CHAPTER 1

Stress Testing at the International Monetary Fund: Principles, Concepts, and Frameworks

ANDREAS A. JOBST • LI LIAN ONG

1. MOTIVATION

Stress testing at the IMF has evolved into an integral aspect of financial sector surveillance over the past two decades. It is a key component of the Financial Sector Assessment Program (FSAP) and was used during the global financial crisis (GFC) to support estimations of banking system recapitalization needs of countries with the IMF’s crisis programs. Stress testing has also become an important forward-looking risk-management tool for financial supervisors and macro-prudential authorities to identify vulnerabilities of individual banks and financial systems to the impact of adverse changes to the operational and market environment.

The intensified interest in stress testing has underscored the need for developing and communicating a coherent and consistent approach for such exercises as part of the IMF’s financial surveillance. The first book, A Guide to IMF Stress Testing: Methods and Models (which was published in 2014), focuses predominantly on the stress testing models developed by the IMF staff for financial stability analysis and crisis management. It categorizes relevant analytical tools and methods into three distinct but complementary approaches (that is, accounting-based, market-price-based, and macro-financial). It also showcases continuing efforts by the IMF staff to apply more encompassing risk measures and assessment methods and summarizes their advantages and disadvantages in recommending their appropriate application.

The development of comprehensive stress testing policies has facilitated better harmonization and comparability of approaches, methods, and models. This anthology complements the first book by compiling selected papers written by several IMF staff members (and their external coauthors) on principles, concepts, and frameworks of stress testing. They illustrate the anatomy of IMF stress testing (Figure 1.1) and discuss several areas where the IMF staff has made progress, notably: (1) the design of and model selection for stress tests; (2) key gaps in the coverage of risks and information, which were highlighted during the GFC; (3) practical approaches to implementing important aspects of stress testing; and (4) widening the scope of financial activities beyond banks, namely, nonbank financial institutions and financial market infrastructures (FMIs).

That said, this book does not claim to be comprehensive, and much remains to be done on the stress testing front as an active area of work for the IMF. It does not capture the work-in-progress owing to the timing of its production, such as the IMF staff’s work on ongoing financial stability assessments and related implications of the current COVID-19 pandemic as well as the 2020 Review of the FSAP.1 The empirical analysis underlying these chapters has not been substantially updated since the issuance of the respective working papers. However, references have been updated throughout to reflect important regulatory and research developments. In all, the chapters in this volume provide helpful insights into the IMF staff’s evolving thinking on the principles, concepts, and frameworks of stress testing over the past eight years.

2. ANATOMY OF STRESS TESTING

Essential Building Blocks

“Best practice” principles provide a baseline against which individual stress testing exercises may be assessed. In Chapter 2, Schumacher and Oura (with others) discuss the

---

1 More recent studies, such as those by Krznar and Matheson (2017) and Bouveret (2018), have not been included.

©International Monetary Fund. Not for Redistribution
Financial auxiliaries are financial corporations that are principally engaged in activities associated with transactions in financial assets and liabilities or with providing the regulatory context for these transactions, which do not involve the auxiliary taking ownership of the financial assets and liabilities being transacted according to the taxonomy of the Global Monitoring Report on Non-bank Financial Intermediation (FSB 2019).

Other financial intermediaries include captive financial institutions and money lenders, investment funds, trust companies, central counterparties, broker-dealers, and securitization vehicles.

Figure 1.1. Anatomy of Stress Testing: Principles, Concepts, and Frameworks
practical guidelines derived from the IMF staff’s stress testing experience up to 2012 and a survey of stress testing practices among selected national central banks and supervisory authorities conducted in 2011. The authors argue that overarching principles can effectively guide many of the choices in designing and implementing system-wide stress tests. These areas include (1) the coverage of institutions, risks, and scenarios; (2) the specification of a suitable quantitative framework to link various shock scenarios to solvency and liquidity measures; (3) a strategy for communicating the results; and (4) follow-up measures, if warranted. These considerations complement the updated supervisory principles published by the Basel Committee on the use, implementation, and oversight of stress testing supervisory principles published by the Basel Committee in 2013.

The implementation of stress testing principles and concepts for macroprudential surveillance and financial stability analysis is closely intertwined with the development of systemic risk models. In Chapter 3, Demekas discusses the required features of tools that can effectively assess system-wide risks; notably, they would have to (1) incorporate general equilibrium dimensions, and (2) focus on the resilience of the system as a whole. However, while stress testers have made significant progress in the first area (some of which are covered in this book), they have achieved much less in addressing the second.

