A WIDER Field of Vision

Stress tests must be adapted and broadened to assess the stability of the financial system as a whole

Dimitri G. Demekas

WHEN engineers want to make sure a structure or a system is well designed, they employ a technique called stress testing: they expose the system to shocks and strains that are far greater than what will be experienced in normal use to confirm specifications are met, determine breaking limits, or examine modes of failure.

Managers of financial institutions and, more recently, financial regulators adapted the tool of stress testing to measure the strength of individual financial institutions. They do so by subjecting portfolios to numerical simulations of large hypothetical "shocks," such as a severe recession, housing price decline, or stock market crash, and estimating their effect on profits, capital, or the ability of financial institutions to continue meeting their obligations.

But using stress tests to assess the resilience of the financial system as a whole is not as simple as adding up the results for the individual institutions. New approaches and techniques are needed to make stress tests a useful tool for financial stability analysis.

Simple start

Stress tests were first used for banks in the early 1990s (see box). These early models were relatively simple: they assumed an exogenous shock and traced the impact of associated losses on the capital of the individual bank. They made simplistic assumptions about how the bank would react to the shock—in terms of profit distribution, credit expansion, or debt reduction, for example. They focused on the solvency of the bank (how much capital it had left after the shock). The risk that an institution would run out of cash (liquidity risk) was treated independently from solvency, if at all, and interactions among banks and the feedback effects on the economy as a whole were generally ignored.

These stress tests had what economists call a microprudential, or single-institution, focus: their objective was to assess the likelihood of failure of individual institutions under adverse conditions. This, in turn, it was thought, would ensure the stability of the financial system as a whole.

Too much and too little

But even as bank regulators were adopting stress tests, many understood that ensuring the soundness of each institution was neither necessary nor sufficient to ensure that the financial system as a whole would remain stable and continue to function. As the late Andrew Crockett, then general manager of the Bank for International Settlements, put it, the microprudential approach to financial regulation may "strive for too much and deliver too little."

It may strive for too much because the occasional failure of an individual institution is not a problem if other institutions can step in and serve its clients, borrowers, and depositors. Building a regulatory system designed to avoid any failures risks providing excessive protection.

And it may deliver too little because firm-level regulation takes into account neither the potential for contagion among

individual institutions nor how each institution pursues compliance with capital rules. When, for example, a regulator pushes a troubled bank to restore its capital-to-assets ratio, the regulator does not care whether the bank increases its capital or shrinks its assets. But if a substantial proportion of the banking system shrinks assets simultaneously to meet capital requirements, the damage to the economy as a whole may be considerable. Unless the regulators take into account the interconnectedness and collective behavior of institutions in response to a shock and their possible impact on the financial system and the economy, they may fail to minimize the risk of distress for the system as a whole and the associated economic costs—in short, systemic risk (Crockett, 2000).

The recent global financial crisis underscored dramatically the importance of systemic risk and the failure of microprudential regulation to contain it. In 2008, U.S. Federal Reserve Chairman Ben Bernanke called for a widening of the "field of vision" of regulators and supervisors to incorporate systemic risk (Bernanke, 2008). Or, as Crockett had put it, "marrying the microprudential and macroprudential dimensions of financial stability."

A new generation

Moving from traditional microprudential stress tests toward a "new generation" of macroprudential stress tests presents two challenges:

• Introducing systemwide or *general equilibrium dimensions,* so that the outcome of the stress tests depends not only on the size and nature of the initial shock and the buffers of each financial institution but also on the behavioral responses of these institutions to the shock and on their interactions with each other and with other economic agents, including borrowers, depositors, and investors. This is particularly important if the stress tests cover a long time horizon, say three or five years, during which the effect of these interactions can be sizable.

New approaches are needed to make stress tests a useful tool for financial stability analysis.

• Shifting the focus of stress tests from individual institutions to the resilience of the system as a whole—its ability to continue functioning and providing financial intermediation services to the real economy.

How much progress have stress testers made in tackling these challenges? How much has this "wider field of vision" been adopted in practice?

A review of the experience of central banks, supervisory agencies, and the IMF since the crisis finds that stress testing has made significant progress in tackling the first of these two challenges but much less in dealing with the second.

