

Stress Testing the Banking Agencies

Matthew C. Turk*

ABSTRACT: One of the major regulatory innovations that has emerged over the decade following the financial crisis is the development of regulatory stress tests for large financial institutions. But the role of stress tests as a pillar of financial regulation has been placed in jeopardy by a recent wave of reforms within Congress and the Trump Administration. Existing legal scholarship provides minimal guidance for evaluating this development, because it lacks a coherent account of what the Dodd–Frank Act’s stress testing programs can and should do. This Article fills that gap.

First, it provides a comprehensive analysis of the promise and limits of financial stress tests. That analysis reveals that both Dodd–Frank’s architects as well as its reformist skeptics have misconceived the vices and virtues of the post-crisis stress testing rules. As it stands, the current procedures bear surprisingly little relation to the systemic risks they were designed to address. At the same time, claims that those rules represent a harmful escalation of regulatory burdens, discretion or uncertainty are overstated.

Second, the Article moves beyond critique and charts a practical path forward by identifying a simple yet fundamental twist to the administration of stress tests which would enable them to effectively perform the functions they were intended to serve. Specifically, it outlines a set of reforms that transform stress tests into tools for diagnosing weaknesses in the regulatory requirements promulgated by federal banking agencies, rather than in the banks themselves. By stress testing for regulatory failure, the market failures which lead to financial crises are more likely to be prevented.

The broader contribution of this Article is to highlight the need for a genuinely interdisciplinary approach to financial regulation, which focuses on how subtle aspects of legal structure interact with the underlying economic

* Assistant Professor of Business Law, Indiana University - Kelley School of Business. I am grateful for valuable comments on earlier drafts by Bill Bratton, Vince Buccola, Nakita Cuttino, George Georgiev, Jill Fisch, Gina-Gail Fletcher, Merritt Fox, Todd Haugh, Claire Hill, Jeremy Kress, Anita Krug, Josh Macey, Jeremy McClane, Michael Ohlrogge, Frank Portnoy, Christina Skinner, and Robert Thompson. I would also like to thank participants at the Canadian Law & Economics Association Annual Conference, the Academy of Legal Studies in Business Annual Conference, the National Business Law Scholars Conference, and the Indiana Maurer School of Law workshop on Financial Regulation and Innovation. For excellent research assistance, I thank Peter Gauss.

principles governing financial markets. The post-crisis stress tests present a classic case on why taking both the law and economics of financial regulation seriously is easier said than done. But they also show that without such an approach, regulatory costs and benefits are misapprehended, basic policy questions prove impossible to answer, and unintended consequences abound.

I.	INTRODUCTION.....	1703
II.	LEGAL & HISTORICAL BACKGROUND ON FINANCIAL STRESS TESTS	1707
A.	STRESS TESTING FROM 1987–2008.....	1707
1.	At Firms.....	1707
2.	In Regulations	1708
B.	THE FINANCIAL CRISIS STRESS TEST.....	1710
C.	THE DODD–FRANK ACT STRESS TESTING RULES	1713
D.	STRESS TESTS RESULTS FROM 2012–2018.....	1715
E.	RECENT PROPOSALS TO ROLL BACK THE DODD–FRANK RULES.....	1717
1.	Executive Branch Reforms	1717
2.	Statutory Reforms.....	1718
III.	EVALUATING THE PROMISE AND LIMITS OF STRESS TESTS: AN UNFINISHED TASK	1719
A.	THE INCOMPLETE CASE AGAINST.....	1719
1.	Quantitative Skepticism.....	1719
2.	Regulatory Discretion, Opacity, & Uncertainty	1722
3.	Compliance Costs.....	1725
B.	THE INCOMPLETE CASE FOR	1729
1.	The Information-Producing Function.....	1729
i.	Information for Markets.....	1729
a.	Stress Testing in Wartime	1730
b.	Stress Testing in Peacetime	1733
ii.	Information to Banks.....	1736
iii.	Information to Regulators.....	1738
a.	Macro-Surveillance	1738
b.	Regulatory Learning	1740
2.	Safeguards & Stabilizes Banks	1742
3.	Complements other Regulations	1745
IV.	A PROPOSAL: STRESS TESTING THE FEDERAL BANKING AGENCIES	1747
A.	SUMMARY OF PROPOSAL	1748
1.	Methodological Changes	1748
2.	A New Role for Banks	1750

3.	A New Role for Regulators	1752
B.	ADVANTAGES OF THE PROPOSAL.....	1754
1.	Addressing Critiques of Stress Testing	1754
2.	Making Good on the Purported Benefits of Stress Testing	1756
V.	CONCLUSION	1757

I. INTRODUCTION

Over the decade since the collapse of Lehman Brothers in 2008, financial regulation has undergone its greatest transformation since the Great Depression. During that period, one legal development that stands out as particularly novel and important is the emergence of banks stress tests as regulatory tools.¹ Indeed, the rapid rise of regulatory stress testing is difficult to overstate. The use of those procedures was essentially non-existent until first introduced by the Federal Reserve on an emergency basis in early 2009.² Yet, due to the perceived success of that experiment, regulatory stress tests were soon after enshrined as a cornerstone of the post-crisis legal architecture established pursuant to the Dodd–Frank Act.³

Within the past few years, however, the role of stress tests has come under attack from a wave of reforms which call for those procedures to be rolled back in substantial part or eliminated in full. One source of that pushback is an Executive Order issued by the Trump Administration in 2017, which instructed the federal banking agencies to undertake a comprehensive review of post-crisis financial regulations.⁴ In response, the Treasury Department has released a series of policy memoranda which set forth a detailed roadmap for

1. Generally speaking, stress tests are statistical simulations which estimate the impact of a potentially adverse economic scenario on a bank's balance sheet and overall financial stability. See *infra* Section II.A (providing background on the origins of financial stress tests); see also Chester S. Spatt, *Regulatory Conflict: Market Integrity vs. Financial Stability*, 71 U. PITT. L. REV. 625, 627 (2010) ("One of the most interesting innovations to emerge in the bank supervision model during the financial market crisis is the use of stress tests.").

