

RETHINKING THE ROLES OF STRESS TESTING

**Daniel K. Tarullo
Harvard Law School**

**2022 Federal Reserve Stress Testing Conference
Boston, Massachusetts
October 7, 2022**

The very fact that the Federal Reserve has hosted a conference on stress testing each year for the last decade is testimony to its richness and significance as a regulatory tool. A look back at the agendas of previous conferences shows that there is never a dearth of relevant topics. This year's conference usefully addresses some ongoing issues such as how to integrate second-order effects into the models; some relatively new issues needing attention in the short term, notably the interaction of CECL with the stress capital buffer; and some new issues that will be with us for some time to come, such as how to adapt stress testing techniques to an evaluation of the financial impact of climate change.

Rather than adding my two cents to these discussions of current issues, I want to use the platform graciously extended by the conference organizers to step back a bit and ask some basic questions about the roles of stress testing in the financial regulatory system. Precisely because there are so many interesting and important issues raised by stress testing – some technically quite complex – there is a risk that we lose focus on these more basic questions. Yet without a disciplined articulation of the purposes for which stress testing is, or may, be used, we might not be providing the most appropriate answers to some of these very issues.

Of course, even to refer to “roles” – plural – of stress testing is to highlight the obvious. One role dominates all others. The use of the Fed's annual stress test as a critical input into the determination of regulatory capital requirements for the nation's largest banks likely accounts for much of the interest in stress testing, as reflected in academic work, policy papers, and

attendance at these conferences. With a decade's experience using stress testing in the Comprehensive Capital Assessment Review (CCAR) and now the stress capital buffer approach (SCB), and especially in light of Vice Chair Michael Barr's very sensible plan to have the Board of Governors take stock of capital requirements as a whole, this seems an opportune time to ask my first question: How stress testing has filled out this dominant role. It is also a good occasion to ask a second question – whether more should be done to foster other roles for stress testing in oversight of the financial system. My third question links the topics of the prior two: To what extent might the dominance of the capital regulation role inhibit development of other useful roles.

To be clear at the outset: It is not my intention to attempt full answers to these questions. My aim, rather, is to prompt discussion – certainly among the members of this highly expert audience, but ideally involving a wide range of government officials, academics, policy commentators, and bank officials. In recognition of the fact that a good way to get a debate started is to offer a proposition that is at least slightly provocative, the observations I offer today are all related to the hypothesis that there is an ongoing conflict between the dynamism that lies at the heart of good stress testing and the inertial tendencies associated with the routinization of stress testing in setting normal time capital requirements.

Stress Testing in Setting Capital Requirements

As a forward-looking approach to setting capital requirements, stress testing has some obvious advantages over standardized approaches. It is not based on fixed, potentially outdated risk categories. It partially corrects for the stickiness of book values that can produce misleadingly favorable capital ratios. It factors in anticipated changes in bank revenues, as well as losses. It can take some account of the degree of loss correlations among asset classes. By

testing all large bank balance sheets simultaneously, it can give a more accurate picture of risks to the financial system as a whole.

Although, as these annual conferences demonstrate, there is still plenty of room to advance the technology of stress testing further, the Federal Reserve's annual stress test is immeasurably more refined than it was a decade ago. In this sense, the stress testing regime has indeed been a dynamic one over the years. The expert staff who administer the test certainly understand where additional improvement would be desirable, such as in calculating PPNR and in overcoming the procyclical "resistance" in the model to generating higher losses when the economy has been growing for a number of years. And, while it is disappointing that second-order effects such as the impact of fire sales have not yet been incorporated into the model, one has the sense that if the modelers were instructed to do so, they could make progress fairly quickly.¹

There is another, less linear form of dynamism needed if the potential of stress testing is to be realized – a more or less continual adaptation to external changes. This form of dynamism is central to stress testing's claim to be a superior approach to evaluating and setting capital standards. Optimally, the supervisory model will be regularly updated to take account of developments such as shifts in the risk characteristics of assets, including those that may result from changes in the composition of bank balance sheets. Dynamism in scenario design is arguably even more important. Here, after all, is the essence of the forward-looking nature of stress testing. The amount of resiliency required by the results of a stress test depends in no small part on the degree to which the scenario(s) have anticipated the actual stresses to which the financial system may someday be exposed.

