

SPEECH

Economic Research and Stress Testing

June 24, 2015

James McAndrews, Executive Vice President and Director of Research

Remarks at the Fourth Annual Stress Test Modeling Symposium, Federal Reserve Bank of Boston, Boston, Massachusetts¹

As prepared for delivery

Stress testing has become an important tool for bank supervision in the wake of the financial crisis of 2007-09. A prominent example of stress testing is the quantitative assessment conducted by the Federal Reserve as part of its annual Comprehensive Capital Analysis and Review (CCAR) of large banking organizations operating in the United States. The Federal Reserve uses its stress test to determine whether the banking organizations, given their current capitalization and future planned capital actions, would be likely to remain adequately capitalized under stressful macroeconomic scenarios. The results of the quantitative assessment figure importantly in Federal Reserve policy decisions regarding whether to object to firms' plans for future capital distribution and retention—for example, planned dividend payouts. More generally, stress tests provide valuable information about the safety and soundness of individual firms, and, significantly, allow comparisons and aggregation across a range of firms. For example, the stress tests conducted by the Federal Reserve under the terms of the Dodd-Frank Act—the DFAST stress tests—provide a very useful perspective on both individual firms and the banking system. Stress tests are therefore used to gain insights applicable to both microprudential supervision of individual firms and macroprudential assessments of broader financial system vulnerabilities.

Given the now-central role of supervisory stress tests, it is appropriate to step back and assess how the science and intellectual framework underlying stress testing is evolving, and how economic research can ensure that these underpinnings are sufficiently robust to justify the continued confidence placed in stress tests by central banks and supervisors. Stress testing is inherently challenging, for a number of reasons. It requires specifying one or more macroeconomic scenarios that are stressful but not implausibly disastrous. Scenarios must have certain elements of realism, but are certainly ahistorical, and may represent structural breaks in the processes that are being stressed. Stress testing requires forecasting earnings and capital conditional on the nature of the scenario; not only is this more difficult than unconditional forecasting, but the objects of interest are the tails of the distribution—the lower quantiles, which, almost by definition, are infrequently observed in historical data. Stress tests rely on well-specified models of the processes that generate earnings for banks, and there are myriad ways to build such models; for example, they can be extremely disaggregated, built up from models of individual loans or securities, or they can be much more aggregated. Supervisors and firms are, separately, both engaged in building these models; the supervisor has the benefit of having access to data from a panel of firms, while each firm mainly knows its own history of asset performance.

The complexity and difficulty of stress testing does not end there. Two very important elements, bookends one might say, are the theory that lies behind stress testing, and the development and implementation of stress test models. Both of these elements are crucial to using the output of stress tests in useful ways and to delivering outputs that reflect the risk of each bank's assets. I'll have more to say on these points later in my remarks.

How can economic researchers help meet the daunting challenge of designing and implementing stress tests that will usefully reveal vulnerabilities in individual firms and in the financial system more broadly? I'll suggest that there are various ways economic researchers can improve the science and practice of stress testing. It seems clear to me that quantitative stress testing is an area in which the insights of economists are particularly valuable. Good stress testing practices should call upon the singular skills and knowledge of economists in econometric modeling, macroeconomic forecasting, banking theory, models of firm and household behavior, and many other areas of specialization. The high demand for economists to work on stress tests at the Federal Reserve Bank of New York and throughout the Federal Reserve System during the last six years has clearly reflected these deep underlying needs.

Obviously, these skills are critical for the practical implementation of stress testing—calculating inputs to revenue and loss forecasts and evaluating the robustness of the models. More broadly, these skills provide the logic for the intellectual framework that supports stress testing as a useful and current policy tool. In what follows, I'll suggest several specific areas in which economic research is needed to continue to build the rigor and credibility of supervisory stress testing.

Areas for Economic Research

Theory

Stress testing cannot become a purely mechanical or statistical process. Instead, the design of stress tests must be informed and guided by sound economic theory, both in terms of identifying the goals of stress testing and determining the specifics of how the tests should be implemented. This need for economic theory is seen in many areas. Let me describe a few.

Procyclicality: How should the severity of stress tests vary over the economic cycle? A good case can be made for setting tough tests during boom periods, when vulnerabilities may be growing but hidden. For example, although mortgage delinquency rates were low in 2004 and 2005, significant risks were building up in the mortgage market at that time.