2. See *infra* Section II.B.

3. Dodd–Frank Wall Street Reform and Consumer Protection Act, Pub. L. No. 111–203, § 165, 124 Stat. 1376, 1430–31 (2010); see *infra* Section II.C; see also Patrick Jenkins & Brooke Masters, *Banks: Again Under Strain*, GONZALO RAFFO INFONEWS (July 7, 2011, 8:29 PM), <http://gonzoraffoinfonews.blogspot.com/2011/07/banks-again-under-strain-financial.html> [<https://perma.cc/AX5H-YKZ5>] ("[S]ince the financial crisis . . . stress-testing [has] become a vital part of the regulatory arsenal.").

4. See Exec. Order No. 13,772, 82 Fed. Reg. 9965 (Feb. 8, 2017).

overhauling the current stress testing requirements under Dodd–Frank.⁵ Congress has been equally active on legislation to the same effect.⁶

This all raises an obvious and important question: Does the ongoing push to aggressively reverse the post-crisis expansion of regulatory stress tests make sense? Surprisingly, existing legal scholarship has little to say on that point.⁷ Although a modest law literature has developed around the specific subject of stress testing banks, those papers are largely descriptive and focus on formalistic legal aspects of the Dodd–Frank rules.⁸ To the extent that commentary turns to normative policy evaluation, the discussion tends to proceed in survey style, cataloguing the most prominent arguments without engaging them in sufficient depth to reach concrete conclusions. Nor has legal scholarship integrated the vast literature in financial economics on stress testing⁹—at least not in a way that would allow the discrete, technical

5. See STEVEN T. MNUCHIN & CRAIG S. PHILLIPS, U.S. DEP'T OF TREASURY, A FINANCIAL SYSTEM THAT CREATES ECONOMIC OPPORTUNITIES: BANKS AND CREDIT UNIONS 3 (2017) [hereinafter TREAS. BANK'G REP.], available at <https://www.treasury.gov/press-center/press-releases/Documents/A%20Financial%20System.pdf> [<https://perma.cc/3RA5-3G6H>]; see also Pete Schroeder, *Federal Reserve to Re-Examine U.S. Bank Stress Tests at July Conference*, REUTERS (Jan. 28, 2019, 1:04 PM), <https://www.reuters.com/article/us-usa-fed-stresstests/federal-reserve-to-re-examine-us-bank-stress-tests-at-july-conference-idUSKCN1PM28Q> [<https://perma.cc/LNQG-QEX8>].

6. See Economic Growth, Regulatory Relief, and Consumer Protection Act, Pub. L. No. 115–174, § 401(e), 132 Stat. 1296, 1359 (2018); see also Financial CHOICE Act of 2017, H.R. 10, 115th Cong. (introducing a bill to “repeal[] the provisions of the Dodd–Frank Act”); cf. Jeff Cox, *House Passes Choice Act That Would Gut Dodd–Frank Banking Reforms*, CNBC, <https://www.cnbc.com/2017/06/08/house-has-votes-to-pass-choice-act-that-would-gut-dodd-frank-banking-reforms.html> [<https://perma.cc/7RH8-CVUK>] (last updated June 8, 2017, 7:09 PM) (predicting that the Financial CHOICE Act would likely fail in the Senate).

7. This, despite the Dodd–Frank stress tests being far from an obscure legal topic. A Westlaw search of law journal publications indicates that the financial stress tests have received at least passing reference in over 7,000 articles since January 1, 2009.

8. For articles that focus on the stress tests specifically, see generally Mehrsa Baradaran, *Regulation by Hypothetical*, 67 VAND. L. REV. 1247 (2014); James F. Powers III, *Stress Testing Under Dodd–Frank: Easing the Regulatory Burden for Midsize Financial Companies*, 20 N.C. BANKING INST. 361 (2016); Margaret Ryznar, Frank Sensenbrenner & Michael Jacobs, Jr., *Implementing Dodd–Frank Act Stress Testing*, 14 DEPAUL BUS. & COM. L.J. 323 (2016); Robert F. Weber, *The Comprehensive Capital Analysis and Review and the New Contingency of Bank Dividends*, 46 SETON HALL L. REV. 43 (2015) [hereinafter *Comprehensive Capital Analysis*]; and Robert Weber, *A Theory for Deliberation-Oriented Stress Testing Regulation*, 98 MINN. L. REV. 2236 (2014). For articles that address stress tests or related issues in substantial part, see generally John Crawford, *Wargaming Financial Crises: The Problem of (In)Experience and Regulator Expertise*, 34 REV. BANKING & FIN. L. 111 (2014); Ronald J. Gilson & Reinier Kraakman, *Market Efficiency After the Financial Crisis: It's Still a Matter of Information Costs*, 100 VA. L. REV. 313 (2014); Henry T.C. Hu, *Disclosure Universes and Modes of Information: Banks, Innovation and Divergent Regulatory Quests*, 31 YALE J. REG. 565 (2014); and Eugene A. Ludwig, *Assessment of Dodd–Frank Financial Regulatory Reform: Strengths, Challenges, and Opportunities for a Stronger Regulatory System*, 29 YALE J. REG. 181 (2012).