On this dimension of dynamism, the use of stress testing in the calculation of normal time capital requirements has been less impressive. After a good many important adjustments in the model in the early years of CCAR, improvements and adaptations to changed economic circumstances appear to have slowed. Though it has not released the actual code of the supervisory model, the Fed has now revealed enough relevant information that the banks probably have a working approximation of the model itself. The resulting optimization of balance sheets that has likely occurred across banks may, in Lucas critique fashion, have undermined somewhat the robustness of stress test results.

Even at the level of theory, scenario design is a meaningful challenge in fashioning an effective stress test. As you all know, though, there is an especially significant limitation inherent in the scenarios used by the Fed and regulators in other jurisdictions. In normal times – that is, when the financial system is not under significant stress – the Fed must construct a severely adverse scenario that is wholly hypothetical. But neither the Fed nor anyone else knows the actual stress banks will one day face. How likely do you think it is that an annual stress test in 2003 would have posited a scenario in which there was a precipitous, nationwide decline in housing prices? Because the impact of different possible shocks on losses and revenues can vary enormously for banks, both individually and collectively, this matters a great deal.

Theoretically, a stress test could include multiple scenarios – more on that in a few minutes. But, if it utilizes only a single severely adverse scenario each year, the Fed is decidedly operating in the realm of third- or fourth-best. Under this constraint, the most the Fed can do to take account of the manifold possible sources of stress is to vary the scenario materially from year to year, so as to reflect the variety of shocks that might lead to serious economic and financial stress. In recent years the severely adverse scenario has changed somewhat, but not

enough to make the loss functions for asset classes vary as much as the losses that might be realized from differing economic shocks.

The Fed has derived the core elements of its severely adverse scenario from “conditions that characterize post-war U.S. recessions,” with primary reliance on the unemployment rate.² For the market shock component, the Fed uses what it calls a hybrid approach, which begins with key features of the 2008 market shock but can also include hypothetical elements reflecting perceptions of current risks.³ These are certainly justifiable starting points for constructing what we might characterize as a generic stress scenario. When the Board materially varies its add-on elements for each year’s test based on concerns with specific asset classes, such as commercial real estate, the result can be a reasonable appraisal of bank resiliency under serious though familiar conditions. When it doesn’t – that is, when the test becomes reasonably predictable – the CCAR or SCB process drifts closer to being a mere compliance exercise. But it is important to note that in neither case is this been a regime that tries to test the potential impact of tail events that have not been experienced before.

To some extent, of course, the reduced dynamism observed in recent years reflected policy decisions. But it’s also important to recognize that while the conceptual foundations of stress testing call for more dynamism in its models and scenarios, institutional factors push towards less. One obvious anti-dynamic factor is the sustained and substantial lobbying by banking interests. A second, less obvious factor may be institutional pressures within the Fed itself.

The desire of some banks for a less dynamic capital regime has been strongly in evidence almost since the inception of the annual CCAR exercise. They want to have the model so that they can anticipate more precisely the capital consequences of their portfolios, and perhaps

adjust them to optimize their requirements. With only a single severely adverse scenario, significant changes in that scenario and in the supervisory model mean that the required capital for a given portfolio of assets will vary over time. When capital requirements vary in ways that are not reasonably predictable, banks must maintain an additional cushion of capital above regulatory requirements, including the stress capital buffers. If they do not, they might find themselves compelled to reduce capital distributions when their required stress capital buffers grow.

One can regard this annual fluctuation in capital requirements as a feature, not a bug, of the CCAR and SCB regimes. But it is entirely rational for banks to seek to limit the dynamism of the stress test, since a more predictable capital requirement facilitates their capital planning and communications with shareholders and analysts. And even those of us who generally favor higher capital requirements can have some sympathy for the chief financial officers trying to manage capital planning and investor relations in this third- or fourth-best world in which their SCBs can go up significantly in one year and down in the next, depending on changes in the supervisory scenario.

Both directly and indirectly, some large banks expend considerable resources in an effort to limit the variability of capital requirements from year to year. In their regular interactions with agency officials, bank employees who are directly involved in the stress tests did, in my experience, make some valid criticisms, especially on technical matters. But the opacity and complexity of the stress testing process – not to mention the concentrated interests and substantial resources of banks compared to the more diffuse public interest – mean that countervailing pressures from non-bank interests and observers on these important technical points are likely to be considerably less significant.

