide qualitative analysis to enhance transparency.

50. Another shortcoming of traditional approaches, regardless of the size of assumed shocks, is that they ignore the interdependence among shocks and among affected institutions or systems. Modest but correlated shocks can cause a system to break down and generate extreme outcomes if the correlated nature of shocks is not taken into account. For stress tests to provide a reliable assessment of the resilience of an institution or an entire financial system, they must consider not just the potential size of individual risk factors but also their dependence under all plausible scenarios. In addition, risk factors may be weakly correlated under normal economic circumstances but highly correlated in times of distress.²⁶ Similarly, the joint default risk within a system varies over time and depends on the individual firms' likelihood to cause and/or propagate shocks arising individual risk factors. Given that large shocks are transmitted across entities differently than small shocks, measuring non-linear dependencies in stress tests can deliver important insights about the joint tail risks.

51. But while it is relatively straightforward to generate stress test results based on the effect of a single risk on a static measure of capital, combining multiple risks and the extent to which firms' default risk may be correlated under different scenarios is very complex.

- In the context of balance sheet-based models, one way to capture these interdependencies and hence possibly extreme outcomes is to use network models (discussed in Box 2), that enable the measurement of spillovers and the probability of many institutions defaulting. However, most conventional balance sheet-based stress test models do not formally account for default dependencies across institutions. In this case, complexity arises from the amount of information necessary to identify the right network.
- In market price-based models, a full modeling of the distribution of institutions' joint default probabilities is possible. In this case, the key challenge is modeling the dependence structure. One approach used frequently is estimating Conditional Value-at-Risk (CoVaR),

²⁶ In these instances, there is a considerable shift of the average away from the median ("excess skewness") and a narrower peak ("excess kurtosis") of the probability distribution. If distributions become highly skewed, large tails may even cause the mean to become undefined, which is an important complication when using stress tests.

which measures the VaR of a financial institution conditional on the distress of other institutions in the system. The value of an institution and distress correlations among institutions are estimated using equity or other market price data. Another related approach is the Marginal Expected Shortfall (MES) (Acharya et al., 2010) that specifies historical expected losses conditional on having breached some systemic risk threshold. Adjusting MES by the degree of firm-specific leverage and capitalization yields the Systemic Expected Shortfall (SES), which yields the average, linear, bivariate dependence between banks when the entire banking sector is undercapitalized.

- More complex approaches require a significant departure from conventional statistical methods (Box 6).

52. Actual stress testing practices in this area are evolving rapidly in light of the experience with the crisis and in line with the development of new analytical tools.

- Recent FSAPs are making a systematic effort to include severe shocks that are comparable across countries.²⁷ One rule of thumb applied in many FSAPs is to assume a two-standard deviation shock on GDP growth rate for two years based on a long (20–30 years) history, unless a different magnitude of shock (in either direction) are justified. Among country authorities, most survey respondents adopt tail events with small probability (ranging from 1–5 percent), while some use qualitative criteria, such as in line with, or worse than, historical worst. Since the global economy has just experienced extremely sharp economic deterioration in 2008/09, shocks calibrated using a sample that includes the crisis time observations should generate fairly severe shocks. Another frequently used is to test for a shock of similar magnitude as the recent crisis. At the same time, all historical data-based approaches always carry the risk of complacency. This needs to be managed by more qualitative approaches (Principle 6).
- Variants of these approaches have been followed in choosing the severity of scenarios in recent stress tests. In the U.S. SCAP, while the macro scenario was rather moderate compared to the actual turnout, the corresponding rise in loan loss rate was comparable to the Great Depression. On the other hand, EBA exercises assumed only a moderate deterioration in sovereign yields, which was quickly exceeded by market developments; indeed in the 2011 exercise, EBA had to provide revised assumptions in the middle of the exercise. The 2010 U.S. FSAP examined a moderate shock on top of the baseline, pointing out that it would push the output gap to one of the lowest levels in the post-war history. The [2012 Spain FSAP](#) complemented quantitative stress testing with qualitative assessments.

²⁷ The effort is partly in response to [IEO \(2006\)](#) report that said FSAP should aim for more severe stress scenarios applied in a similar manner across major countries.

Box 6. The Importance of Risk Interdependence in System-Wide Stress Testing*

Estimating the dependence of risk factors to account for the low-probability of (negative) extreme outcomes with no or little historical precedent is not straightforward and requires a significant departure from conventional statistical methods. The traditional correlation coefficient detects only linear dependence between two variables (or risk factors) whose marginal distribution is assumed to be normal. This statistical inference presupposes an empirical relation (or the lack thereof) based on relatively more central (and more frequent) observations, and also implies that the bivariate distribution of these variables is elliptical, which is hardly encountered in reality. Alternative measures of dependence between risk factors can capture the non-linear dynamics of changes in variables far removed from the median. For example, an expedient non-parametric method of investigating the bivariate empirical relation between two random vectors is to ascertain the incidence of shared cases of cross-classified extremes via a refined quantile-based *Chi*-square statistic of independence. Similarly, copula functions and other non-parametric methods provide the possibility to combine two or more distributions of variables based on a more flexible specification of their dependence structure at different levels of statistical significance. These approaches generate measures of so-called “joint asymptotic tail dependence”, which define the expectation of common extreme outcomes.

The measure of extreme dependence across institutions can be combined with a severity distribution of each risk factor (or the risk profile of each firm) to derive a tail-sensitive estimate of joint default risk and/or the system-wide capitalization under stress. As part of the valuation of the different risks affecting the operating performance of a bank, expected losses can be modeled by explicitly taking into account extreme events. For instance, closed-form methods under extreme value theory (EVT), such as the generalized Pareto distribution, are frequently used to help define the limiting behavior of extreme observations. Generalized extreme value theory (GEV) models the asymptotic tail behavior of the order statistics of normalized maxima (or minima) drawn from a sample of dependent random variables, whereas the generalized Pareto distribution is an exceedance function that measures the residual risk of extremes beyond a given threshold (i.e., a designated maximum (or minimum)) as the conditional distribution of mean excess. The estimates of joint default risk under extreme scenarios would be used to cross-validate results from traditional stress testing approaches.

The Systemic CCA framework extends contingent claims analysis (CCA) to measure systemic risk from the interlinkages between individual risk-adjusted balance sheets of firms, and represents a suitable approach of quantifying non-linear changes of systemic financial sector risk during times of stress. This market-based framework applies option-pricing theory to market information—thus explicitly acknowledging the potential role of non-linearities in the measurement of default risk—to derive estimates of joint market-implied expected losses in a system of financial institutions using a multivariate set-up.¹ For this purpose, the concept of EVT is employed to generate a multivariate extreme value distribution (MGEV) that formally captures the potential of tail realizations of market-implied joint potential losses. The analysis of dependence is completed independently from the analysis of marginal distributions, and, thus, differs from the classical approach, where multivariate analysis is performed jointly for marginal distributions and their dependence structure by considering the complete variance-covariance matrix, such as the MGARCH approach.

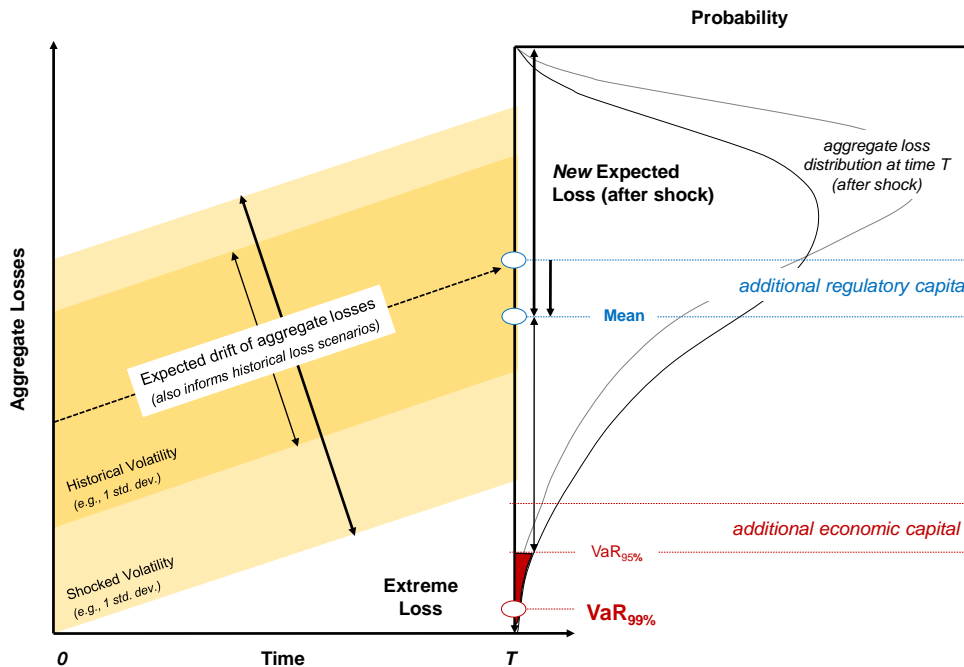
The estimation of joint expected losses within the Systemic CCA framework follows a two-step process. After estimating the non-parametric dependence function of individual expected losses, it is combined with their marginal distributions, which are assumed to conform to a GEV distribution. These marginal distributions are estimated via the Linear Combination of Ratios of Spacings method, which identifies the asymptotic tail behavior of normalized extremes. The dependence function is estimated iteratively on a unit simplex that optimizes the coincidence of multiple series of cross-classified random variables—similar to a Chi-statistic that measures the statistical likelihood of observed values to differ from their expected distribution. Finally, the conditional Value-at-Risk (VaR) estimate of joint expected losses is determined as the probability-weighted residual density beyond a pre-specified statistical confidence level.