9. The volume of research in this area is so extensive that no handful of references would do it justice. According to one estimate, there has been on average one full-length book published per month on the statistical methods associated with financial stress testing since 1987. Andrew G. Haldane, Exec. Dir., Fin. Stability, Bank of Eng., Speech at the Marcus-Evans

questions which frame research in that field to inform the broader dilemmas of regulatory-and-institutional design that must now be confronted.

This Article remedies that gap by presenting a comprehensive analysis of the promise and limits of bank stress testing. Such an analysis turns out to be sorely needed. As it stands, there is no coherent critique of the post-crisis stress tests to support the reformist push for a fundamental overhaul.¹⁰ At the same time, none of the leading justifications for the Dodd–Frank stress testing procedures make a persuasive case that those rules contribute to a more stable or efficient financial system.¹¹ The Article then shows a way out of that intellectual impasse by providing pragmatic policy guidance for moving forward. It argues that, despite shortcomings of the current rules, a useful role can be salvaged for the Dodd–Frank stress tests. The specific proposal consists of a simple yet fundamental twist: Reorient those procedures 180-degrees, so that they test the federal banking agencies and their rules, not the banks.

The basic intuition behind the proposal is as follows.¹² When a stress test result indicates that a bank (which is otherwise in full compliance with supervisory standards) is so fragile that it would pose systemic risks in the event of a financial downturn: Who, really, has failed this test? Under the implicit logic of Dodd–Frank, the auto manufacturer has always erred when the crash-test dummy meets an ugly fate, never the airbag or seatbelt mandates designed to prevent that outcome.¹³ But if banks are legally authorized to take risks which benefit their shareholders while destabilizing the system as a whole, surely the regulatory structure itself has been exposed as inadequate. As will be shown, by transforming stress tests into a tool that calibrates the banking agencies’ regulatory requirements, rather than micromanages the balance sheets of particular financial institutions, the result is to restore the benefits that the current rules purport to provide, but do not deliver.

A broader ambition of this Article is to advance a more dialectical approach to financial regulation, which views policy questions as essentially a mapping exercise, where progress is made by connecting the economic dynamics underlying financial markets to the legal design of the regulatory structure at a granular level. It may sound trite to observe that both the law and the economics matter, especially when it comes to the financial system.

Conference on Stress-Testing: Why Banks Failed the Stress Test 2 (Feb. 9, 2009), *available at* <https://www.bis.org/review/rogo219d.pdf> [<https://perma.cc/ED4E-4USU>].

10. For a critical review of these arguments, see *infra* Section III.A.

11. For a critical review of these arguments, see *infra* Section III.B.

12. The details of the proposal are spelled out in Part IV.

13. The crash-test analogy is open to interpretation and only meant to be suggestive of the questions that motivates this Article’s analysis, rather than supplying the content of the argument itself. One interpretation that would be inapposite is that the shift of focus from the banks to regulators means that the proposal has an inherently “de-regulatory” posture. In the crash-test hypothetical, a “failed test” would mean that regulators should impose stricter safety requirements on the car companies.

But that lesson is put into practice less frequently than might be assumed and, more often than not, the two perspectives are sprinkled together in a superficial fashion which squanders the strong complementarities they share. The sub-field known as “law-and-finance,” which was explicitly conceived as an interdisciplinary research agenda, serves as a cautionary tale that is instructive on this point.¹⁴ While the first generation of studies in that literature were hailed as a powerful synthesis that generated wide-ranging insights, the subsequent scholarly debate has since been consumed by claims that the findings from those canonical studies are moot because the relevant legal or economic variables were misinterpreted in various ways.¹⁵

This difficulty in arriving at reliable policy conclusions stems from both sides of the law-and-finance coin. On one hand, the devil is never in the legal details. Unless the policy analysis constantly circles back to economic first principles, the legal materials can be documented ad nauseam without yielding any clue as to how those rules actually function in practice. On the other hand, however, legal minutia is critical. Core social science principles remain arid textbook abstractions until they are filtered through a rich account of the relevant legal context where their assumptions are thought to apply. An upshot of this dilemma is that subtle features of the financial or institutional environment at issue can change everything.

14. The law-and-finance literature adopts a comparative approach which asks how the variation in legal rules across different jurisdictions impacts outcomes in their financial markets. For the seminal law-and-finance publications by economists, see generally Rafael La Porta et al., *Legal Determinants of External Finance*, 52 J. FIN. 1131 (1997) (noting that countries with poorer investors protections have smaller and narrower capital markets); and Rafael La Porta et al., *Law and Finance*, 106 J. POL. ECON. 1113 (1998) (examining legal protections of investors in 49 countries). For some important contributions by legal scholars, see generally Dan Awrey, *Law and Finance in the Chinese Shadow Banking System*, 48 CORNELL INT'L L.J. 1 (2015) (using legal theory of finance to explore the emergence, growth, and risks of the Chinese shadow banking system); Emiliano M. Catan & Marcel Kahan, *The Law and Finance of Antitakeover Statutes*, 68 STAN. L. REV. 629 (2016) (evaluating studies of the effect of Antitakeover statutes on firm and managerial behavior); and Katharina Pistor, *A Legal Theory of Finance*, 41 J. COMP. ECON. 315 (2013) (developing a legal theory of finance for contemporary financial systems).