The effective implementation of stress tests requires frameworks that adequately cover both solvency and liquidity risks. Within the banking sector, the focus had traditionally been on solvency risk, but liquidity risk came to the fore at the onset of the GFC. In Chapters 15 and 16, Jobst, Ong, and Schmieder illustrate the application of stress testing principles and concepts by reviewing key elements of the IMF staff’s solvency and liquidity stress tests and cataloging their actual implementation in FSAPs for jurisdictions with systemically important financial sectors between 2010 and 2016. While a similar solvency stress testing framework may be used for crisis stress tests, Ong and Pazarbasioglu underscore in Chapter 14 that the design of some of those elements is invariably different to ensure the credibility of the exercise at a most critical time.

Another important aspect of stress testing is the simulation of shock scenarios. Banking sector solvency stress tests typically apply adverse—that is, severe but plausible—macro-financial scenarios. These scenarios describe the evolution of various macro-financial variables over a specific horizon, following shocks that move them away from their respective baselines. The shocks may originate domestically or abroad, with the banking sector exposed to them either directly, via determinants of profitability and capitalization, or indirectly through macro-financial linkages to, between, or among these determinants. However, the consistent calibration of shocks that affect the interrelationships across various macro-financial variables tends to be challenging.

The difficulty in designing scenarios increases if the shocks are assumed to reverberate across several countries, affecting financial institutions that have significant cross-border activities. Vitek addresses this issue in Chapter 4, using a structural macroeconomic model of the world economy, which features a range of nominal and real rigidities, extensive macro-financial linkages, and diverse spillover transmission channels. More recently, the growth-at-risk framework was used to inform stress testing scenario design based on how financial conditions, sectoral imbalances, and external factors influence the tail risks to future GDP growth (IMF 2017f; Prasad and others 2019). This concept was applied in the specification of both the baseline and adverse scenarios over a three-year risk horizon in the context of the Peru FSAP (IMF 2018c).

Quality information is essential for useful and credible stress tests. In Chapter 5, Heath and Bese Goksu, in a 2016 paper, observe that the availability of “right data” might have made it possible to better detect risk buildups prior to the GFC (even though the lack of data was not the main cause of the crisis). Information gaps could create additional uncertainty about the performance and soundness of financial institutions during periods of systemic stress. While prudential data are generally available to national supervisors, they remain confidential and often cannot be shared with supervisors in host countries, much less disclosed to the general public. This may create challenges for country authorities who may want to assure investors and depositors of the health of financial institutions during times of stress when trust in the financial system is critical. There are several key examples of how information gaps have undermined market confidence during the GFC:

- The variation and limited transparency in the calculation of risk-weighted assets across banks and jurisdictions created uncertainty about banks’ capital adequacy. This issue is covered by Le Leslé and Jobst in Chapter 10.
- The crisis stress tests conducted by the European authorities did not disclose important information on banks’ sovereign risk exposures, which was widely regarded as a major vulnerability. This issue is discussed by Ong and Pazarbasioglu in Chapter 14.

Data limitations may also require reliance on rules of thumb for stress tests. In Chapter 7, Hardy and Schmieder show that rules of thumb may have to be imported in the design of stress tests for a country in situations such as when national authorities or bank management are unable to estimate behavioral relationships robustly based on available data.

---

2 The design of a stress test and the communication of its results should be fully aligned with the policy objectives and the applicable restructuring and resolution regime, which Baudino and others (2018) confirm in their review of the three main building blocks (governance, implementation, and outcomes) of system-wide bank stress tests in the euro area, Japan, Switzerland, and the United States.
Addressing Gaps in Risk Coverage and Assessment

FSAP stress tests attempt to cover the relevant sources of risk affecting capital and liquidity conditions in the financial system in adverse macro-financial scenarios. The outcomes of stress tests are driven by the initial identification of these risks—in detecting, monitoring, and mitigating their buildup based on known vulnerabilities from common exposures, risk concentrations, and interdependencies within the financial system. The IMF staff has made significant efforts to close important gaps in risk coverage that were highlighted by the GFC to ensure that FSAP stress tests are fit for this purpose and encompass the following four essential domains, namely: (1) the dynamic approach to modeling institutional behavior under stress, (2) the interaction between solvency and liquidity risks, (3) feedback loops with the real economy, and (4) spillover effects from interconnectedness.

Dynamic approach to modeling institutional behavior under stress

A more dynamic approach considers changes in institutional behavior that can impact both capital and liquidity conditions under adverse macro-financial scenarios. In stress situations, banks may ring-fence liquidity, adjust their balance sheets, and/or restrict profit distribution, or be required to do so. These measures could include (1) limiting flows of funds within a banking group, (2) increasing liquidity through asset sales and/or slowing credit growth, (3) raising capital, and/or (4) reducing dividends. Hence, stress tests should adopt a more dynamic approach that considers changes in regulatory requirements or institutional behavior that could affect banks’ financial statements. In Chapter 6, Cerutti and Schmieder show that the use of both consolidated and unconsolidated balance sheet data is necessary to estimate the potential impact of ring-fencing on international banking groups. This dynamic approach was included as part of the bank stress tests in the Spain FSAP, which took place during the European sovereign debt crisis (IMF 2012b).