* Prepared by Andreas (Andy) Jobst.

¹ In contrast to the traditional (pairwise) correlation-based approach, this method links the univariate marginal distributions of expected losses in a way that formally captures both linear and non-linear dependence in joint asymptotic tail behavior over time.

Box 6 (contd.). The Importance of Risk Dependence in System-Wide Stress Testing

Key Conceptual Differences in Loss Measurements—Implied Capital Requirement under “Distributional Approaches”



In general, incorporating time-varying extreme dependence of risk factors that determine system-wide default risk—ideally together with a market-derived measure of capital adequacy—offers a more realistic capital assessment in stress tests. This approach identifies the amount of capital shortfall to current capital levels based on market expectations on solvency that far exceed the minimum regulatory requirements. For instance, the assumption of higher uncertainty about the realization of expected losses (i.e., greater historical volatility) reflects the notion that, especially during distress periods, firms would need to have higher capital buffers in place to absorb the realization of losses above existing provisioning levels so that current capital levels remain unaffected over a specific risk horizon.

- Incorporating risk interdependence in stress tests continues to remain a difficult technical challenge. Based on the survey results, a number of central banks and supervisory authorities attempt to capture risk interdependence heuristically by applying common macro and financial shocks and assessing interbank contagion through network models. The Fund has also been increasingly using network models on stand-alone basis or as a part of a stress testing model. Beyond these heuristic approaches, some of the recent FSAPs tried to include some of these non-linear dependences formally in the stress testing framework. In the FSAPs in [Germany](#), [Sweden](#), the [United Kingdom](#), and the [United States](#), institution-level stress tests were combined with an attempt to derive capital assessments from a system-wide perspective, which considered losses from a portfolio of key financial institutions after controlling for the dependence structure of risk factors and the stochastic nature of input parameters. Technical limitations, however, mean that it will take some time before these approaches are mainstreamed in FSAPs.

Principle No. 6: When Communicating Stress Test Results, Speak Smarter, Not Just Louder

53. **Since the onset of the global crisis, public communication of stress test results has gained momentum.** Several countries were disseminating the results of stress tests in their Financial Stability Reports even before the crisis, although the degree of detail in describing the coverage, assumptions, and results varied.²⁸ FSSAs—most of which are published with the consent of country authorities—always included stress test assumptions and results, the latter on an aggregate basis or in a way that preserved the anonymity of individual financial institutions. But the aftermath of the crisis saw an unprecedented degree of disclosure of stress test results in the U.S. and Europe, where policy-makers saw it as a way to shore up market confidence. In the U.S., Congress enshrined this into law.²⁹ This spurred greater public interest in—and scrutiny of—stress tests which, in turn, increased pressure for greater disclosure.

54. **The experience has demonstrated that public disclosure of stress tests can yield significant benefits, but also highlighted that disclosure alone is no panacea.** The U.S. SCAP was successful in restoring confidence, allowing investors to differentiate between banks based on their resilience, and arguably facilitating the raising of additional capital from private sources. The 2011 EU-wide stress test, on the other hand, did not fully achieve its goals. This was not due to differences in the extent of public disclosure between the two exercises: indeed the 2011 EU-wide exercise was praised for its transparency. Rather, it reflected differences in the design of the stress tests and in the context within which their results were published. The setup of the SCAP was credible, and backstop measures, including government support, were clearly communicated and in place ahead of time. The EU-wide tests were seen as inadequately severe, not fully capturing the risk profile of weaker national banking systems, and—perhaps more importantly—the follow-up actions and policy backstops for banks that failed or passed only marginally were considered ambiguous (Appendix III compares the two exercises in more detail).

55. **Public disclosure may also create difficult tradeoffs in some cases.** Public disclosure of stress test methodologies, underlying exposures, assumptions, and results can help raise public awareness of risks; promote more realistic risk pricing and strengthen market discipline, thereby reducing the probability of future, sudden reversals of investors’ sentiment; and promote a meaningful conversation about financial stability policies. Even when the results are weak, public communication can have a positive impact if it is accompanied by credible contingency plans and support measures for financial institutions that fail the tests, which reflect the authorities’ recognition of the problems and commitment to financial stability. At the same time, disclosure has risks: it may entice financial institutions to make portfolio choices to “game” the tests; increase moral hazard if investors rely excessively on published stress test results—which are always subject to a margin of error (see Principle No. 7)—at the expense of other bank soundness

²⁸ Sweden’s Riksbank was one of the early adopters, and an example of extensive disclosure of stress test details, including bank-by-bank results for the four largest banking groups, on the basis of publicly-available data.

²⁹ The Dodd-Frank Act requires the Federal Reserve Board to disclose summary results of stress tests for large banks.

indicators; and actually undermine confidence, if the necessary support measures are not in place (for political economy or other reasons). Also, as stress tests become more common, disclosure of different results can cause confusion—a problem encountered in recent FSAPs in EU countries that took place concomitantly with EU-wide stress tests.

56. Balancing the benefits and costs of disclosure depends partly on the circumstances and the nature of the tests.

- In the midst of a crisis, when market confidence is at a premium, the case for public disclosure of the details of system-wide stress tests is compelling. In normal times, the case is more finely balanced. Even then, however, regular publication would familiarize market participants to stress tests, making them a more effective tool at times of crisis.
- The case for disclosure of system-wide stress tests, be they for macroprudential surveillance or for crisis management, is much stronger than for microprudential stress tests. The former are more likely to apply consistent assumptions across financial institutions, making the results comparable, and focus on systemic risk factors that are relevant for maintaining or restoring market confidence.

57. In addition, realizing the benefits of greater disclosure depends on a number of crucial pre-conditions. The stress tests should be credible. For this, they need to cover the relevant risks and transmission channels, assume serious shocks, set appropriate hurdle rates, and produce a candid assessment. And crucially, they should be accompanied by a convincing framework for follow-up action, including government support, if needed. If these pre-conditions are not met, disclosure would not be informative and might do more harm than good.

58. Lastly, the disclosure of stress tests should also be set in the perspective of the broader communication strategy of financial stability policies. Just as stress tests should be one of several assessment tools, so should their disclosure be part of a coherent overall communication strategy. Publishing stress test results is likely to be much more effective if done in the context of regular outreach aimed at informing markets and the public about financial stability issues, and complemented by disclosure of a broad set of indicators.

59. The survey of actual practices confirms the trend toward greater disclosure.

- The majority of respondents to the survey communicate the results of macroprudential stress tests (almost 85 percent in the case of solvency tests and 50 percent in the case of liquidity tests). Usually, this takes place in annual or semi-annual Financial Stability Reports, placing stress tests in the context of a broader financial stability assessment. In most cases, results are communicated using system aggregates, often with some distribution metrics, without disclosing the identity of individual institutions—with the notable exceptions of the recent EBA and U.S. Federal Reserve tests. FSAP stress test results are reported in FSSAs, which also typically do not disclose the identity of individual institutions. At 60–70 percent, the publication rate of FSSAs has always been

relatively high, and has risen in recent years. FSAP Technical Notes on stress tests, however, that report much more detailed results, are usually not published.

- Raising public awareness of financial stability issues, achieving greater transparency, and providing information to market participants are mentioned as the objectives of public communication by most respondents that publish stress test results.
- Although public communication is seen positively by all respondents that publish, many expressed concerns about the risk of exaggerated expectations placed on stress tests, inconsistent interpretation of stress tests by mass media, and excessive focus by banks on published stress test results that could undermine their effectiveness as a supervisory tool.

Principle No. 7: Beware of the “Black Swan”³⁰

60. **Regardless of how extensive the coverage of risk factors, how refined the analytical models, how severe the shocks incorporated in the stress tests, and how careful the communications strategy, there is always the risk that the “unthinkable” will materialize.** Stress tests cannot predict the future: they can only provide a measure of the resilience of a financial institution or a system to the shocks assumed by the stress tester. But the shocks of the future will likely arise from completely new products, unexpected events (e.g., the breakup of currency unions), factors that have historically shown little volatility, or risks that have not materialized for such a long time that they have been forgotten (e.g., advanced country sovereign defaults). What practical ways are there to incorporate these factors into stress test design?

61. **One approach is to supplement the traditional ways used to identify possible shocks with expert judgment and new information, where available, rather than simply be guided by history.** The Bank of England (Haldane et al., 2007) has proposed using current vulnerabilities as a guide for the choice of the scenario. This means, for example, that systems that are concentrated in real estate deserve a stress test on the impact of a large decline in real estate prices regardless of the probability of such a shock actually taking place.