15. See Catan & Kahan, *supra* note 14, at 632 (concluding that, contrary to large body of law-and-finance research identifying the impacts of state antitakeover statutes, “[o]ur analysis is consistent with the view that antitakeover statutes do not matter after all”); Ulrike Malmendier, *Law and Finance “at the Origin,”* 47 J. ECON. LITERATURE 1076, 1077 (2009) (“The Roman evidence illustrates the limitations of the existing law and finance theories. In the case discussed here, legal restrictions (or the lack of legal development) per se appear to matter little as long as the law as practiced is flexible and adapts to economic needs.”). See generally Holger Spamann, *The “Antidirectors Rights Index” Revisited*, 23 REV. FIN. STUD. 467 (2010) (exploring the accuracy of index values); Gerhard Schnyder et al., *Twenty Years of ‘Law and Finance’: Time to Take Law Seriously* (Ctr. for Bus. Research, Univ. of Cambridge, Working Paper No. 501, 2018) (providing the first comprehensive discussion of the first twenty years of the Law and Finance School literature), available at https://www.cbr.cam.ac.uk/fileadmin/user_upload/centre-for-business-research/downloads/working-papers/wp501.pdf [https://perma.cc/SLD6-SEG2].

The post-crisis stress tests provide a classic case study on this theme and epitomize how taking the law-and-finance interface seriously is more elusive than it seems. They also illustrate why such an approach is nonetheless indispensable and can overturn many key assumptions and received wisdom when carefully applied. To preview some examples:

- Regulatory stress tests in their current form do not measure systemic risk. As designed, they presuppose no such thing exists;¹⁶
- They also rarely produce new information that markets, banks, or regulators find useful. For some of those audiences, they produce no new information at all;¹⁷
- Complaints about the lack of transparency in the Dodd–Frank stress test procedures are a red herring. Along the most important dimensions, those procedures should aim for maximal opacity;¹⁸
- So too are concerns that the post-crisis rules represent a dramatic escalation in regulatory discretion, interventionism, and uncertainty. In substance, the Dodd–Frank stress tests closely mimic existing banking regulations which have been in place for decades (some since the civil war);¹⁹

and so on. These claims should not be as controversial as they might appear at first glance. To the extent they seem implausible or willfully contrarian, this Article highlights the potential that a genuine law-and-finance perspective has to demystify fundamental points of confusion on important policy questions.

The discussion below proceeds as follows. Part II provides background on financial stress tests. Part III evaluates arguments for-and-against the current stress testing procedures. Part IV presents this Article's proposal. A final Part briefly concludes.

II. LEGAL & HISTORICAL BACKGROUND ON FINANCIAL STRESS TESTS

A. STRESS TESTING FROM 1987–2008

1. At Firms

Stress testing was originally adopted by financial firms on a voluntary basis as an internal risk assessment device. A formative moment came with the stock market crash of 1987, which served as a wakeup call for investors, most

16. See *infra* Section III.B.1.

17. See *infra* Section III.B.1.

18. See *infra* Sections III.A.2, IV.A.

19. See *infra* Sections III.A.2, III.B.2.

of whom had been working under the assumption that such a severe fluctuation was not possible. As one response, the CEO of JP Morgan, Dennis Weatherstone, began to ask for a “4:15 Report”—a daily memo which the trading desk must produce within 15 minutes of the market close at 4 PM, which would estimate the bank’s largest potential market loss for the next day.²⁰

At the time, the “4:15 Report” was seen as an extravagant request, but the inability to formulate a quick answer to such a critical question prompted what came to be known as Value-at-Risk (“VaR”) models.²¹ In its simplest form, VaR modeling is a simple aggregation exercise. The analyst looks up the largest single-day price drop that has occurred in recent history for each kind of asset the bank holds and then adds up those losses to calculate the maximum expected exposure for the portfolio as a whole.²² That figure represents the bank’s VaR.

Soon after the initial use of VaR models at institutions such as JP Morgan, the technical complexity of financial risk assessments grew at an exponential rate.²³ It is therefore possible to draw up endless taxonomies to distinguish among those procedures and debate which of them can be considered a *bona fide* “stress test.”²⁴ But without putting too much strain on definitional categories, it is fair to say that “quantitative risk management” of some sort has been a staple at financial firms since the late 1980s.

2. In Regulations

Stress testing (in the broad sense) entered financial regulations at a surprisingly early stage. The first examples appeared in housing finance. In 1988, the Federal Reserve issued a guidance document, “Thrift Bulletin #12,” which directed the Federal Home Loan Banks to perform a “sensitivity analysis” on their mortgage portfolios.²⁵ Legislation from 1992 included a mandate that Fannie Mae and Freddie Mac address the possibility of changes in interest rates and other key variables when submitting financial projections

20. See Barry Eichengreen, *The Last Temptation of Risk*, NAT’L INTEREST, May 2009, at 8, 8–9.

21. *Id.* at 9. See generally Robert F. Weber, *An Alternative Story of the Law and Regulation of Risk Management*, 15 U. PA. J. BUS. L. 1005 (2013) (reviewing the legislative and regulatory history of risk management).