This asymmetric external pressure to limit the dynamism of stress testing may be reinforced by internal inertial tendencies. Most people who have worked in larger organizations administering regulatory or other programs will recognize the tendency toward routinization, and perhaps some degree of ossification. Norms are established and routines develop for the good reasons of giving direction to the staff who administer a program and clarity to those benefitting from, or affected by, it. They provide staff some insulation from criticism of their role in administering the program, precisely because they are following established practice. Change, on the other hand, must be justified, either *ex ante* in trying to convince principals or *ex post* if principals or stakeholders disagree with a change that has been made.

For parallel reasons, and quite apart from policy preferences as such, the principals who ultimately have to answer for the program will often themselves favor considerable regularity in program administration, so that they are not surprised by staff innovations. Finally, after such a heavy investment by some of the best minds at the Fed in building out the stress testing regime, it was almost inevitable that other imperatives would begin to draw away some talent and energy.

I observed some of these tendencies even before I left the Board in 2017. While we were committed to an ongoing development of both stress testing and what was then the CCAR regime, it was clearly unrealistic as an organizational matter to expect, for example, a complete rethink of the scenarios every year. All organizations are subject to these inertial tendencies. But I think the impact on stress testing – premised as it is on dynamism and adaptability – is especially important.

Notwithstanding the loss of dynamism, from a strictly regulatory perspective the use of stress testing in capital regulation retains some advantages. It very likely remains more risk

sensitive than a standardized risk-weighting approach, even the much more detailed one that emerged from Basel III. And, while the predictability of the scenarios and release of more information on the models have made the stress test a less accurate measure of resiliency in absolute terms, it continues to provide supervisors with a better cross-firm comparison of resiliency levels than does the standardized approach. It may also be the case that conducting an annual test that carries consequences for capital requirements is the best way to ensure that both banks and supervisors keep their stress testing muscles well-toned.

Should we be satisfied, though, drifting into a fairly predictable SCB exercise that has only modest advantages over simply raising the fixed capital conservation buffer for larger banks above the generally applicable 2.5% level? Given the potential of stress testing – and all the work on it by supervisors, academics, consultants, and banks in the past decade – there is a good argument that we should be aiming higher.

But what, exactly, should we be aiming at? More dynamism, to be sure. To answer this question with specifics, though, we need first to answer an even more basic question. What level of resiliency is the SCB regime intended to produce in the nation's largest banks? There has occasionally been some loose, and obviously mistaken, talk suggesting that the stress test establishes the capacity of the banks to undergo even the most severe stress and remain viable financial intermediaries. In contrast, the Fed's official statements on this point are both careful and ambiguous. The release of results this past June stated that the stress test "assesses whether banks are sufficiently capitalized to absorb losses during a severe recession."⁴ In prior years the Board was even less specific, stating that the results demonstrated that "large firms have sufficient capital levels to absorb losses during stressful conditions"⁵ or in "an unfavorable economic environment."⁶

I would suggest that the level of resiliency that the SCB regime is intended to achieve should be more precisely stated. This is important in the first instance just as a matter of being straightforward with the public. As then-Governor Brainard pointed out, the approach to the SCB eventually adopted in early 2020 actually reduced required capital buffers. Even with the somewhat increased severity of this year's scenario, more than a third of the participating banks were assigned an SCB at the minimum 2.5% level required of all banks under current generally applicable capital standards. One might ask how this outcome squares with the Dodd-Frank requirement that capital standards be more stringent for banks with over \$250 billion in assets.

The fundamental question is how far out on the tail the regulatory capital requirements are intended to establish resiliency for systemically important banks. At present, the implicit resiliency standard is probably the capacity to remain a viable financial intermediary in the face of a serious, but generally familiar, shock and ensuing consequences. If there were a policy decision to seek greater resiliency to genuine tail events, a promising avenue would be to use multiple non-baseline scenarios – something that was not practically feasible in the early years of CCAR. Although there would obviously still be no guarantee that any of the constructed scenarios would correctly anticipate a future actual stress event, the Board could at least cover several quite different possibilities each year. Supervisors would have greater insight than at present into the kinds of shocks that might present serious risks to the financial system. They would have a better sense of the vulnerabilities of specific banks to different kinds of shocks.

Of course, the question that immediately arises is how the results of multiple scenarios would figure into the calculation of the SCB. I suspect many of you are already thinking of the options – such as using the highest loss scenario for each bank; or, perhaps, using the highest loss scenario for the tested banks as a group. That would be the subject for discussion once a

clear policy decision on the standard of resiliency has been made. This and other implementation decisions might be complicated, but there would be a principle to guide them.