- The Federal Reserve Board typically uses two scenarios for stress tests: one is unique to each institution and chosen by it; the other is common. In this way, the institutions are assessed under scenarios that they themselves consider particularly damaging.
- There are also policy trade-offs in designing shocks. On one hand, it is important not to let weak banks pass the test (Dexia is a good example of a bank that passed the July 2011 EBA stress tests and failed shortly afterwards). On the other hand, it is not useful to design shocks that are too large and fail many banks that are sound under many scenarios.
- Reverse stress test by individual institutions (Appendix I) and surveys of such exercises across institutions could help extend the frontier of tail risks.

³⁰ The term “black swan” was first used by Nassim Taleb in his 2004 book *Fooled by Randomness* to indicate highly improbable events that have a major impact.

62. **Another approach is the application of distribution theory to the scenarios themselves, as opposed to the current practice of choosing just one adverse scenario.** This reflects the recognition that the future is stochastic and can be represented by a number of event combinations, each of which with a probability of realization. A scenario distribution approach was used by IMF staff in the first South Africa FSAP. Based on the statistical properties of the historical distribution of price changes (key distribution moments) in South Africa, and using Monte Carlo simulations, each scenario was represented by a combination of changes in prices, including credit spreads, which were used to revalue bank assets. The final outcome was represented a distribution of bank capital ratios for each bank, in which each point of the distribution was associated with a particular scenario.³¹

63. **Ultimately, this principle is more about the context and proper use of stress tests than about the mechanics of their design and implementation.** It is thus not easy to assess to what extent actual stress testing practice complies with it. Instead, this principle should serve as a reminder that stress tests should not be undertaken in isolation and their results should not be taken too literally. No matter how much a stress tester tries, stress tests always have margins of error. Their results will almost always turn out to be optimistic or pessimistic ex post. In addition, there will always be model risk, imperfect data access, or underestimation of the severity of the shock. One should therefore set stress test results in a broader context.

- Stress testing is just one of the many tools to assess key risks and vulnerabilities in financial institutions or entire systems. They should be treated as complements to other tools that can provide information about potential threats to financial stability, such as qualitative and quantitative bank risk analysis, early warning indicators, models of debt sustainability, and informed dialogue with supervisors and market participants, among others. Final conclusions about the resilience of the institution or system should draw on all these sources and not just on the results of stress tests.
- Stress test design, models, and implementation should be back-tested to the extent possible and regularly re-assessed. Back-testing can take the form of a comparison of stress tests outcomes under baseline scenarios with actual outcomes. For adverse scenarios, one could have stress tests results reviewed by a panel of external experts to assess their rationale and consistency of results across banks. Checking the robustness of the results for variations of key parameters (in other words, stress testing the stress test), assessing the impact of new tools and new approaches and, last but not least, remaining vigilant for the emergence of new risks are crucial to ensuring more reliable tests.

³¹ The methodology for the first South Africa FSAP is discussed extensively in Barnhill et al. (2002). This approach is similar to the line taken in [Borio, Drehmann, and Tsatsaronis \(2012\)](#) that suggest that stress tests themselves need to be stress tested by assessing the sensitivity of results to changes in assumptions.

V. CONCLUSIONS AND OPERATIONAL IMPLICATIONS

64. **This paper contributes to the debate on stress test design and implementation by proposing a set of operational best practice principles.** The wide variety and somewhat ad hoc nature of stress testing approaches and underlying assumptions thus far has given rise to questions about the interpretation of their results and their comparability, at a time when stress tests are being put to new—and publicly more prominent—uses. Setting best practice principles and ensuring adherence to these could improve the integrity of the exercise, promote greater candor and comparability of stress tests across countries and over time, and ultimately contribute to more meaningful and effective financial stability assessments.

65. **A key goal of the paper is also to set realistic expectations about what stress tests can and cannot accomplish.** Stress tests are not intended to forecast crises. They are forward-looking tools to assess financial institutions' solvency and liquidity and the resilience of the entire financial system under possible adverse scenarios, but do not predict the likelihood of these scenarios materializing. As such, regardless of refinements and improvements, they will always remain hypothetical statements. One should therefore always be cautious about using stress test results in isolation: a well-rounded risk assessment should use stress tests in conjunction with other tools to broaden the understanding of vulnerabilities.

66. **The discussion highlighted a few critical decisions that stress testers need to make at the outset.** These are the choice of risk scenarios—in terms of both the coverage of all relevant risk factors and their severity; the design of the tests so that they cover all relevant transmission channels and include realistic assumptions about buffers; and the choice of hurdle rates. These decisions are key for the effectiveness of the stress tests and the reliability of their results. When pressures from the industry (especially in the case of bottom-up tests) or political economy or other considerations unduly influence these decisions, stress tests can do more harm than good. Deriving benign conclusions from stress tests that assume modest shocks, include optimistic projections of future income, and use trivial hurdle rates can lead to complacency that may prove very costly in the event of a crisis. Disseminating publicly such results could undermine the credibility of the exercise and convey the impression that the authorities are in denial, or are not prepared to take the measures that may be required to shore up the system.

67. **Ultimately, however, the success of stress tests cannot be reduced to the choice of a few parameters but should be seen in the broader context outlined by the proposed principles.** Stress tests are complex exercises with many “moving parts.” Their effectiveness does not depend on just a few parameters, however critical these may be, or on the degree of public disclosure of their results, but also on the context within which they are conducted. This context includes a clear understanding of the stress tests' scope and objectives; knowledge of the key individual financial institutions in the system, their business models, and main channels of risk transmission; appropriate decisions on the tests' perimeter and coverage; 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.

68. **Adherence to these principles in practice is uneven.** A survey of central banks and supervisory authorities in 23 countries and stress tests in FSAPs shows that despite major improvements since the crisis, practices still fall short of these principles. Shortcomings are particularly notable in three areas: identifying the channels of risk propagation, using the investors' viewpoint, and focusing on tail risks. They reflect both gaps in the analytical toolkit and weaknesses in implementation. Appreciating the implications of these gaps is crucial for the proper interpretation of the results of stress tests. And needless to say, closing these gaps is the key priority for the community of stress testing practitioners today.

69. **The table below summarizes key operational implications that flow from each principle.** For principles 1–3, the recommendations focus on the preparatory work that needs to be conducted by the stress tester to conduct tests with an adequate coverage of institutions, risks, and channels of risk transmission. For principles 4–7, the recommendations focus on stress test design and communication.

Table 4. Practical Implications of “Best Practice” Principles for Stress Testers

| |
|---|
| <p>Principle 1. <i>Define appropriately the institutional perimeter for the tests</i></p> |
| <ul style="list-style-type: none"> ▪ For system-wide stress tests, if universal coverage is not an option, identify systemically important institutions to cover in the tests, including relevant nonbanks and financial market infrastructures. Criteria for assessing systemic importance should include an assessment of interconnectedness, for example through the use of a network model. ▪ Gain a basic understanding of the structure of financial conglomerates, if present, and cover in the tests any of their banking or non-banking activities that may have a major impact in a stress scenario. |
| <p>Principle 2. <i>Identify all relevant channels of risk propagation</i></p> |
| <ul style="list-style-type: none"> ▪ Identify and understand the main channels of risk propagation, including but not limited to the relation between solvency and the cost of funding for institutions with large wholesale funding; the risk transfers between banks and the sovereign; potential bank reactions to stress; and feedback effects between financial sector stress and the macroeconomy (e.g., through deleveraging and lower GDP growth). ▪ Incorporate as many of these channels in the design of the stress tests as possible. If modeling them explicitly is not feasible, use heuristic means, such as a second-round top-down test that maps the impact of first-round tests on individual institutions onto key macro variables (asset prices, credit, aggregate liquidity, etc.); or expert judgment, to capture all potential channels of risk transmission. |
| <p>Principle 3. <i>Include all material risks and buffers</i></p> |
| <ul style="list-style-type: none"> ▪ Understand the tested institution’s—or, for system-wide tests, the systemically important institutions’—business model, including markets where they operate and sectoral or cross-border exposures. ▪ Be as comprehensive as possible in including potential sources of risk in the tests. Do not omit risks because they may be “too big to mitigate.” ▪ Assess whether any banks may be exposed to risks coming from cross-border exposures, non-banking activities (such as employee pension funds, investment funds) or other institutions that may be legally separate but with significant economic links (such as parent companies, subsidiaries, and off-balance sheet vehicles), and include as many of these as possible in the tests. ▪ If stress tests include nonbanking institutions, consider relevant non-economic risks that may have a major impact on them and the rest of the system (such as natural disasters for insurance companies). ▪ In stress tests with longer time horizon, model explicitly likely buffers, such as pre-impairment income. Pre-impairment income should be scenario-dependent and conservatively projected. |
| <p>Principle 4. <i>Make use of the investors’ viewpoint in the design of stress tests</i></p> |
| <ul style="list-style-type: none"> ▪ Adopt mark-to-market methodologies to the valuation of bank assets and liabilities and point-in-time parameters to measure expected and unexpected losses. ▪ Adjust institutions’ initial capital for supervisory practices or other factors that result in systematic over-reporting of capital, such as underprovisioning, collateral overvaluation, and forbearance. Use economic capital estimates as a complement to statutory capital in the stress tests. ▪ Stress market risk appetite by including explicitly shocks involving increases in the market price of risk. ▪ Use hurdle rates reflecting market views (e.g., target funding costs) in addition to regulatory minima. |

Principle 5. *Focus on tail risks*

- Use a variety of approaches to determine “extreme but plausible” shocks to be used in stress tests, including calibration on historical data, cross-country experience, “worst-ever” events, etc.
- Whenever possible, adopt methodologies that capture tail events and risk dependencies. Ideally, use both balance sheet-based models and market price-based models, as the latter can incorporate more easily measures of risk dependencies using market prices. If using balance sheet-based models alone, supplement stress tests of individual financial institutions with tools that measure spillovers and joint default probabilities, such as network models.