22. See MICHEL CROUHY, DAN GALAI & ROBERT MARK, *THE ESSENTIALS OF RISK MANAGEMENT* 157–58 (2006).

23. See Haldane, *supra* note 9, at 2.

24. See Weber, *A Theory for Deliberation-Oriented Stress Testing Regulation*, *supra* note 8, at 2250–68 (taking up that task).

25. Investment Portfolio Policy and Accounting Guidelines, 54 Fed. Reg. 23,457, 23,461, 23,471 (May 19, 1989) (to be codified at 12 C.F.R. pts. 563c, 571); FED. HOME LOAN BANK SYS., THRIFT BULL. NO. 12, MORTGAGE DERIVATIVE PRODUCTS & MORTGAGE SWAPS 2–3 (1988); see also Weber, *A Theory for Deliberation-Oriented Stress Testing Regulation*, *supra* note 8, at 2280–81.

to their regulatory supervisor, then known as the Office of Federal Housing Enterprise Oversight.²⁶

For traditional deposit-taking banks, the first stress testing requirements were introduced through guidance from the Office of the Comptroller of the Currency in 1993, which required bank managers to “facilitate stress testing” within their derivatives trading departments.²⁷ Similar mandates trickled up to the international level as early as 1996, when the Basel Committee on Banking Supervision (“Basel Committee”)—a forum of financial regulators from the United States and other advanced economies—amended its cross-border banking protocols, known as the “Basel I” rules.²⁸ By the time the Basel Committee finalized an updated set of “Basel II” rules in 2004, its directive that globally-active banks maintain a “sound stress testing process[]” merely reiterated what was already a commonplace supervisory standard in domestic financial regulations.²⁹

A final noteworthy episode in the development of stress testing was spearheaded by the International Monetary Fund (“IMF”) in the aftermath of a series of sovereign currency and debt crises which swept across Latin America and East Asia during the late 1990s.³⁰ Within the next few years, the

26. See Federal Housing Enterprises Financial Safety and Soundness Act of 1992, 12 U.S.C. §§ 4511, 4611 (2012); Housing and Community Development Act of 1992, Pub. L. No. 102-550, §§ 1311, 1361, 106 Stat. 3672, 3944, 3972; see also Press Release, Fed. Hous. Fin. Agency, OFHEO Issues Risk-Based Capital Stress Test Results for Fannie Mae and Freddie Mac (June 27, 2002), available at <https://www.fhfa.gov/Media/PublicAffairs/Pages/OFHEO-Issues-Risk-Based-Capital-Stress-Test-Results-for-Fannie-Mae-and-Freddie-Mac.aspx> [https://perma.cc/E45W-9DCF].

27. OFFICE OF THE COMPTROLLER OF THE CURRENCY, BANKING CIRCULAR NO. 277, RISK MANAGEMENT OF FINANCIAL DERIVATIVES 9 (1993), available at <https://www.occ.gov/static/news-issuances/bulletins/pre-1994/banking-circulars/bc-1993-277.pdf> [https://perma.cc/SE7V-EW84]. In 1996, the Office of the Comptroller of the Currency (OCC) followed up with another push toward stress testing oversight, designed “to evaluate the bank’s exposure in a highly stressed market scenario.” OFFICE OF THE COMPTROLLER OF THE CURRENCY, BULL. NO. 1996-43, CREDIT DERIVATIVES: GUIDELINES FOR NATIONAL BANKS (1996), available at <https://www.occ.treas.gov/news-issuances/bulletins/1996/bulletin-1996-43.html> [https://perma.cc/2D68-V6P5].

28. See BASEL COMM. ON BANKING SUPERVISION, BANK FOR INT’L SETTLEMENTS, AMENDMENT TO THE CAPITAL ACCORD TO INCORPORATE MARKET RISKS 1 (1996), available at <https://www.bis.org/publ/bcbs24.pdf> [https://perma.cc/7UUE-DMAR] (amending the BASEL COMM. ON BANKING SUPERVISION, BANK FOR INT’L SETTLEMENTS, INTERNATIONAL CONVERGENCE OF CAPITAL MEASUREMENT AND CAPITAL STANDARDS (1988) (Basel I)); see generally DANIEL K. TARULLO, BANKING ON BASEL: THE FUTURE OF INTERNATIONAL FINANCIAL REGULATION (2008) (providing a brief history of Basel I, as well as its successes and failures).

29. BASEL COMM. ON BANKING SUPERVISION, BANK FOR INT’L SETTLEMENTS, INTERNATIONAL CONVERGENCE OF CAPITAL MEASUREMENT AND CAPITAL STANDARDS: A REVISED FRAMEWORK 89 (2004), available at <https://www.bis.org/publ/bcbs107.pdf> [https://perma.cc/F6Vg-QTTG] (Basel II).

30. See Winfrid Blaschke et al., *Stress Testing of Financial Systems: An Overview of Issues, Methodologies, and FSAP Experiences* 3 (Int’l Monetary Fund, Working Paper No. 01/88, 2001), available at <https://www.imf.org/external/pubs/ft/wp/2001/wp0188.pdf> [https://perma.cc/Z795-YFG4]; Marco Sorge, *Stress-Testing Financial Systems: An Overview of Current Methodologies* 1 (Bank for Int’l Settlements, Working Paper No. 165, 2004), available at <https://www.bis.org/publ/work165.pdf> [https://perma.cc/SK4G-QC4U].