Many models that incorporate some systemwide "general equilibrium" dimensions into stress tests are available and widely used. They fall into two broad categories.

• ***Balance-sheet-based models*** use individual bank balance sheet data to assess the impact of a shock on asset quality, income, and—ultimately—capital (for solvency tests) or various measures of cash flow or liquidity (for liquidity tests) of individual banks. The results are then aggregated to give an idea of the vulnerability of the system as a whole.

In this approach—common across central banks and supervisory agencies around the world—the dimensions the stress tester intends to capture, whether solvency-liquidity interactions, behavioral responses, or macroeconomic feedback effects, are built explicitly into the model. This makes it possible to trace the effect of the shock through various channels and figure out how much each channel contributes to the final outcome.

This benefit comes at a price. First, analytical and computational complexity and data requirements increase rapidly as features are added to the models. This renders them slow, cumbersome, and costly to construct and run. Second, because they rely on bank balance sheet data, they depend crucially on the availability and quality of these data.

But by far the biggest flaw in this approach is the fact that, given the different ways banks are interconnected, the sum of the losses or capital shortfalls of individual banks is not representative of the vulnerability of the system as a whole: correctly aggregating individual shortfalls requires some knowledge of the complex interdependence between individual bank balance sheets.

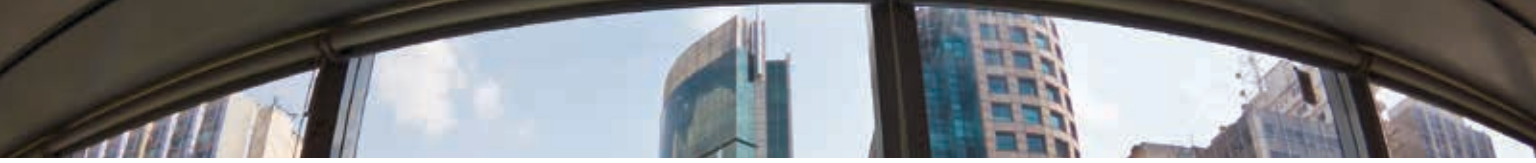
• ***Market-price-based models*** use (mostly) market data to infer the probability of distress or default of individual institutions. They capture—at least in principle—all sources of

Origins of financial stress testing

One of the early adopters of stress tests was the U.S. financial services firm JPMorgan Chase & Co., which in the early 1990s used what is called value at risk (VaR) methodology to measure the market risk of a given shock—how much the changes in asset prices would affect the value of the bank's portfolio.

Regulators soon caught up. It had long been understood that banks financing themselves with government-insured deposits have an incentive to take excessive risks. So the goal of capital regulation was to force banks to internalize at least some of the unexpected losses should these risks materialize, thus mitigating moral hazard and protecting depositors. Regulators saw that stress testing was a way to estimate potential losses under adverse scenarios, and could be a key input in capital regulation.

Stress tests became a regulatory staple in the early 2000s, when the international rules on capital adequacy known as Basel II required banks to perform stress tests for market risk and, in some cases, credit risk. These tests had to be "plausible, severe, and relevant" to help banks evaluate their capacity to absorb losses and identify steps they could take to reduce risk and conserve capital (BCBS, 2005). Equipped with this tool, regulators could ensure the soundness of each institution by requiring it to hold a minimum amount of capital in proportion to its risky assets.



vulnerability and contagion, including the risk of bank runs triggered by investors' self-fulfilling fears. Such risk might not reflect the real financial condition of a bank, which may have been healthy before the run. Another advantage is their computational simplicity.

An obvious weakness of these models is their reliance on market data, which are "noisy" and may overestimate or underestimate risks. That means that bank risk indicators estimated from these data may be excessively volatile,

The obvious remedy is to use many extreme but plausible scenarios for stress tests.

and may not provide a sound basis for bank management or supervisory action. Another pitfall is that by extracting information from market data and constructing a summary metric of bank soundness, market-price-based models do not allow the stress tester to differentiate between the various factors—initial shock, risk interdependence, common exposures, and cross-institutional contagion—that contribute to the final result: all these factors are lumped into the probability of default or distress generated by the model. This has led some critics to dismiss such models as "black boxes."