Some statistical models, however, may instead predict *lower* losses during good economic times, particularly if current loan or security performance is used as conditioning information in the model. This raises an important question: should other aspects of the stress test, such as the severity of the macroeconomic scenarios, be adjusted to compensate for any such potential procyclicality? If so, how, and by how much? And does setting tougher tests during boom periods necessarily imply having weaker tests during periods of stress? If it does, will having weaker tests during stress periods undermine the credibility of the stress testing regime? Theory is needed to help guide thinking on these issues. Such work could draw on the lessons learned from related research on the procyclicality of Basel II risk-based capital requirements.²

Feedback Effects and Fire Sales: A second area in which further theoretical work can be focused is the feedback effects of stress at individual banks on other financial firms and markets. This theme has been explored in many papers, such as Greenwood, Landier, and Thesmar (2015) and Duarte and Eisenbach (2013) and the sources cited in those papers. The papers examine the role of spillovers in asset price declines among banks—that is, situations in which the abrupt sale of assets by some banks causes other banks, holding concentrated positions in the asset sold, to experience significant declines in the value of their holdings, leading to further asset sales. Such “fire sales” are emblematic of the complexity of the economics in the financial system, in which banks often hold similar assets and decisions by one firm to hold or sell an asset can affect the range of possible actions by other firms.

To date, most supervisory stress tests don’t directly include feedback effects; instead, the macroeconomic scenario is itself intended to be stressful enough to include an assumption of difficult market conditions, consistent with negative feedback being already incorporated in the environment. Whether this approach is sufficiently informative to policymakers, and whether it leads to a sufficient focus on managing the risk of exposure to fire sales, are questions that could benefit from further theoretical work by financial and macro economists.

Combined Liquidity and Asset Stress: A third area meriting further exploration is the role of liquidity in stress tests. The Federal Reserve conducts two types of stress tests: the CCAR, which stresses asset values and bank capital, and the Comprehensive Liquidity Analysis and Review, or CLAR, which focuses on stresses to funding and liquidity. But as we saw in 2008, periods of financial instability are likely to involve a combination of stresses to capital *and* liquidity, with important feedback loops between the two. This issue is, of course, related to the topic I just discussed—Firms forced to sell assets in a fire sale because of liquidity problems may realize losses, reducing their capitalization and perhaps exacerbating liquidity problems. A colleague and coauthor, Thomas Eisenbach, delivered our joint research on this topic at this conference last year. Our research suggested one way that combined capital and liquidity stress tests, so-called dual stress tests, might be conducted. But much more thought is needed here.

Lucas Critique: Another theoretical issue that could be explored further is the reliance on historical data to estimate econometric models that are then used by policymakers to project stressed capital ratios and to mandate additional capital retention or capital raising by firms subject to the stress test. This procedure runs afoul of the Lucas Critique (Lucas 1976), which suggests that the relationships that are revealed by the use of historical data are likely to change if new policies are imposed based on the use of the estimated relationships. While many papers question the empirical relevance of the Lucas Critique, it is nonetheless important to have a coherent theory of the policy process when stress testing is used to determine firms’ minimum regulatory capital levels. The problem might even be more severe when firms may be actively attempting to circumvent the intent of the policy. Such a coherent theory might provide a better guide to policymakers on how to use the results of stress tests.

Effects of Bank Capital on Financial Intermediation and the Real Economy: More broadly, additional work is needed to further our understanding of the relationship between bank capitalization and the provision of credit and other financial intermediation services to promote growth during a period of economic stress. Stress testing as currently practiced by the Federal Reserve is oriented toward ensuring that banks have sufficient capital. An understanding of the theoretical costs (and possible benefits) of different levels of bank capital, both for the individual firm and the financial system, will help ensure that we make the right trade-offs as a society.

Empirical Work

As I mentioned earlier, stress testing necessarily involves conditional forecasting of the tails of distributions of losses and revenues. Many opportunities exist to develop new approaches and to incorporate existing insights from macroeconomic

forecasting. Two approaches that might be usefully applied in the stress testing context are Bayesian techniques and methods to incorporate structural breaks in the stochastic process being estimated.

The Federal Reserve has collected an unprecedented amount of granular cross-sectional data from the banks participating in the stress tests. The data allow for the construction of structural models—models that incorporate loan-by-loan or asset-by-asset projections of cash flows with valuable details on asset and loan characteristics. These models may sidestep some of the concerns that arise with more aggregated modelling. For example, better information on borrower risk allows stress tests to be more responsive to shifts in bank risk-taking and other changes in asset composition. This information also represents an opportunity for economists to better understand bank risks since, historically, many models of banks and their assets have not had such rich data to work with. Making use of these data to inform policy choices and monitor bank risk-taking, while ensuring the confidentiality and integrity of firm data, is both a challenge and an opportunity.