Interaction between solvency and liquidity risks

Liquidity and solvency risks faced by individual institutions are increasingly intertwined during times of stress. They tend to be influenced by system-wide liquidity conditions associated with the interconnectedness and network effects within the financial system. Stress tests that do not account for the interaction between solvency and liquidity shocks substantially underestimate the risk exposure of individual banks and banking sectors (Puhr and Schmitz 2014; BCBS 2015). Empirical evidence suggests that the interaction between solvency and funding costs (1) is indeed statistically significant and (2) might be economically relevant, especially during periods of stress. In Chapter 8, Schmitz, Sigmund, and Valderrama show that the interactions between liquidity and solvency risks are material and argue that failure to incorporate the solvency-liquidity nexus in stress tests could lead to significant underestimation of shocks on bank capitalization. They also find evidence of nonlinearity between solvency and funding costs.

The practical application of the solvency-liquidity linkage in stress tests remains at an early stage. A limited set of existing stress testing models contain liquidity and solvency interactions or network modules, in considering contagion and systemic risk from a cross-functional perspective (Barnhill and Schumacher 2011; Schmieder and others 2012; Babiluga and Spaltro 2014; Jobst 2014; Krznar and Mattheson 2017; Cont, Kotlicki, and Valderrama 2019). While these models have remained necessarily simple, given various empirical and data constraints, they could be extended to accommodate prudential data, which would help refine scenario specifications such as the duration of shocks and the adjustment dynamics as the financial system stabilizes over time (Grillet-Aubert 2018). They could also be integrated with relevant analytical tools, such as general equilibrium models, agent-based models, networks, and behavioral analysis.

Feedback loops with the real economy

The two-way interaction between the real economy and financial activities, and related feedback effects generated by financial institutions’ reaction function to stress, requires a dynamic specification of transmission channels. It also necessitates the consistent and comparable design of macro-financial scenarios. Such specifications could be enriched with insights into the adjustment process of economic agents to price and output shocks from full (or partial) equilibrium macroeconomic models. Work in this area remains at an early stage, with the IMF staff developing models that can credibly capture important feedback effects. In Chapter 12, Hesse, Ferhan, and Schmieder simulate the potential impact of spillovers from a crisis on banks’ liquidity and capital positions, and then examine their impact on the real economy. They find that spillovers have a highly nonlinear impact on both aspects (liquidity and solvency) of bank soundness.

Spillover effects from interconnectedness

The interlinkages within financial systems (such as interbank markets) or interactions between financial and nonfinancial entities within and across national boundaries (through common exposures, such as the property market) can result in financial contagion and spillovers. Prior to the GFC, credit risk shocks in FSAP stress tests focused largely on the capital impact of banks’ local exposures to firms and households without considering cross-border exposures (through branches and subsidiaries). Since then, spillover effects have been introduced through network analysis, for example, in the FSAPs for Australia, France, Japan, Spain, and the United States (IMF 2012a, 2012b, 2013a, 2017c, 2019); the simulation of ring-fencing, for example, the Spain FSAP (IMF 2012b); and shocks to business activities in...
other countries, for example, the United Kingdom FSAP (IMF 2016c). In Chapter 13, Kim and Mitra explore the significance of spillover effects from cross-border banking linkages. They use the network model of Espinosa-Vega and Solé (2011) to estimate the impact of credit risk and funding shocks on bank capital through direct and indirect transmission channels. They also model the relationship between spillover effects from cross-border credit and funding shocks and GDP growth rate surprises and find that funding vulnerabilities have implications mostly for the latter.

Other facets of solvency risk

Solvency stress tests remain largely focused on credit and market risks. The most common among the latter are interest rates, foreign exchange rates, and credit spreads, as well as equity, real estate, and commodity prices. However, the GFC highlighted additional facets of solvency risk, which have since been included in FSAP stress tests, notably: (1) the nonlinear impact of solvency on funding costs, which is demonstrated by Schmitz, Sigmund, and Valderrama in Chapter 8; (2) the potential valuation losses of sovereign exposures and other low-default assets, which have influenced stress tests via the analysis of the bank-sovereign nexus (IMF 2018b, 2018d) and are covered by Jobst and Oura in Chapter 9; and (3) the realization of contingent liabilities (such as guarantees, commitments, and derivatives), which have been mostly addressed via the impact of net cash outflows on the capital assessment of banks under stress in the United Kingdom FSAP (IMF 2016b).