IMF began to implement stress tests of member-countries as part of its “Financial Stability Assessment Program.” The Financial Stability Assessment Program is rarely referenced by commentary on the post-crisis stress tests under Dodd–Frank, perhaps because the IMF intended it as a form of administration surveillance for developing economies with relatively unsophisticated financial systems. But it provides a valuable template for the proposal introduced later in this Article. The key feature of the Financial Stability Assessment Program, in contrast to the more familiar stress testing requirements, is that it looks beyond the portfolios of individual financial institutions and attempts to gauge the stability of a country’s financial sector as a whole.³¹

Despite the extensive list of initiatives surveyed above, stress testing remained at the margins of U.S. banking regulation prior to the financial crisis of 2008. For one, regulators did not play a direct role in the stress testing process itself, which was entirely carried out at the bank level. Neither did bank supervisors undertake an effort to monitor how rigorously those tests were run nor verify the reliability of the results: It is difficult to find an instance where there was a determination that a bank had fallen short of meeting the vague standards that were in place and a concrete sanction imposed as a result.³² One of the best summaries of the pre-crisis landscape for stress testing can be found in a Government Accounting Office report on the subject from 2006, which concluded that “there was neither a well-developed set of best practices nor supervisory guidance in this area at the time.”³³ Internal stress testing of some sort was assumed to be an industry best practice, but anything a bank could shoehorn under that label was considered acceptable. Thus, regulatory supervision of stress testing during the pre-crisis period largely consisted of a formalistic box-checking exercise which left the relevant legal requirements toothless as a practical matter.³⁴

B. THE FINANCIAL CRISIS STRESS TEST

A turning point in the evolution of stress tests came while the financial crisis was still ongoing. On February 10, 2009, the Obama Administration’s

31. Blaschke et al., *supra* note 30, at 8.

32. U.S. GOV’T ACCOUNTABILITY OFFICE, GAO-09-499T, FINANCIAL REGULATION: REVIEW OF REGULATORS’ OVERSIGHT OF RISK MANAGEMENT SYSTEMS AT A LIMITED NUMBER OF LARGE, COMPLEX FINANCIAL INSTITUTIONS 23 (2009); *see also* COMM. ON THE GLOB. FIN. SYS., BANK FOR INT’L SETTLEMENTS, STRESS TESTING AT MAJOR FINANCIAL INSTITUTIONS: SURVEY RESULTS AND PRACTICE 3 (2005), *available at* <https://www.bis.org/publ/cgfs24.pdf> [<https://perma.cc/3FC9-EQ7U>] (reflecting similar findings made by the Basel Committee).

33. U.S. GOV’T ACCOUNTABILITY OFFICE, *supra* note 32, at 23.

34. *See Lessons Learned in Risk Management Oversight at Federal Financial Regulators: Hearing Before the Subcomm. on Sec., Ins., & Inv. of the S. Comm. on Banking, Hous., & Urban Affairs*, 111th Cong. 8–9 (2009) (statement of Timothy W. Long, Senior Deputy Comptroller, Bank Supervision Policy and Chief National Bank Examiner, Office of the Comptroller of the Currency) (acknowledging the deficiencies of the box-checking posture of stress test oversight).

incoming Treasury Secretary, Tim Geithner, used one of his first public addresses to announce a new “Capital Assistance Program,” part of which involved the Federal Reserve running its own stress tests on “each of the major U.S. banking institutions.”³⁵ The 2009 stress test—eventually to become known as the Supervisory Capital Assessment Program (“SCAP”)—was improvised within the executive branch as an ad hoc crisis management measure. Its primary legal basis was as an extension of the TARP bailout package authorized by Congress under the Emergency Economic Stabilization Act of 2008 after the failure of Lehman Brothers and AIG, rather than any pre-existing banking regulations.³⁶

To administer the SCAP, the Federal Reserve formulated a pair of hypothetical scenarios (one “baseline,” the other more “adverse”), which stipulated a set of economic indicators regarding unemployment, GDP growth, and so on, for the next two years.³⁷ The banks then had to use that data to project their losses, revenue, loan reserves, and capital levels under those conditions.³⁸ A bank would “fail” the test if the results indicated it would fall short of its minimum capital requirements under either scenario, specifically, a risk-weighted leverage ratio of four-to-six percent.³⁹ Banks which failed to meet that threshold were given two options to fill the gap. They had to either: (1) raise the required equity from markets by November 2009; or, (2) if private funding proved unavailable, request a publicly-funded capital

35. Press Release, U.S. Dep’t of the Treasury, Joint Statement by Sec’y of the Treasury Timothy F. Geithner, Chairman of the Bd. of Governors of the Fed. Reserve Sys. Ben S. Bernanke, Chairman of the Fed. Deposit Ins. Corp. Sheila Bair and Comptroller of the Currency John C. Dugan, and Dir. of the Office of Thrift Supervision John M. Reich—Financial Stability Plan (Feb. 10, 2009), *available at* <https://www.treasury.gov/press-center/press-releases/Pages/tg21.aspx> [<https://perma.cc/KL7W-QXJJ>]. The federal banking agencies provided further clarity on February 25, with a series of releases which laid out specific terms and indicated that participating institutions would consist of the country’s 19 largest bank holding companies. Press Release, Bd. of Governors of the Fed. Reserve Sys., Agencies to Begin Forward-Looking Economic Assessments (Feb. 25, 2009), *available at* https://fraser.stlouisfed.org/files/docs/historical/fct/frsbog/fedres_pressrelease_20090225.pdf [<https://perma.cc/TYA6-NAFU>] (appending an “FAQ” document); *see also* Edmund L. Andrews & Eric Dash, *Government Offers Details of Bank Stress Test*, N.Y. TIMES (Feb. 25, 2009), <https://www.nytimes.com/2009/02/26/business/economy/26banks.html> [<https://perma.cc/7VZT-VRF9>].