Another reason to be clear about the resiliency standard of the SCB regime is that explicit articulation of this target helps place it in the broader context of financial stability policies. The less ambitious the level of resiliency, the greater the implicit reliance on other financial stability policies – not only prudential requirements such as liquidity rules, but also lender-of-last-resort and resolution policies.

Even if the policy choice were not to adopt a resiliency standard that covered a broader range of potential tail risks, using multiple scenarios could be a good way to inject more dynamism into the current SCB approach. The information generated by running additional scenarios would increase supervisory insight into potential risks faced by the financial system. Having to run these scenarios would also likely enhance the banks' capacities to project tail risks faced by their own institutions. With a less ambitious resiliency standard, the method for calculating the SCB would presumably be different than with a standard aiming at tail risk – perhaps some form of averaging of all results of the different scenarios.

Unfortunately, the Dodd-Frank stress test has headed in the opposition direction. In 2018, Congress removed the requirement for an “adverse” scenario to accompany the baseline and severely adverse scenarios. Regrettably, the Fed subsequently dropped the adverse scenario and indicated it was unlikely to use multiple non-baseline scenarios in the future. For several years the adverse scenario had been used to test a shock that, while of a lesser magnitude, was quite different in kind from that in the severely adverse scenario. Supervisors and banks, and perhaps some close outside observers such as bank analysts, thereby gained some additional information about bank vulnerabilities. Ironically, the adverse scenario used for those several

years posited sharp changes in interest rates – a scenario that would obviously have been very appropriate this year.

Other Roles of Stress Testing

Let me turn now, much more briefly, to my other two questions, beginning with whether there should be greater emphasis on current and potential roles of stress testing. There is a good case to be made that the highest and best use of stress testing is during a period of actual stress. In the incipient or early stages of stress that might develop into a crisis, supervisory and market need for information on the soundness of banks will be increasing. A stress test can help provide this information, as it most notably did in the first part of 2009, when the Fed conducted the Supervisory Capital Assessment Program (SCAP). Crisis stress testing does not require the Fed to hypothesize the source of stress. Instead, it extrapolates the impact of the conditions that have given rise to the stress, making the scenario much more likely to project the losses that will result if conditions deteriorate. A stress test in the early stages of a possible systemic event allows for timely changes in banks' capital policies to ensure their resiliency should the crisis develop as projected.

There can be variations on this general role of stress testing in the face of emerging or realized stress. Where circumstances are already serious, the Fed can undertake a full-blown special stress test, with the usual public disclosures, on the basis of which it can recalculate the SCBs of participating banks. This is the option that is contemplated in the Fed's regulations. But, with the machinery of stress testing now better developed at both the banks and the Fed, there is also the option of a quicker exercise that would provide supervisors with insight into

vulnerabilities that might be exposed by the developing strains. On the basis of that information, the Fed can decide on appropriate supervisory responses (perhaps including a full-blown test).

Stress testing may even have a role in monetary policy. Consider current circumstances, in which a series of sizeable increases in the federal funds rate has forced market actors to recalibrate everything from anticipated ROEs to the value of collateral to funding costs and availability. Of course, tighter financial conditions are an intended consequence of tighter monetary policy. But when conditions tighten so significantly so quickly, especially after an extended period of very low rates, there is some risk of *too* abrupt a market reaction. Such a reaction may expose leveraged firms to painful – even destabilizing – adjustments, which could ripple across the financial system. A modified stress test using scenarios based on significant interest rate and yield curve changes would usefully inform the FOMC about the potential financial stability effects of alternative paths for the removal of accommodation. Even where the greater vulnerabilities lie with financial firms outside the prudential regulatory perimeter, the results of a stress test could yield considerable insight into the effects of rapid federal funds rate increases on various asset classes.

My point here is that the role of stress testing can be tailored to the circumstances and information needs that the Fed faces at any given moment. There is a host of information-generating uses to which stress testing has, and could, be put within the broader responsibilities of regulators to ensure the stability of the financial system.

Have CCAR and the SCB Inhibited Other Valuable Roles of Stress Testing?

My third and final question is whether the dominance of the ongoing capital requirements role of stress testing has inhibited development and use of other roles for stress testing. Probably not where the results from the use of stress testing techniques or information will be available

only within the Fed – whether for supervisory, financial stability, or other purposes. I have come to suspect, though, that there may be some inhibition with respect to any use of stress testing whose results may become publicly known – either directly or indirectly.