Practical Approaches in the Implementation of Important Conceptual Aspects of Stress Testing

Stress testing is arguably more art than science. While the stress test methods and models covered in the first book are technical in nature, the principles, concepts, and frameworks described in this book often require some expert judgment and experience, especially if policy considerations or empirical constraints require practical and pragmatic solutions at different stages of a stress testing exercise. Heuristics can facilitate the calibration of shocks and the modeling of typical behavioral relationships, complemented by detailed analyses of banks’ financial statements and circumstances. In Chapter 7, Hardy and Schmieder identify several helpful rules of thumb for stress testing bank solvency, with a focus on the elasticity of credit losses, preimpairment income, and credit growth during times of stress.

Stress test results could also be influenced by model errors and parameter uncertainty. In Chapter 11, Taleb and others introduce a simple approach that helps evaluate how well tail risks are captured in stress tests. The authors argue that stress tests capture only first-order effects of negative impacts associated with tail shocks; they propose a heuristic as a second-order (robustness) test to detect nonlinearities in the tail behavior of risks. More specifically, their method identifies potential convexities that could under- or overstate the impact of tail events.

Wider Coverage of Financial Activities

The increasing importance of nonbank financial services for macroprudential surveillance has gradually widened the perimeter of stress testing beyond the banking sector. Improved data availability, enhanced statutory reporting, and supervisory coordination facilitate the integration of a wider range of nonbank financial institutions and markets into stress tests. However, limited data access and various other data-related factors (such as standards and classifications) continue to constrain the development of comprehensive stress testing models and frameworks for nonbank financial intermediaries and auxiliaries.

Building a complete map of funding interconnections between money and derivatives markets and various financial entities remains challenging. While a macroprudential perspective to stress testing and regulations could have helped prevent the GFC, Aikman and others (2019) argue that it would still have been difficult to understand the fragility of funding flows across the system and their knock-on effects on nonbank financial entities without covering the entire financial system to reveal the full extent of existing vulnerabilities. Clearly, there are still significant gaps in the coverage of all relevant market participants and the interlinkages between solvency and liquidity conditions across banks and important FMI elements.

At the IMF, stress testing of the insurance sector has become a more regular exercise during FSAPs since the GFC. Indeed, the IMF staff has developed a common approach to stress testing insurance companies. In Chapter 17, Jobst, Sugimoto, and Broszeit review the state of system-wide solvency stress tests for insurers, comparing national practices and drawing on experience from FSAPs to derive principles and concepts, and distill practical guidelines for a more comparable and consistent implementation of system-wide insurance stress tests.

The IMF staff has also run stress tests on other nonbank financial institutions. They include pension funds, investment funds, and critical FMI elements, including central counterparties (CCPs) and central securities depositories. The technical guidelines for stress testing defined benefit pension plans are covered in the first book. Technical work on liquidity stress testing for investment funds and its potential integration with bank stress tests was undertaken in the context of the FSAPs for Ireland, Luxembourg, Sweden, and the United States (Bouveret 2017; IMF 2015b, 2016e, 2017c, 2017d). Detailed risk assessments of critical FMI elements have been undertaken in several FSAPs, focusing on the largest central counterparties and central securities depositories in Europe and the United States. ³ The financial stability analysis of CCPs varies across countries and is largely driven

³ For instance, Euroclear (Belgium), Eurex (Germany), CC&G (Italy), Clearstream (Luxembourg), Euro CCP (Netherlands), Nasdaq (Sweden), LCH (United Kingdom), and CME (United States) (IMF 2013a, 2013b, 2013d, 2015a, 2016a, 2016d, 2016f, 2017a, 2017b, 2017d, 2018a).
by existing stress testing frameworks developed by local regulators (Anderson, Cerezetti, and Manning 2018; ESMA 2018). The FSAPs encouraged national authorities to develop standardized stress testing to support their assessment of CCPs’ loss-absorbing capacity.

The nature of stress testing continues to expand. Its scope is becoming more diverse to account for different characteristics of financial services across countries. For instance, the IMF staff has been developing conceptual guidelines for the coherent implementation of solvency stress testing of Islamic banks. The formulation of stress tests also considers the evolving nature of risks. For example, new types of risks are emerging, such as climate change and technological disruptions. The IMF staff has provided initial considerations on these risks (He and others 2017; Jobst and Pazarbašioglu 2019), but a fuller discussion on their incorporation into stress tests is outside the coverage of this book.

Stress testing exercises comprise many “moving parts” of continuously evolving risks and the methodologies required to adequately capture them. Managing these dynamics is even more challenging in the context of FSAPs, which need to be sufficiently comprehensive and comparable for a wide range of countries with varying characteristics of financial systems. FSAP stress tests must also remain sufficiently flexible to accommodate the specific nature of different local regulatory requirements (including on data) and political sensitivities. Hence, the development of common principles and concepts for well-designed stress tests, all within coherent frameworks, ensures that the financial stability analysis in FSAPs remains relevant and appropriate while complementing the efforts of national authorities in developing their own stress tests.

REFERENCES


Andreas A. Jobst and Li Lian Ong


8