36. *Supervisory Capital Assessment Program & Capital Assistance Program*, U.S. DEP’T OF THE TREASURY, <https://www.treasury.gov/initiatives/financial-stability/TARP-Programs/bank-investment-programs/scap-and-cap/Pages/overview.aspx> [<https://perma.cc/F8T5-VDKD>] (last updated Dec. 9, 2013, 4:41 PM).

37. BD. OF GOVERNORS OF THE FED. RESERVE SYS., THE SUPERVISORY CAPITAL ASSESSMENT PROGRAM: DESIGN AND IMPLEMENTATION 5 (2009), *available at* <https://www.federalreserve.gov/bankinforeg/bcreg20090424a1.pdf> [<https://perma.cc/Y3Q6-LXY8>].

38. *Id.* A team of over one hundred financial supervisors and economists were recruited to adjust those figures as appropriate and evaluate the banks’ performance against a common benchmark. *Id.* at 10–11.

39. *See id.*

injection directly from the government, backed by the sale of certain preferred securities to the Federal Reserve.⁴⁰

When the Federal Reserve released the SCAP results on May 7, 2009, it took what was considered an extraordinary step by publicly disclosing an extensive portion of the bank's projections as well as the underlying stress testing methodology used to compute them.⁴¹ It also announced that 10 of the 19 participating institutions had failed.⁴² Despite the low passage rate of participating banks, the market response upon the release of the SCAP results was positive.⁴³ And due to improving economic conditions over the course of 2009, only one of the failing institutions was forced to receive funding from the Fed.⁴⁴ As a result, the SCAP was therefore generally viewed as a success at the time.⁴⁵ It remains widely credited as a major source of stability that restored confidence to a fragile financial system which had not yet exited the crisis atmosphere of 2008.⁴⁶

40. *Id.* at 2.

41. CONG. OVERSIGHT PANEL, JUNE OVERSIGHT REPORT: STRESS TESTING AND SHORING UP BANK CAPITAL 3–4 (2009) [hereinafter CONG. STRESS TEST REP.] (“Typically, . . . bank supervisory examination results are kept strictly confidential. . . . [B]ut, because the [SCAP] stress tests were undertaken in order to restore confidence in the banking system, they included an unprecedented release of information.”).

42. BD. OF GOVERNORS OF THE FED. RESERVE SYS., THE SUPERVISORY CAPITAL ASSESSMENT PROGRAM: OVERVIEW OF RESULTS 3 (2009), available at <https://www.federalreserve.gov/newsevents/files/bcreg20090507a1.pdf> [<https://perma.cc/3N6P-SB9L>]. As part of those results, the Federal Reserve estimated that the total capital shortfall those institutions needed to make up was roughly seventy-five billion dollars. *Id.*

43. See, e.g., *After the Financial Stress Tests: Relief but Still Some Uncertainty*, CNBC, <https://www.cnbc.com/id/30640189> [<https://perma.cc/9R3G-GDLL>] (last updated Aug. 5, 2010, 12:07 PM).

44. Press Release, Bd. of Governors of the Fed. Reserve Sys., Federal Reserve Board Makes Announcement Regarding the Supervisory Capital Assessment Program (SCAP) (Nov. 9, 2009), available at <https://www.federalreserve.gov/newsevents/pressreleases/bcreg20091109a.htm> [<https://perma.cc/X6LL-TXY2>].

45. Gary Gorton, *Stress for Success: A Review of Timothy Geithner's Financial Crisis Memoir*, 53 J. ECON. LITERATURE 975, 977 (2015) (“The stress tests of the largest banks are widely viewed as a great success.”); CONG. STRESS TEST REP., *supra* note 41, at 50 (concluding that “[t]he short-term effect of the stress tests was positive, and the financial markets have calmed to some extent”).

46. Ben Bernanke is one of the more outspoken champions of the SCAP. See Ben S. Bernanke, Chairman, Bd. of Governors of the Fed. Reserve Sys., Speech at the “Maintaining Financial Stability: Holding a Tiger by the Tail” Financial Markets Conference Sponsored by the Federal Reserve Bank of Atlanta, Stone Mountain, Georgia: Stress Testing Banks: What Have We Learned? (Apr. 8, 2013) [hereinafter Maintaining Financial Stability], available at <https://www.federalreserve.gov/newsevents/speech/bernanke20130408a.htm> [<https://perma.cc/Z46B-JGS3>] (“In retrospect, the SCAP stands out for me as one of the critical turning points in the financial crisis. It provided anxious investors with something they craved: credible information about prospective losses at banks. Supervisors’ public disclosure of the stress test results helped restore confidence in the banking system and enabled its successful recapitalization.”); Ben S. Bernanke, Chairman, Bd. of Governors of the Fed. Reserve Sys., The Supervisory Capital Assessment Program (May 11, 2009), available at <https://www.bis.org/review/r090512a.pdf> [<https://perma.cc/EUE9-JGL4>].