The design of macroeconomic stress scenarios is an area of interest in itself. As the economy evolves over time, relevant potential scenarios are largely “ahistorical,” in that they will contain constellations of asset price movements and macroeconomic conditions that might bear little resemblance to historical patterns. Crafting such scenarios will nonetheless entail a great deal of empirical work to gauge the range of possibilities and to determine how probable such configurations of events might be.

Combining ahistorical stress test scenarios and models developed with historical data presents a further empirical challenge. Since macroeconomic variables generally move together, it can be hard to identify or calibrate the importance of any single variable. Models calibrated using historical data may be challenged, therefore, to capture fully the impact of ahistorical movements in macroeconomic variables. Finding ways to address this challenge is another area where the science of stress testing could be advanced.

Performance Testing and Model Validation

Precisely because stress tests are attempts to project unlikely outcomes—those that occur under stressful macroeconomic conditions, the actual evolution of events generally won’t closely follow the stress scenario, nor will bank losses necessarily accrue according to stressed loss estimates. Notwithstanding that stress testing is a test of possible risks, not likely outcomes, the performance of the underlying models can and should be tested in various ways. Here, too, economic theory and empirical work can be combined to suggest new ways to accomplish this important exercise. How can stress testing model performance be measured and tested if severely adverse macroeconomic outcomes do not arise? How can we estimate the precision of forecasts and quantify increases to firm and system stability that stress tests promise?

A closely related dimension of stress testing that requires further investigation is the practice of model validation. Models are validated in a number of ways—for example, by checking model codes and data sources—but economic models should also be validated in a larger sense by checking underlying assumptions and the efficiency of the model in measurement. Such an undertaking requires a dialogue with the broader academic community of economists, which validates models by peer review and replication of results. And indeed, the Federal Reserve sponsors just such a group of academic economists—the Model Validation Council, some of whom are here today—to provide expert and independent advice on stress test model validation.

How Economists Can Help

Now that I’ve outlined, surely in an incomplete way, a number of areas in which economists can usefully improve the intellectual framework of stress testing, the question arises, Which economists should be employed in those pursuits? Here, I’ll suggest that both academic economists and economists in central banks and other regulatory agencies, who work closely with their supervisory colleagues, are needed to make progress.

Furthermore, I’ll recommend that the central bank economists, like those at the Federal Reserve Bank of New York, be required as part of their career development and evaluation to seek to publish papers in peer-reviewed scientific journals. There is a strong long-run complementarity between research and policy analysis. The skills needed to be successful in research and in policy analysis are largely the same. It is vital that economists continue to invest in new knowledge and maintain their human capital.

The peer-review process is an important external signal that economists are obtaining these skills. The best way, or should I say the most economical way, to build and retain a staff of highly qualified economists capable of making vital contributions to policy work is to encourage the economists to pursue research with the aim of publishing in peer-reviewed journals. To this end, central bank economists should be given a significant amount of independent research time, with the appropriate accountability for their research output. In the specific context of stress testing, this also means a commitment on the part of the Federal Reserve to make confidential data available for research purposes and to encourage collaboration between system economists and outside academics.

Central bank economists from research departments also bear a special responsibility to bring their full complement of knowledge, technical expertise, and inventiveness to important policy efforts such as stress testing. At the New York Fed, this expectation ranks equally with the idea that economists must publish. Effective policy analysis of the type necessary to improve stress tests requires a close collaboration between economists and key staff from other areas of the Federal Reserve, such as those charged

with supervision and market monitoring. This collaboration is facilitated by establishing and nurturing long-term relationships with colleagues outside of research departments. In contrast with academic economists, an internal staff of central bank economists is in the best position to establish such relationships and to cultivate the necessary trust. Academic economists have less incentive to develop relationships or to acquire institutional knowledge that is necessary to the analysis of specific, and sometimes acute, policy problems but that does not provide a general return in the broader market.

At the same time, academic economists are often expert in areas requiring a significant investment in innovative theoretical and empirical methods. Their work, at a more general level, can be very useful for assessing the work of central bank economists and for suggesting promising new avenues of development. Consequently, there are complementary roles to be played by central bank and academic economists in developing stress testing as a discipline.