Principle 6. *When communicating stress test results, speak smarter, not just louder*

- Tailor the communication of stress test assumptions, methodologies, and results to circumstances and the goals of the tests. Disclosure can yield substantial benefits for macroprudential stress tests, especially in crisis situations, when market confidence is at a premium.
- Before disclosing stress test results, ensure that the tests cover relevant risks and transmission channels, assume serious shocks, set appropriate hurdle rates, produce a candid assessment, and are accompanied by a credible framework of follow-up action, including government support, if needed.
- To the extent possible, set the communication of stress test results in the context of a broader communication strategy for financial stability objectives and policies, and complement them by other financial soundness indicators.

Principle 7. *Beware of the “black swan”*

- Supplement statistical approaches to the identification of potential tail events with expert judgment and new information. Consider analyzing the sensitivity of results to variations in the assumed shock, instead of choosing just one adverse scenario.
- Supplement stress tests with other assessment techniques, such as qualitative and quantitative bank risk analysis, early warning indicators, models of debt sustainability, and informed dialogue with supervisors and market participants, among others. Final conclusions on the resilience of an institution or a system should be based on all these sources, not just the results of stress tests.
- Evaluate and regularly re-assess the institutional coverage, risks, and channels of risk transmission included in the tests; as well as the models and other approaches. Constantly assess the suitability of new tools. Remain vigilant for the emergence of new risks.

APPENDIX I. EXISTING SUPERVISORY GUIDELINES FOR STRESS TESTING BY BANKS³²

70. **In the Basel II framework, banks are asked to apply rigorous stress testing programs.** Pillar 1 focuses on minimum capital requirements for market risk, credit risk, and operational risk. The Capital Requirements Directive (CRD) that introduced Basel II in the EU requires banks that apply the Internal Ratings-based approach to conduct stress tests under Pillar 1 in order to determine the regulatory capital for market risk. Internal ratings-based (IRB) banks are further asked to run credit risk stress tests in order to examine the robustness of their internal approach. However, banks are not only exposed to the three risks covered under Pillar 1: they are also exposed, for example, to securitization, concentration, liquidity, business, and residual credit risks. The Supervisory Review Process, Pillar 2 of the framework, requires banks to take a forward-looking and more comprehensive view on risk, which in turn should determine both capital and strategic planning. Since banks are relatively free in choosing their approach, supervisors should make regular and comprehensive assessments of the processes, strategies, mechanisms, and systems that banks integrate in their Internal Capital Adequacy Assessment Process (ICAAP). The Supervisory Review and Evaluation Process (SREP) by supervisors is supposed to review and evaluate the banks' internal stress testing framework according to the principle of proportionality. Supervisors should consider banks' results to assess capital adequacy and, if necessary, require additional capital and liquidity buffers.

71. **The Basel Committee and other supervisors have issued guidance on the implementation of stress tests by banks.** According to these guidelines, stress testing should be an integral part of corporate governance and the risk management culture of a bank ([CEBS, 2010](#)). While the board is ultimately responsible for the stress testing program in general, senior management is accountable for the implementation and management of the entire system. Senior management is required to identify and communicate the level of risk appetite according to the bank's business model, to identify relevant and plausible stress scenarios that are tested on a firm-wide level, and to ensure that results feed into the bank's decision-making process, including strategic business planning and capital and liquidity planning. Stress scenarios should reflect bank-specific risks, take into account system-wide interactions, and be flexible enough to adapt to changes in portfolio composition, the emergence of new risks, and specific risks related to businesses, entities, and products ([BCBS, 2009](#)). Also, scenarios should feature a wide range of alternatives, from optimistic forecasts (baseline) to tail events that challenge the bank's business model (reverse stress tests).

72. **Nevertheless, these guidelines were not followed systematically.** According to [BCBS \(2009\)](#), risk managers were not able to communicate the purpose, results, and implications of their assessments to the risk-takers within the banks. In many cases, stress testing had been a very technical exercise performed in a rather isolated and mechanical manner not sufficiently interacting with business units. Consequently, the results were often interpreted as unrealistic, not

³² Prepared by Emanuel Kopp.

credible, or just too technical. In many cases, these tests were performed separately for different units, disregarding the requirement of having a comprehensive, firm-wide perspective across different units and risks. As a consequence, risks were largely underestimated (BCBS, [2009](#)). Moreover, [BCBS \(2012\)](#) argued that the stress scenarios analyzed prior to the outbreak of the crisis were most often based on pre-crisis data that did not cover heavy downturn scenarios and did not take into consideration the possibility of simultaneous realization of several risks. This, in turn, caused a systematic underestimation of the relevance and consequences of shocks. The Committee further argued that banks failed to integrate guidelines on reputational risk and risks arising from off-balance sheet vehicles, as well as the risks arising from highly leveraged counterparties and deficiencies in risk mitigating techniques.

73. **The supervisors share part of the blame for these deficiencies.** Based on comprehensive assessments within the SREP, banks should have been required to take corrective action, where necessary. A constructive and open dialogue with other national and international public institutions would have helped in identifying systemic vulnerabilities that individual banks could not uncover (see BCBS, [2009](#) and [2012](#)).

APPENDIX II. STRESS TESTING OF NONBANK FINANCIAL INSTITUTIONS: INSURANCE COMPANIES AND FINANCIAL MARKET INFRASTRUCTURES³³

A. Insurance Companies

74. **Stress testing of the insurance sector is gaining growing acceptance as a risk management tool.** Insurers increasingly use it to assess and manage their risks. Insurance supervisors use it to assess the risks facing specific insurers and to identify possible vulnerabilities of the sector as a whole. From a financial stability perspective, stress testing of insurance companies is also important. If an insurer is significantly involved in activities such as providing protection against credit exposures—which are closely linked to the broader financial sector—then its failure could have systemic implications similar to those of the failure of a large, highly interconnected bank. Similarly, the failure of a large insurance company for which there is no quick substitution in the market, would also make a systemic case.³⁴

75. **In spite of these concerns, insurance stress testing has been performed infrequently in FSAPs.** Insurance stress testing has only been performed in only nine FSAPs between 2003 and 2009, as compared with the banking sector, where stress testing is an integral part of every FSAP.

76. **In FSAP insurance stress tests, efforts have been made to align the economic shocks to those used for the banking stress tests.** However, although economic turbulence is generally considered to be the most likely cause of widespread financial stress in the banking sector, insurers can also be vulnerable to other types of events. For example, natural events, such as floods, earthquakes, and windstorms, can be important to nonlife insurers; and life insurers might be severely affected by a pandemic or by long-term improvements in mortality. The insurance stress testing also incorporate insurance-specific shocks that are relevant to the specific jurisdiction or the insurance sector.

77. **Insurance stress testing tends to rely heavily on bottom-up tests.** This was by far the most common approach in the FSAPs that included stress testing of insurance. One reason for this is that the effects of stresses on the insurance sector are particularly difficult to test on a top-down basis because of the contract-level linkage between assets and liabilities for many life insurance products and the effects of insurer-specific reinsurance programs on the financial condition of nonlife insurers. Another reason is that many insurance supervisors—even in developed markets—do not have the detailed data and models needed to perform such tests. It might be argued that top-down insurance stress testing at a meaningful level of granularity is impossible in most jurisdictions, at least until supervisory data and modeling capabilities have evolved, and that efforts should instead be made to improve the quality of bottom-up stress testing.

³³ Prepared by Christine Sampic and Rodolfo Wehrhahn.

³⁴ Systemic Risk in Insurance—an analysis of insurance and financial stability, special report of the Geneva association systemic risk working Group, March 2010.

78. **The need for stress tests to analyze the vulnerabilities of global insurers has become clear as a result of the recent crisis.**³⁵ Their interconnectedness and cross-border activities are still not well understood and could pose a financial stability risk. The IMF is in the process of developing appropriate stress test methodologies and tools to gain a deeper understanding of the key risks affecting global insurers.