In contrast to the progress made toward incorporating general equilibrium dimensions into the traditional microprudential stress-testing framework, relatively few advances have been made in tackling the second challenge: correctly measuring the resilience of the financial system as a whole and its ability to continue providing financial intermediation services under stress.

This measurement must be done in a way that allows individual banks and their supervisors to take action on the results. It is hard to build a model that correctly measures systemic risk and the contribution of individual institutions to that risk and then relates the results to each bank's established regulatory framework, such as capital adequacy ratios or liquidity rules. And it is even harder to make this model robust enough to use in a variety of environments and for a variety of financial institutions, but simple enough to explain to supervisors, bank managers, and market participants.

Moving the dial

How do we move from where we are today toward more effective macroprudential stress tests?

Use a variety of models: Given the limitations of the existing stress-testing frameworks, it is surprising to see several central banks and regulatory agencies relying on single individual models. This makes the outcome of the stress test hostage to the limitations of a single analytical framework.

Instead, a variety of models should be used for macroprudential stress testing. The challenge would then be to interpret and synthesize the results of the different models into a coherent and persuasive narrative. Should the different results be combined or averaged according to a strict rule? Should qualitative judgment be used in weighing different—and potentially contradictory—results? These are complex questions on which there is no consensus among practitioners. But this is a challenge well worth tackling, because it would enhance insight into systemic risk and the quality of the ensuing conversation about financial stability, both within the supervisory agency and with the banks.

Run more—and smarter—stress scenarios: Most stress-testing exercises are limited to one or two macroeconomic stress scenarios (for instance, an "adverse" and a "severe" recession scenario). This approach has a major problem: resilience to a shock of a given probability does not imply resilience to all shocks with the same probability. It also ignores the increasingly important cross-border nature of risk: banks and other financial institutions are increasingly interlinked across borders and may be vulnerable to shocks that originate in—or propagate through—a foreign country or market. The outcome of a test of a single stress scenario focused on a domestic recession may thus be misleading.

The obvious remedy is to use many extreme but plausible scenarios for stress tests. This would provide a better sense of the resilience of the system to a range of shocks than would a single scenario. Using multiple scenarios (as well as a variety of models) would also have another big advantage: it would minimize the scope for individual institutions to "game the test"—gear portfolio choices toward passing a specific stress test—a risk that regulators recognize (Office of Financial Research, 2012; Bank of England, 2013).

In addition to the *number*, a related issue is the type of scenario used in stress tests. In most cases, the main stress scenario is an adverse macroeconomic shock exogenous to the financial sector, such as a severe recession or a housing price bust. But in many actual crises, the shock originates entirely inside the financial system and is then followed by a recession. In a study of 43 banking crises in 30 countries, Alfaro and Drehmann (2009) show that only about half were preceded by adverse macroeconomic conditions.

Effective macroprudential stress tests should therefore involve a higher number and a wider range of "smart" stress scenarios covering a variety of risks, including domestic macroeconomic shocks, asset price moves, and cross-border contagion. Such testing would require an in-depth understanding of the risks affecting the financial system, including cross-border dimensions, and would complicate the task of synthesizing and communicating the results—especially when accompanied by a variety of models. It is these challenges that have held back many supervisors from moving in this direction. But given the significant pitfalls of limiting the number of scenarios to just one or two, it may be time to reconsider the cost-benefit balance of the current approach.

Expand coverage to nonbank financial entities: Microprudential stress tests have been traditionally applied to banks because these were the predominant agents of financial intermediation. But today the line between banks and nonbanks (such as investment banks that provide commercial-bank-like services) is blurred; the nonbank industry has expanded greatly

in size and importance; and the global financial crisis demonstrated that banks and nonbanks are deeply entwined and risks move easily between the two. So stress tests should cover both banks and nonbanks, and the choice of which nonbank entities to incorporate into the stress-testing framework should depend on country circumstances. Priority should be given to sectors that are closely connected with banks through ownership or financial linkages—typically insurance companies, for which well-established stress-testing models already exist. Asset management companies, mutual funds, and sometimes pension funds can also be important providers of liquidity to banks and could be affected by—or be a channel for—a systemic shock.