A Robust Intellectual Framework for Stress Testing

Stress testing is a relatively new supervisory discipline that requires the development of a robust intellectual framework to achieve its objective—the forceful and effective supervision of systemically risky financial firms. I’ve suggested a number of areas in which work by research economists, both within central banks and in academia, can be used to build that framework over time.

I found two clear lessons in the paper presented at this conference by Frame, Gerardi, and Willen (2015). In that fascinating and important paper, the authors document how the stress tests of Fannie Mae and Freddie Mac, designed over a decade by the Office of Federal Housing Enterprise Oversight, OFHEO, failed in every sense. By its governing law, OFHEO was required to fully disclose its stress testing model, and so published all stress scenarios, empirical specifications, and parameter estimates in the Federal Register. The authors suggest that, from the introduction of the stress tests in 2002 through 2007 when the crisis hit, OFHEO never “re-estimated its mortgage default or prepayment forecasting model, nor introduced new variables, despite well-documented changes in mortgage underwriting practices during this time” (Frame, Gerardi, and Willen, p. 3).

The first lesson I draw is that full transparency in stress testing can be an enormous weakness; it can make tests less flexible and responsive, and give the firms subject to the test opportunities to circumvent the test. The second lesson is that it is important to update stress testing models over time to incorporate changes in borrower behavior, lending standards, and other aspects of the economic environment, as well as improvements in techniques of the kind that I have addressed in these remarks.

What I’ve argued is that to build the robust intellectual framework that stress testing needs requires the sort of dynamism that economic research can bring to such inherently difficult problems. Many commentators, including Schuermann (2013), warn against a sort of “model monoculture,” in which the regulator is forced to focus on a single model and all firms aim to manage to the model rather than to the real risks they face. This is certainly a significant problem, and its dangers are seen in the failures of the OFHEO stress tests. Thus far, my colleagues Bev Hirtle and Anna Kovner have examined this question empirically in posts in the New York Fed’s *Liberty Street Economics* blog and have not found strong evidence of convergence between the Fed’s and the banks’ stress test estimates. But the best method of averting this problem is exactly to have a robust intellectual debate that leads to innovations in stress testing, and that can allow the performance measurement of alternative approaches to stressing a bank’s condition.

Conferences such as this one are a tonic against settling on a narrow modeling approach to the myriad difficulties of stress testing, and I hope that more economists, both in central banks and in academia, will pursue new solutions to the challenges of stress testing.

References

- Duarte, F., and T. Eisenbach. 2013. *Fire-Sale Spillovers and Systemic Risk*. Federal Reserve Bank of New York *Staff Reports*, no. 645, October; revised February 2015.
- Eisenbach, T., and J. McAndrews. 2014. “Dual Stress Testing and Run Vulnerability.” Unpublished paper, Federal Reserve Bank of New York, March; revised June 2015.
- Frame, W. S., K. Gerardi, and P. S. Willen. 2015. *The Failure of Supervisory Stress Testing: Fannie Mae, Freddie Mac, and OFHEO*. Federal Reserve Bank of Atlanta Working Paper no. 2015-3, March.
- Greenwood, R., A. Landier, and D. Thesmar. Forthcoming. “Vulnerable Banks.” *Journal of Financial Economics*.
- Kashyap, A., and J. Stein. 2004. Cyclical Implications of the Basel II Capital Standards. *Federal Reserve Bank of Chicago Economic Perspectives* 28, no. 1 (First Quarter): 18-31.
- Lucas, R. 1976. “Econometric Policy Evaluation: A Critique.” In Brunner, K., and A. Meltzer, eds., *The Phillips Curve and Labor Markets*. Carnegie-Rochester Conference Series on Public Policy, p. 19-46. New York: American Elsevier.
- Repullo, R., and J. Suarez. 2009. “The Procyclical Effects of Bank Capital Regulation.” Working Paper, CEMFI, wp2012_1202.

Schuermann, T. "The Fed's Stress Tests Add Risk to the Financial System." *Wall Street Journal*, March 19, 2013.

¹ The views expressed in this speech are those of the author and do not necessarily reflect the position of the Federal Reserve Bank of New York or of the Federal Reserve System. The author thanks Beverly Hirtle, Anna Kovner, and James Vickery for extensive discussions and assistance in preparing these remarks.

² For example, see Repullo and Suarez (2009) and Kashyap and Stein (2004).