B. Stress Testing of Financial Market Infrastructures (FMIs)³⁶

79. **Stress testing is a tool used to measure Financial Market Infrastructures' resilience to extreme but plausible shocks.** In contrast to what is done for banks, it is not their balance sheet which is tested, but their proper functioning in case a risk materializes. This risk may be of an operational, credit, or liquidity nature. What matters is the immediate reaction of the FMI, the way it will finish the day and be able to operate the next few days following the shock. Both central banks, as FMI operators and overseers, and FMIs themselves conduct stress tests. The Fund helps specify the stress testing requirements embedded in the CPSS/IOSCO standards and checks whether those requirements are met by the FMIs it assesses in the context of FSAPs.

FMI stress testing by central banks

80. **Central banks play a crucial role in ensuring that payments systems are designed and operated in a safe manner, so that they do not compromise financial stability.** They also seek to ensure payment systems operate in a practical and efficient manner for the users and for the economy as a whole. In most cases, the central bank operates the main national large-value payment system, which is the backbone of the entire financial market infrastructure. In addition to this operational role, the central bank is often in charge of overseeing the core payment systems, and often other FMIs, such as central counterparties and securities settlement systems.

81. **As the operator and overseer of the main payment system, some central banks use simulation tools to analyze the underlying payment flows and participant behavior in different scenarios, in addition to a standard-based qualitative approach.** As a result, central banks may propose operational, organizational, or financial changes in the system, such as the implementation of new and more robust risk mitigation facilities, resulting in increased systemic stability. The Bank of Finland has developed a payment system simulator (BoF-PSS2), which is available to other central banks that can customize it, and organizes regular seminars among central banks to share their practical experience in stress testing and define the business requirements of the next version of the simulator. Eurosystem overseers have decided to develop a tailor-made TARGET2-specific simulator based on BoF-PSS2 in order to run quantitative simulation-based stress tests on TARGET2.

³⁵ IMF-IAA Stress Testing Workshop, Washington, DC, September 7–9, 2010.

³⁶ FMIs refer to payments systems, central securities depositories, securities settlement systems, and central counterparties.

82. **Operational disruptions of the FMI itself or a major participant are often tested, as well as the financial default of major participants.** Incidents are simulated in order to identify recovery times, critical participants, contingency measures, and stop-sending limitations under different parameters, such as the concentration level, availability of liquidity, back-up procedures, reactions of non-defaulted banks, and structure of the money market. The simulations can, for example, indicate if the system can complete settlement before the end of the day, as prescribed by CPSS standards, and allow defining the level of the contingency capacity needed. They can also help quantify the impact on liquidity by simulating the suspension of a participant's outgoing payments, which results in liquidity accumulations for the failing participant and liquidity shortages for other participants and thereby disrupts settlements. Financial defaults of major participants are also tested, to check whether the payment system will be able to handle them properly and to assess the impact on remaining participants.

Stress-testing by FMIs

83. **In light of the financial crisis, many FMIs have developed their own financial stress tests as a tool to ensure they can properly manage liquidity and credit risks.** Clearing House Interbank Systems (CHIPS), the U.S. private sector large-value system, has instituted a program to ensure that participants understand the consequences of a failure of one or more banks to honor their closing positions and to encourage participants to develop liquidity contingency plans. Continuous Link Settlements (CLS) Bank, which operates payment versus payment settlement services in 17 currencies and covering more than 60 percent of foreign exchange transactions worldwide, conducts a range of stress and back-testing scenarios to review the adequacy of its risk management procedures. The simulations include *inter alia* the adequacy of haircuts and the failure of the settlement member with the single largest funding obligation in a single currency. Since the recent crisis, CLS Bank has been working on some more extreme scenarios, such as the failure of all settlement members of a given currency to complete their funding and the simultaneous failure of all liquidity providers in the same currency.³⁷ Some securities settlement systems, in particular the International Central Securities Depositories (ICSD),³⁸ whose activities go beyond mere settlement (they are also offering securities loans and credit lines to participants), face large credit and liquidity risks that need to be closely monitored and mitigated and most of them conduct regular stress tests of their financial resources.

84. **Among FMIs, CCPs exhibit the highest concentration of liquidity and credit risks.** Their core service is to become principal to every transaction that they clear, which implies that market participants no longer have credit exposures to their trading counterparties, but only to the CCP. Therefore, CCPs concentrate credit risk and would face large liquidity needs if a participant defaults, because they need to fulfill the settlement obligations of the defaulting participant,

³⁷ For more detail on CHIPS and CLS Bank's stress test, see United States FSAP documentation—Technical Note on Selected Issues on Liquidity Risk management in Fedwire Funds and Private Sector Payment Systems.

³⁸ The two main ICSDs are Euroclear and Clearstream.

potential losses when the cleared position or related collateral are liquidated, and the cash flows relating to possible hedge transactions.

85. **Stress testing is key in managing the credit and liquidity risks of CCPs.** Stress tests take into account extreme but plausible market conditions and are typically framed in terms of the number of participant defaults a CCP can withstand. Current CPSS/IOSCO standards³⁹ prescribe that CCPs should be able to withstand the default of the participant with the largest exposure, but some CCPs have chosen to be more stringent and test for the default of several major participants. CCPs' models that calculate their margin requirements, default fund contributions, collateral requirements and other risk control mechanisms are expected to be subjected to rigorous and frequent stress tests that reflect their product mix and other risk management choices. Key elements of stress tests are the assumed market conditions and default scenarios and the test frequency. A CCP should assume extreme market conditions (that is, price changes significantly larger than the prevailing levels of volatility), and evaluate the potential losses in individual participants' positions. Other stress tests may consider the distribution of positions between the defaulting participants and their customers in evaluating potential losses. These should take into account the resources of the potential defaulters that are available to a CCP (margins, clearing fund contributions or other assets), as well as the CCP's own resources, to provide perspective on the potential size of the losses and liquidity gaps of the CCP.

86. **Since the financial crisis, regulators, overseers, and supervisors' attention has been drawn to CCPs' stress testing, in particular when clearing OTC derivatives.** Central clearing of OTC derivatives presents more challenges than clearing listed or cash-market products because of their complex risk characteristics. Following the G-20 commitment to strengthen regulation of the OTC derivatives markets, improved rules are to enter into force by end-2012. In some jurisdictions, this will include guidance on stress testing. For example, it is foreseen that the European Commission will adopt technical standards that will specify the type of tests to be undertaken for different classes of financial instruments and portfolios, the involvement of clearing members or other parties in the tests, the frequency of tests, and the time horizon. In addition, CCPs' supervision and oversight have generally been strengthened since the crisis, for example by systematically analyzing CCP's internal risk management models, including stress testing parameters.

FMI stress testing and the Fund

87. **The Fund often examines FMI stress testing arrangements in the context of FSAP and ROSC missions.** The Fund has not so far conducted stress tests on FMIs. Rather, it helps specify the stress testing requirements embedded in the CPSS/IOSCO standards and checks whether those requirements are met by the FMIs it assesses in the context of FSAPs.

³⁹ CPSS/IOSCO Recommendations for central counterparties, BIS, November 2004.

APPENDIX III. RECENT PROMINENT CRISIS MANAGEMENT STRESS TESTS⁴⁰

88. **Since the recent financial crisis, stress tests have increasingly been used as instruments for crisis management.** Their results feed into political decision-making processes, and in many cases determine the allocation of public support measures.

89. **The best known examples for crisis management stress testing exercises were conducted in the United States and Europe.** These were (i) the Federal Reserve’s Supervisory Capital Assessment Program ([SCAP](#)) completed in May 2009; (ii) the European Union-wide stress tests performed by the Committee of European Banking Supervisors (CEBS) in [2010](#) and the European Banking Authority (EBA) in [2011](#); and (iii) the EBA’s 2011–2012 [Capital Exercise](#). In contrast, the Fed’s Comprehensive Capital Assessment Reviews ([CCAR](#)) in 2011 and 2012, as well as tests required under the Basel frameworks, constitute typical supervisory stress tests.

Key Features of Crisis Management Stress Tests

90. **Crisis management stress testing exercises incorporate characteristics of both microprudential and macroprudential or surveillance stress tests.**

- Like in supervisory stress tests, the individual institutions’ resilience to shocks is assessed on a bank-by-bank basis. The tests incorporate systemic risk through macroeconomic and market-level shocks, and evaluate the banks’ performance in light of current or future regulatory requirements, or according to alternative thresholds and definitions for capital ratios specifically designed for the particular exercise.⁴¹ Potential follow-up actions for banks that do not pass the test include requiring recapitalization from the private market and/or the acceptance of governmental support packages, restructuring, and changes in business models.
- At the same time, crisis management stress tests do not limit their focus on assessing the health of individual banks but are equally concerned about the stability of the system as a whole. The shocks result from a common macroeconomic scenario and are of systemic nature. In order to evaluate the resilience of the banking system and the potential need for macroprudential or system-wide measures, aggregate indicators are taken into consideration. National and/or international macroprudential authorities are often involved in these exercises, in addition to national supervisors. And the results are usually subject to some form of publication, but with varying degrees of granularity and transparency.

⁴⁰ Prepared by Emanuel Kopp.