Explore agent-based models: Micro- and macroprudential stress-testing models—like all traditional economic models—share a fundamental trait: they assume that individuals and institutions always behave rationally in ways that can be modeled based on past experience, and that policy decisions influence this behavior in the same way for all market participants. These assumptions miss some critical points about financial crises, notably

- the fact that market participants are heterogeneous and often make less-than-rational decisions, especially under stress;
- the emergence of a new dynamic under stress, when relationships among financial institutions can change suddenly; and
- the fact that the response of regulated institutions to policy signals depends partly on the conditions they are facing. For example, raising the regulatory capital requirements put in place in normal times to ensure banks have sufficient capital buffers may have no positive effect on systemic stability in times of crisis.

Agent-based models can capture many of these aspects. An agent-based model assumes autonomous, heterogeneous agents with limited information and specifies simple rules that dictate how they will act under different circumstances. These rules can vary across different types of agents (for instance banks, depositors, investors) and allow for herd behavior and panics. The model determines how these agents can interact (for example, how they form networks) and can explore various types of shocks. Agent-based models are increasingly used for macrofinancial modeling, and relatively simpler versions have been used to explore the impact of stress scenarios on bank solvency, liquidity, and contagion.

Agent-based models are complex and have their own pitfalls. Implementing them would require a shift in the approaches traditionally taken by (and the skills traditionally required of) stress testers. Nevertheless, the limited experience so far suggests that they can provide unique insights into the aspects that matter most in a stress scenario: the behavior of banks and other economic agents in times of crisis.

Embed stress tests into the financial stability policy framework: The recent explosion of interest in stress testing is creating a risk. Policymakers, regulators, market participants, and the broader public may focus excessive attention on stress tests, take their results out of context, and give them much greater weight than they merit in guiding policy action. This risk is evident in the way stress test results tend to dominate the public debate on the health of banks in the United States following the introduction of the Dodd-Frank Act, the centerpiece of postcrisis regulatory reforms in the United States, as well as in Europe, following a string of highly publicized tests by the European Banking Authority. This unprecedented attention on stress tests seems at times to overshadow, rather than inform, the conversation about financial stability among all relevant stakeholders in society.

This risk has been noted before. In setting out best-practice principles for macroprudential stress testing, the IMF put it this way (IMF, 2012, pp. 44-45):

"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.

[…] 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."

The call to embed stress tests firmly in the financial stability framework is essentially a call for caution. Macroprudential stress testing is just one of many tools to assess systemic resilience. It should be treated as a complement to other tools, such as early warning indicators, and—crucially—should be combined with the insights gained by the ongoing supervision of individual financial institutions. ■

Dimitri Demekas is an Assistant Director in the IMF's Monetary and Capital Affairs Department.

This article is based on the author's IMF Working Paper 15/146, "Designing Effective Macroprudential Stress Tests: Progress So Far and the Way Forward."

References:

Alfaro, Rodrigo, and Mathias Drehmann, 2009, "Macro Stress Tests and Crises: What Can We Learn?" BIS Quarterly Review *(December).*

Bank of England, 2013, "A Framework for Stress Testing the UK Banking System," Bank of England Discussion Paper (London).

Basel Committee on Banking Supervision (BCBS), 2005, "International Convergence of Capital Measurement and Capital Standards: A Revised Framework," BIS Technical Report (Basel: Bank for International Settlements).

Bernanke, Ben S., 2008, "Reducing Systemic Risk," speech at the Federal Reserve Bank of Kansas City's Annual Economic Symposium, Jackson Hole, Wyoming, August 22.

Crockett, Andrew D., 2000, "Marrying the Micro- and Macro-Prudential Dimensions of Financial Stability," speech at the Eleventh International Conference of Banking Supervisors, Basel, September 20–21.

International Monetary Fund (IMF), 2012, "Macrofinancial Stress Testing: Principles and Practices," IMF Monetary and Capital Markets Department Policy Paper.

Office of Financial Research, 2012, Annual Report *(Washington: U.S. Treasury Department).*