One can readily see why this may be the case where the annual Dodd-Frank test is concerned. While there is a good argument for the Fed to experiment with its methodology as part of this exercise, the Board may feel that it will need to explain, defend, and possibly act upon the results. This may be true even where it is clear that the results of these variant scenarios or tests will not count in the calculation of SCBs.

So, for example, the Fed would gain a valuable perspective on its methodology by constructing a parallel exercise that used more market-based measures, rather than book value, as the starting point for a wider range of bank assets. Similarly, constructing a scenario that extrapolates a trend of fast growth in certain products or asset classes might help supervisors better see the strengths and shortcomings of current methodologies. But the Fed may want to avoid the potentially awkward situation where the results of an exercise of this sort suggested that capital should be higher than in the conventional DFA test.

To the degree there is something to my supposition, the Fed is foregoing information that would increase insight into potential tail risks faced by the financial system. The use of materially different scenarios or starting points for analysis could also prompt banks to improve their own capacities to project tail risks that are especially relevant for their own institutions.

My suspicion of some inhibiting effect is of course largely speculative. A more tangible piece of evidence in favor of this inference may be found in the organizational response of the Fed to the financial turmoil in March 2020. The annual stress test was just underway as the dimensions of the COVID crisis became clear. As was immediately apparent from the

staggering declines in employment and GDP, the stress scenario was not nearly as severe as the path the economy seemed to be on. Yet the Fed continued with what it admitted was a stale scenario, all the way to the announcement of results in June, as originally scheduled. Only then did it undertake what it called a “sensitivity analysis” of bank balance sheets that applied three possible scenarios for the COVID-afflicted economy. A public comment from the Board that there wasn’t time to restart the process and still meet the June deadline seemed odd, given that the Board acknowledged that the scenario was stale.

Was the Fed’s institutional commitment to its regular process so strong that it was reluctant to pivot to a special test – or even to the half measure of a sensitivity analysis which it eventually adopted – despite the unprecedented and massive nature of the shock that had struck the financial system? I don’t know the answer, but I do know that the three months that were lost might have been crucial. In the event, the extraordinary fiscal stimulus and Fed emergency programs protected the banking system. In other circumstances the failure to determine banks’ potential capital needs at an earlier stage may have made things much worse later.

Conclusion

Stress testing is arguably the most important supervisory innovation to have emerged from the Global Financial Crisis. Of necessity, today I have only scratched the surface of its potential. To realize its promise as part of an approach to setting capital requirements that is superior to point-in-time calculations, it needs to be more or less continually adapting to changes in assets, bank practices, and the economy. With the fresh look at capital requirements being conducted by the Board, perhaps there will be some opportunities for reinjecting some dynamism into the SCB process.

To realize its potential in crisis management, macroprudential policy, and perhaps monetary policy, stress testing needs to be at least partially liberated from the annual Dodd-Frank and SCB process. I hope that the Fed will find ways to foster the potential of stress testing in these other areas, so that future conference organizers will have an even wider array of interesting topics from which to choose.

¹ Willian F. Bassett and David Rappoport, “Enhancing Stress Tests by Adding Macroprudential Elements,” in J. Farmer, A. Kleinnijenhuis, T. Scheurmann and T. Wetzer, *Handbook of Financial Stress Testing* 455-483 (2022)

² Board of Governors of the Federal Reserve System, Amendments to Policy Statement on the Scenario Design Framework for Stress Testing, 84 Fed. Reg. 6651, 6658 (Feb. 28, 2019).

³ *Id.* at 6661-6662.

⁴ Board of Governors of the Federal Reserve System, 2022 Supervisory Stress Test Results (June 2022) at v, available at <https://www.federalreserve.gov/publications/files/2022-dfast-results-20220623.pdf>.

⁵ Board of Governors of the Federal Reserve System, Dodd Frank Act Stress Test 2021: Supervisory Stress Test Results (June 2021) at 3, available at <https://www.federalreserve.gov/publications/files/2021-dfast-results-20210624.pdf>.

⁶ Board of Governors of the Federal Reserve System, Dodd-Frank Stress Test 2016: Supervisory Stress Test Methodology and Results (June 2016) at 5, available at <https://www.federalreserve.gov/newsevents/pressreleases/files/bcreg20160623a1.pdf>.