⁴¹ For instance, in the 2011 EU-wide stress test, the EBA focused on banks’ core capital and considered a specific Core Tier 1 capital definition that was based on the Capital Requirements Directive (CRD) II definition of Tier 1 capital net of deductions for participations in financial institutions, excluding hybrid instruments, but including existing preference shares and existing governmental support measures. This definition should not be confused with the Basel III Common Equity Tier 1 definition, according to CRD IV. In the 2009 SCAP, the Fed applied a modified definition of Tier 1 capital that excluded preferred stock, minority interest in subsidiaries, and less qualifying trust preferred securities.

91. **Crisis management exercises are tailored to examine current risks.** In contrast to both surveillance and supervisory stress testing, crisis management stress tests are conducted for specific (crisis management) purposes and do not necessarily take place on a regular basis. The shock scenarios usually involve higher-probability shocks that are more likely to materialize than the extreme-but-plausible tail risks typically tested in surveillance and supervisory tests. Moreover, the scenarios are specifically tailored to current risks. The exercise is comprehensive, covering both the banking and trading book, and usually taking into account off-balance sheet positions. Portfolios are broken down by region and sector, which in turn implies that potential losses can differ across banks not only because of their different portfolio composition, but also reflecting differences in asset quality within specific geographical regions and asset classes.

92. **The methodology of crisis management stress tests combines elements of the bottom-up and the top-down approach.**⁴² According to a detailed methodology designed by the supervisor, banks examine the impact of consistent, common macroeconomic scenarios on their portfolios. The results are checked for completeness, consistency, and plausibility by the supervisor. But the tests also include centralized components.⁴³ The idea is to make use of the particular advantages of top-down and bottom-up approaches and, at the same time, to cross check or validate the banks' results. A key objective is to ensure quality control, not least because banks have incentives to represent their results in the best possible light.⁴⁴

93. **Crisis management stress tests are mainly focused on solvency, are designed as traditional balance sheet-based stress tests, and apply some form of static balance sheet assumption.** The SCAP, EU-wide stress testing exercises, and the EBA Capital Exercise focused on solvency risk in individual banks: liquidity risk was either not tested ([SCAP](#)), or assessed in a separate exercise ([EBA 2011](#)).⁴⁵ While the frameworks did not directly account for spillover effects and default dependencies across institutions, these properties were, to a certain extent, indirectly considered through the design and structure of adverse scenarios that typically take into account a number of propagation channels. These exercises applied a form of constant balance sheet assumption in order not to allow banks to shrink balance sheets or adapt business models

⁴² As banks are in fact constrained by the supervisors' methodology in assessing the impact of the scenarios on their institution, the EBA has coined the term "constrained bottom-up" tests.

⁴³ The 2009 [SCAP](#) basically followed a top-down approach and, at the same time, incorporated a number of decentralized components of bottom-up frameworks, like the integration of banks' own projections of losses, operating profits, and loan loss provisions under the given scenario. According to the detailed methodology designed by the supervisor, the CEBS/EBA tests asked banks to examine the impact of two scenarios on their portfolios. The results were checked and challenged internally by [CEBS \(2010\)](#) and within a multilateral review process that was also flanked by top-down calculations by the European Systemic Risk Board in the [2011 exercise](#).

⁴⁴ This is most crucial for pre-impairment income, which serves as a first, and substantial, cushion against losses. Since banks have a better understanding of their business conditions, they are in a better position when discussing their bottom-up forecasts with the supervisor. While the Fed applied top-down stress tests in order to challenge the banks' submissions, the EBA chose to cap net interest income at 2010 levels in its [2011 EU-wide test](#).

⁴⁵ The [2011 EU-wide stress test](#), however, did consider stress on banks' funding costs, i.e. funding liquidity, within the solvency stress testing framework. A traditional liquidity stress test assessing banks' liquidity profiles was performed separately. The results of this assessment were not published.

during the forecasting period. Such approaches eliminate the scope for strategies to boost capital ratios by reducing risk-weighted assets. At the same time, this comes at the price of disregarding behavioral response functions.

Comparing the Recent Crisis Management Stress Testing Exercises

94. **The three most prominent crisis management exercises had different purposes.** The SCAP was designed to “estimate losses, revenues, and reserve needs” for the large bank holding companies, and evaluated the size of governmental capital injections contingent on the banks’ performance in the stress test (Federal Reserve, [2009](#)). The EU-wide stress testing exercises aimed at examining the resilience of the banking sector, improving transparency, identifying vulnerabilities, and informing policy-makers about the current capacity of banks to absorb shocks and the banking system’s dependence on public support measures ([EBA 2011](#)). EBA’s Capital Exercise was specifically designed to “create an exceptional and temporary capital buffer to address current market concerns over sovereign risk and other residual credit risk related to the current difficult market environment” ([EBA 2011](#)).⁴⁶

95. **The design and severity of macroeconomic stress scenarios depended on the objectives and the timing of each exercise.** The adverse scenarios in both the SCAP and the EU-wide stress tests involved a simultaneous realization of several risk factors, as in a typical macroeconomic scenario-based stress test. EBA’s Capital Exercise, in contrast, was based on a EU-wide baseline scenario,⁴⁷ and banks were required to conservatively assess the value of European Economic Area sovereign debt exposures held in the banking book⁴⁸ according to market prices as of September 2011 and current bond yields by maturity for loans and non-traded assets. This exercise therefore specifically focused on the risks stemming from European sovereign debt markets. The adverse scenario of the 2011 (2010) European-wide stress test translated into a -4.1 percentage points (-3.1 percentage points) cumulative GDP shock, compared to a -2.8 percentage points shock in the SCAP. Since the SCAP was organized at the peak of the crisis, however, the assumed shock was equivalent to a deeper contraction compared to the European tests (a cumulative two-year contraction in GDP of -2.7 percent under the SCAP, compared to a cumulative contraction of -0.2 percent in the 2011 EBA exercise and -0.4 percent in the 2010 exercise).

96. **Another key difference between these exercises was the organizational complexity of the EU-wide tests.** The EU exercises involved 21 countries, 24 national supervisors and authorities, around a dozen European institutions, and 91 participating banks (Figure 2). They had

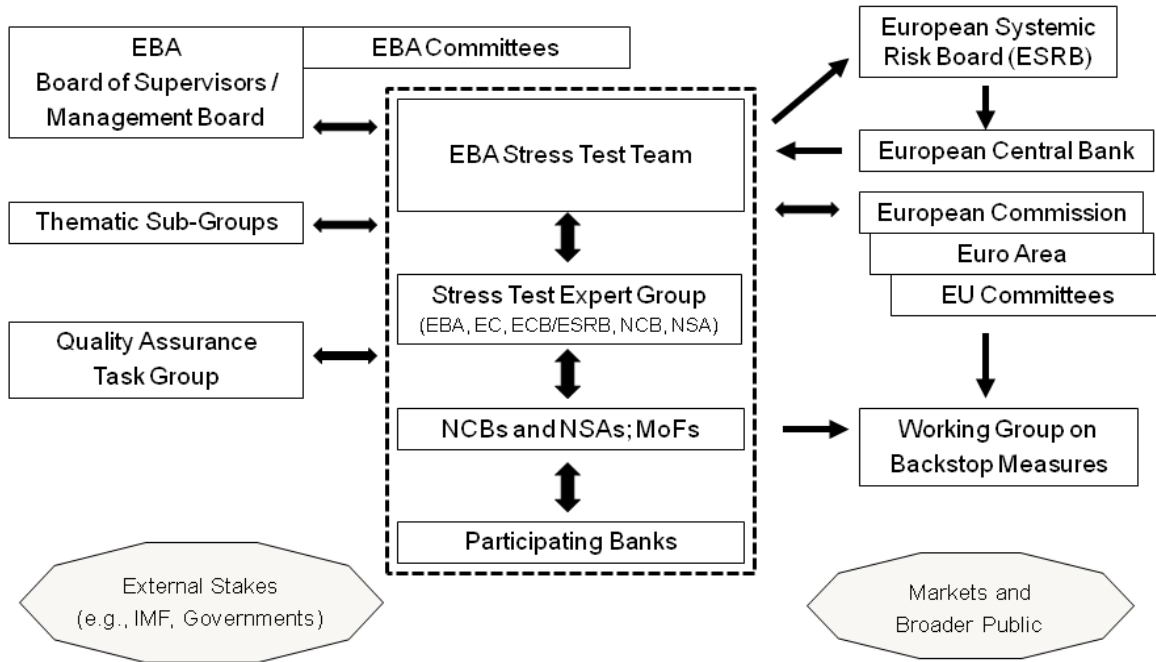
⁴⁶ EBA’s assessment was the first international crisis management exercise that, to some extent, considered the investors’ perspective, including for sovereign risk.

⁴⁷ Autumn 2010 European Commission Forecast.

⁴⁸ In the held-to-maturity and loans-and-receivables portfolios.

to deal with 19 different languages and seven currencies. None of these complexities applied to the SCAP.

Figure 2. Institutional Setup of EBA Stress Tests



97. **The experience with these exercises was mixed.** The goal of the SCAP was to “[e]nsure adequate system capital to promote lending and restore investor confidence” (Federal Reserve, 2009). The goal of the 2011 EU-wide stress test was to assess the resilience of the banking system and to restore confidence in EU banks and the European banking system as a whole (EBA, 2011). But while the goals were (almost) identical, the impact of the two exercises was not.⁴⁹

98. **The SCAP demonstrated convincingly that stress testing can be a powerful instrument for crisis management.** The setup of the SCAP was credible, and backstop measures (including, crucially, government support) were clearly communicated and in place ahead of time. The Federal Reserve has argued that the test was “an important turning point in the financial crisis” and that confidence improved as banks raised capital, mainly in the private markets. The results allowed markets to differentiate between banks based on their ability to withstand the

⁴⁹ It would be premature to evaluate the success of the 2011–2012 EBA Capital Assessment exercise as its results have not yet been published and it is not possible to judge its full impact.

shocks: after publication of the results, the correlation among major banks' stock prices fell by more than 10 percent. Capital injections for banks not passing the tests were mandatory. Banks unable to raise capital in the private markets were asked to accept governmental capital injections. Consequently, both goals of the SCAP were achieved: confidence in the U.S. banking sector was restored and capital adequacy ensured.

99. **The 2011 EU-wide stress test, on the other hand, although in many aspects a successful exercise, did not fully achieve its goals.** The 2011 EU-wide stress tests were widely praised for their risk coverage, tighter definition of capital components, quality assurance process, and detailed disclosure of exposures and results. The fact that the EBA included stress on funding costs, sovereign exposures, and securitization positions expanded the test's risk coverage substantially. The tougher capital definition was welcomed, too, but was also criticized for being inconsistent with the Basel III Common Equity Tier 1 definition. The comprehensive publication of methodology and results made the test "an important exercise in disclosure",⁵⁰ allowing analysts to duplicate the results, and is seen today as the most important part of the exercise. The whole exercise, however, was seen as flawed in some key respects. *First*, it was criticized because of the way sovereign exposures were stressed and for the severity of stress scenarios in general: while sovereign risk was, in principle, covered, the stress applied on sovereign exposures did not reflect the strong tensions in European sovereign debt markets.⁵¹ *Second*, the adverse scenario was designed as a common scenario for all jurisdictions, but this meant that it did not adequately reflect some national banking sectors' risk profiles and, at the end, was not sufficiently credible. Cases like Dexia in 2011 and the Irish banks in the 2010 exercise further undermined the credibility of the results. And *third*, backstops that were seen by the markets as ambiguous, along with uncertainty over actions for banks ending up slightly above the hurdle rate, considerably undermined attempts to restore investor confidence.

⁵⁰ The Economist, July 18, 2011.

⁵¹ "The banking component can no longer be separated from sovereign and institutional developments. This is why Friday's publication of stress tests results, while useful, is unlikely to be the game-changer it could have been two years ago." (Financial Times, July 14, 2010)

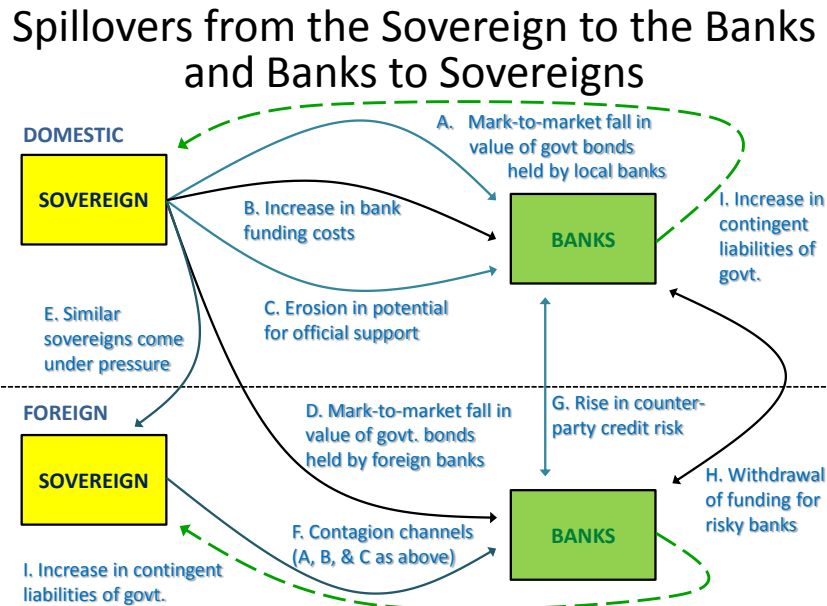
APPENDIX IV. FRONTIERS OF STRESS TESTING⁵²

100. **This Appendix presents current conceptual challenges for stress testing**, especially in light of the global financial crisis and ongoing European crisis, and various analytical approaches that might be used to address them, including work being done at the Fund.

Integrating Sovereign Risk and Banking-Sovereign Feedbacks in Stress Tests

101. **Conceptual and technical difficulties related to how to account for sovereign risk exposures in bank stress testing became evident in the [EBA 2011 stress tests](#)**. These tests highlighted the need to properly account for the credit risk associated with sovereign bonds held on bank balance sheets. However, the bank and sovereign interactions are complex: what is needed is models that account for the banks' impact on sovereign risk *and* feedbacks from sovereigns to banks in stress testing. There are numerous channels of interaction between the sovereign and the banks as shown in Figure 3. The mark-to-market fall in the value of sovereign bonds held by banks reduces bank asset values, and distressed banks lead to an increase in government contingent liabilities. High sovereign spreads lead to increased bank funding costs. If the sovereign is distressed enough, the value of official support (guarantees) to bank is eroded. This can have additional contagion effects to foreign banks and sovereigns, as shown in Figure 3.

Figure 3. Sovereign and Bank Interactions



Source: [IMF \(2010b\)](#), GFSR

⁵² Prepared by Dale Gray.

102. **To include sovereign risk in stress testing, the key interlinked risk exposures between the government and financial sector should be analyzed in a comprehensive framework.** A stylized framework starts with the economic, i.e., risk-adjusted, balance sheets of the financial sector (portfolio of financial institutions) and is then linked to, and interacts with, the government's balance sheet. For example, distressed financial institutions can lead to large government contingent liabilities, which in turn reduce government assets and lead to higher risk of default on sovereign debt. Dynamic macrofinancial linkage models used for bank stress tests can also be linked to sovereign risk models, together with the feedback of banking risk to sovereigns via contingent liabilities and sovereign spreads affecting bank funding costs.

Enhancing Analysis of Macroeconomic and Banking Sector Feedbacks and Contagion between Financial Institutions

103. **A typical stress testing exercise uses macroeconomic scenarios to assess the impact on the banking/financial sector risk and capital adequacy without feedbacks to the macroeconomy.** One of the lessons from the recent crises is that, in many cases, the banking/financial sector becomes distressed first and credit supply contracts, leading to lower GDP growth only afterwards. These feedbacks are not usually included in stress tests. Going forward, it would be useful to build stress test models that have this feedback channel explicitly incorporated; for example, a shock to the financial sector might be used to estimate the reduction in GDP growth and other factors, which would in turn have negative impacts on corporate and household borrowers, which would in turn increase credit risk on banks' balance sheets. Dynamic factor models can help enhance the modeling of the feedback between the banks and the macroeconomy.

104. **The recent crisis has shown how strong contagion can be among financial institutions.** Traditional stress tests that use macro factor models to link to banking risk have a certain built-in correlation between banks, which comes from their common correlations to the macro factors. However, this does not capture correlations, dependencies, and feedbacks between the institutions. Contagion effects can be modeled with networks, joint probabilities of default, systemic CCA, co-risk indicators, and a variety of other models that have been developed recently. Enhanced stress testing can include some of the features of these systemic risk models to improve the analysis of interdependence and joint risk.

Techniques to Improve Modeling of Non-Linearities in Changes in Bank Assets, Capital, Credit Risk and Funding Cost

105. **The risk-adjusted balance sheet of a bank can be helpful in illustrating the tradeoffs between changes in bank's assets, (risky) debt, credit/funding spreads, and bank equity.** The fundamental conceptual framework of the risk-adjusted balance sheet comes from the contingent claims approach, CCA (Merton, 1973) and risk-neutral valuation (Cox and Ross, 1976). A bank's liabilities (equity and (risky) debt) are claims on underlying bank assets that are uncertain over this time horizon, and the degree of uncertainty (i.e., volatility) affects the risk premiums and

values of equity and debt liabilities. There are different ways to construct the risk-adjusted balance sheet. One can use the estimated loan portfolio loss distribution and other components of the bank's balance sheet. Using risk-neutral valuation, the probability distribution of the bank's risky loans and distribution of assets (over a specific time horizon) can be estimated. This asset distribution is then combined with the promised payments on debt and deposits to construct a risk-adjusted (CCA-type) balance sheet. This technique does not rely on market prices. A second method estimates bank implied asset level and asset distribution from the observed market value and volatility of the bank's equity and its book value of debt and deposits. Comparing the risk indicators and tail risk shapes from the two methods can provide insights into the dynamics and differences between the "fundamental" loan portfolio loss approach and the market implied view (which will vary between calm and stress periods).

106. **Stress tests can be enhanced by including the impact on banks' funding costs.** Higher bank funding costs leads to higher lending rates for corporates and households, credit rationing, and lower credit growth. This can have a negative impact on economic output, which can in turn feed back, causing further distress in the banking system. Funding costs are dependent on the risk-free interest rate, the bank's credit spread, market risk appetite, and the impact of the government's (implicit and explicit) guarantees. The risk-adjusted balance sheet models are useful tools for calculating bank funding costs.

107. **More work is needed on capturing the non-linearities between changes in bank assets and capital.** In balance sheet-based stress testing models, fixed correlations between exposures and static risk weights can lead to underestimation of the capital shortfall under stress. This can be improved by using time-varying correlation between exposures when estimating portfolio loss distributions. In market price-based models, such as the contingent claims model, the dynamic changes between assets and market capital are already built in. Changes in assets lead to changes in market equity and changes in expected losses to bank creditors or guarantors. The magnitude of these losses depends on the level of distress of the bank. The change in bank market capitalization given a change in asset values is analogous to an aggregate RWA adjustment factor in the Basel rules: it can thus be seen as a sort of "quasi-RWA" dynamic adjustment factor.

REFERENCES

- Acharya, V. V., Pedersen, L. H., Philippon, T, and Richardson, M. P., 2010, “Measuring Systemic Risk”, in *Regulating Wall Street: The Dodd-Frank Act and the New Architecture of Global Finance*, Acharya, V. V., Cooley, T., Richardson, M, and Walter, I. (Eds.), John Wiley & Sons.
- Adrian, Tobias and Markus Brunnermeier, 2008, “[CoVaR](#),” Staff Reports 348, Federal Reserve Bank of New York.
- Aikman, David, Piergiorgio Alessandri, Bruno Eklund, Prasanna Gai, Sujit Kapadia, Elizabeth Martin, Nada Mora, Gabriel Sterne and Matthew Willison, 2009, “Funding Liquidity Risk in a Quantitative Model of Systemic Liquidity,” Bank of England Working Paper, June, No. 372.
- Allen, Franklin., and Douglas. Gale, 2000, “Financial Contagion,” *Journal of Political Economy* No. 108 (1), pp. 1–33.
- Avesani, Renzo, Kexue Liu, Alin Mirestean and Jean Salvati, 2008, “Review and Implementation of Credit Risk Models of the Financial Sector Assessment Program,” IMF Working Paper No. 06/134.
- Bank of International Settlements, 2011, “[Sovereign risk in bank regulation and supervision: where do we stand?](#)” Speech by BIS Deputy General Manager Herve Hannoun.
- Barnhill, Theodore, Panagiotis Papapanagiotou, and Liliana Schumacher, 2002, “Measuring Integrated Credit and Market Risks in Bank Portfolios: An Application to a Set of Hypothetical Banks Operating in South Africa,” *Journal of Financial Markets, Institutions and Instruments*, 11(5) pp. 401–443.
- Barnhill, Theodore and Liliana Schumacher, 2011, “Modeling Correlated Systemic Liquidity and Solvency Risks in a Financial Environment with Incomplete Information,” IMF Working Paper No. 11/263.
- Basel Committee on Banking Supervision, 2004, “[Principles for the management and Supervision of Interest Rate Risk](#).”
- _____, 2009, “[Principles for sound stress testing practices and supervision](#)”, May.
- _____, 2011, “[Global systemically important banks: Assessment methodology and the additional loss absorbency requirement](#)”, November.
- _____, 2012, “[Peer review of supervisory authorities’ implementation of stress testing principles](#)”, April.
- Board of Governors of the Federal Reserve System, 2012, United States Comprehensive Capital Analysis and Review 2012: Methodology and Results for Stress Scenario Projections (CCAR).
- _____, 2009, [United States Supervisory Capital Assessment Program \(SCAP\)](#).
- Borio, Claudio, Mathias Drehman, and Kostas Tsatsaronis, 2012, “[Stress-testing Macro Stress Testing: Does it Live up to Expectations?](#)” BIS Working Paper No. 369.

- Committee of European Banking Supervisors, 2010, "[CEBS Guidelines on Stress Testing](#)."
- Čihák, Martin, 2007, "[Introduction to Applied Stress Testing](#)," IMF Working Paper No. 07/59.
- _____, and Li Lian Ong, 2010, "[Of Runes and Sagas: Perspectives on Liquidity Stress Testing Using an Iceland Example](#)," IMF Working Paper No. 10/156.
- _____, Sonia Muñoz and Ryan Scuzzarella, 2011, "[The Bright and the Dark Side of Cross-Border Banking Linkages](#)," IMF Working Paper No. 10/105.
- Cox, J. and S. Ross, 1976, "The Valuation of Options for Alternative Stochastic Processes" *Journal of Financial Economics*, No. 3 (Jan–Mar), pp.144-166.
- European Banking Authority, 2011, [EU-wide Stress Testing. Aggregate Report](#).
- Espinosa-Vega Marco and Juan Sole, 2010, "[Cross Border Financial Surveillance: A Network Perspective](#)," IMF Working Paper No. 10/105.
- Espinoza, Raphael and Miguel Segoviano, 2011, "[Probabilities of Default and the Market Price of Risk in a Distressed Economy](#)," IMF Working Paper No. 11/75.
- Gray, Dale F., 2008, "A New Framework for Measuring and Managing Macrofinancial Risk and Financial Stability," Harvard Business School Working Paper No. 09-015.
- _____, and Andreas A. Jobst, 2010, "[United States: Technical Note on Stress Testing](#)," IMF Country Report No 10/244, Section IV.
- Greenlaw, David, Anil Kashyap, Kermit Schoenholtz, and Hyun Song Shin, 2012, "[Stressed Out: Macroprudential Principles for Stress Testing](#)," Chicago Booth Paper No. 12-08.
- Haldane, Andrew, 2009a, "*Why banks failed the stress test*," Speech at Marcus-Evans Conference on Stress-Testing, London, February 9–10, 2009.
- _____, 2009b, "*Small Lessons from a Big Crisis*," Remarks at the Federal Reserve Bank of Chicago 45th Annual Conference "Reforming Financial Regulation," May 8, 2009.
- _____, Simon Hall, and Silvia Pezzini, 2007, "A New Approach to Assessing Financial Stability," Bank of England Financial Stability Paper No. 2.
- Hardy, Daniel, and Christian Schmieder, forthcoming, "Rules of Thumb for Solvency Stress Tests with a Global Case Study," IMF Working Paper.
- International Monetary Fund, 2006, [Report on the Evaluation of the Financial Sector Assessment Program](#), Independent Evaluation Office of the International Monetary Fund.

- _____, 2010a, [“Integrating Stability Assessments under the Financial Sector Assessment Program into Article IV Surveillance: Background Material,”](#)
- _____, 2010b, [Sovereigns, Funding and Systemic Liquidity](#), Global Financial Stability Report, October Chapter 1.
- _____, 2011a, [“How to Address the Systemic Part of Liquidity Risk,”](#) Global Financial Stability Report, April, Chapter 2.
- _____, 2011b, [“Toward Operationalizing Macroprudential Policies: When to Act?”](#) Global Financial Stability Report, Chapter 3, September.
- Ishikawa, Atsushi, Kichiro Kamada, Yoshiyuki Kurachi, Kentaro Nasu, Yuki Teranishi, 2012, [“Introduction to the Financial Macro-econometric Model,”](#) Bank of Japan Working Paper Series, No. 12-E-1.
- JP Morgan (1995). *Creditmetrics*.
- Le Lesle, V. and S. Avramova, 2012, [“Revising Risk-Weighted Assets: Why do RWAs Differ Across Countries and What Can be Done About It?”](#) IMF Working Paper No. 12/90.
- Merton, R. C. 1973, “Theory of Rational Option Pricing,” *Bell Journal of Economics and Management Science*, Vol.4 (Spring), pp. 141–83.
- Office of Financial Research, 2012, [Annual Report](#), United States Treasury, Washington, DC.
- Rosch, Daniel and Harald Scheule, 2008, *Stress Testing for Financial Institutions: Applications, Regulations, and Techniques*, Risk Books, Incisive Media, London.
- Schmieder, Christian, Claus Puhr, and Maher Hasan, 2011, [“Next Generation Balance Sheet Stress Testing,”](#) IMF Working Paper No. 11/83.
- _____, Hesse, Heiko, Neudorfer, Benjamin, Puhr, Claus, and Stefan W. Schmitz, forthcoming, [“Next Generation System-Wide Liquidity Stress Testing,”](#) IMF Working Paper No. 12/03.
- Segoviano, Miguel, and Charles Goodhart, 2009, [“Banking Stability Measures,”](#) IMF Working Paper No. 09/04.
- _____, 2006, “The Consistent Information Multivariate Density Optimizing Methodology,” Financial Markets Group, London School of Economics, Discussion Paper No. 557.
- Tressel, Thierry, 2010, [“Financial Contagion through Bank Deleveraging: Stylized Facts and Simulations applied to the Financial Crisis,”](#) IMF Working Paper No. 10/236.