C. THE DODD–FRANK ACT STRESS TESTING RULES

In large part due to the SCAP experiment's perceived success, stress tests were enshrined as a central component of the Dodd–Frank Act of 2010.⁴⁷ They were also incorporated in the Basel Committee's newest generation of international banking protocols,⁴⁸ and became a mainstay of the European Central Bank's post-crisis oversight framework as well.⁴⁹

The key statutory provision in Dodd–Frank is section 165(i), which mandates that federal banking agencies conduct an annual stress test of large financial institutions.⁵⁰ A further requirement is that banks undertake pairs of “company-run” stress tests, which are to be performed in-house on an annual or semiannual basis.⁵¹ Although section 165(i) is light on specifics, the substance of what these new stress testing procedures entail was laid out in an agency rule promulgated in 2012, “Reg YY,”⁵² which outlined a program now known as the Dodd–Frank Act Stress Test (“DFAST”). As implemented under Reg YY, the DFAST works much the same as SCAP. One major twist is that it asks banks to model a third scenario, which stipulates the details of a “severely adverse” recession, in addition to the “adverse” and “baseline” hypotheticals used in the SCAP.⁵³ Another distinction is that, technically, banks do not pass or fail the DFAST. Instead, that assessment and its accompanying regulatory

47. Daniel K. Tarullo, Governor, Bd. of Governors of the Fed. Reserve Sys., Speech at the Yale University School of Management Leaders Forum, New Haven, Connecticut: Next Steps in the Evolution of Stress Testing (Sept. 26, 2016), *available at* <https://www.federalreserve.gov/newsevents/speech/tarullo20160926a.htm> [<https://perma.cc/4H3Q-8CEN>] (“During the financial crisis, the success of an ad hoc stress test in assessing the capital needs of, and restoring confidence in, the nation's largest financial institutions encouraged Congress to make stress testing a required and regular feature of large firm prudential regulation.”).

48. *See generally* BASEL COMM. ON BANKING SUPERVISION, BANK FOR INT'L SETTLEMENTS, STRESS TESTING PRINCIPLES (2018), *available at* <https://www.bis.org/bcbss/publ/d450.pdf> [<https://perma.cc/VL4S-DAZW>] (making stress testing recommendations for international banks).

49. In 2009, the European Central Bank's Committee of European Banking Supervisors administered the first stress test in the Eurozone. *See* Press Release, European Banking Auth., CEBS's Statement on Stress Testing Exercise (May 12, 2009), *available at* <https://eba.europa.eu/cebs-s-statement-on-stress-testing-exercise> [<https://perma.cc/24GF-5NQZ>]. The European Banking Authority, a new regulatory entity formed in 2011, implemented a biannual stress testing program beginning in 2014. *See generally* Alexander Abramovich, Note, *Comparative Analysis of Stress Testing in the United States and Europe*, 15 N.C. BANKING INST. 333 (2011) (analyzing stress test systems in Europe and the United States).

50. Dodd–Frank Wall Street Reform and Consumer Protection Act, Pub. L. No. 111–203, § 165(i)(1)(A), 124 Stat. 1376, 1430 (2010).

51. *Id.* § 165(i)(2)(A); 12 C.F.R. § 252.51 (2015).

52. Supervisory and Company-Run Stress Test Requirements for Covered Companies, 12 C.F.R. pt. 252.III.C. (2012).

53. For the agency supervised tests, the Federal Reserve develops the scenarios, provides them to the banks, and then asks the banks to provide the relevant balance sheet information for regulators to run the test. Policy Statement on the Scenario Design Framework for Stress Testing, 12 C.F.R. pt. 252 (2013). In the company-run stress tests, the banks themselves develop and run scenarios. *Id.*

remedies are available pursuant to a related program, the Comprehensive Capital Analysis Review (“CCAR”).⁵⁴

The CCAR requires a subset of the larger banks subject to the DFAST to submit annual “Capital Plans” to the Federal Reserve.⁵⁵ In its capital planning document, a bank must provide an overview of the shareholder distributions it anticipates to make within the next year and outline its strategy for raising further capital if economic conditions deteriorate during that time.⁵⁶ The CCAR has both a qualitative and quantitative component, and banks can fail either one or both. Similar to the SCAP, the quantitative test looks to see if a bank’s DFAST results suggest that it would maintain a capital ratio above the regulatory minimum of five percent in the event a severely adverse recession occurs.⁵⁷ With the qualitative test, the federal banking agencies review the narrative aspects of a bank’s Capital Plan and issue one of three determinations—“approve,” “reject,” or “resubmit”—based on whether those documents are found to be credible.⁵⁸ For either a quantitative or qualitative stress test failure, the primary regulatory response is for the agencies to impose limitations on the banks following four quarters of dividend distributions and stock repurchases, with both measures meant as a way to boost a financial institution’s capital levels through retained earnings.⁵⁹

A third stress test procedure which rounds out the Dodd–Frank Act’s alphabet soup is the Comprehensive Liquidity Assessment Review (“CLAR”).⁶⁰ The conceptual distinction between bank “liquidity” (as in the CLAR) and bank “capital” (as in the CCAR) is notoriously fragile, and in many cases collapses into two sides of the same coin. However, the CLAR is unique in that it does not ask how a bank’s capital buffer would fare during a generalized recession in the real economy.⁶¹ Instead, it focuses on whether a bank’s funding model is designed to withstand the kind of short-term panic in credit markets that might otherwise lead to a run on its debt.⁶² A further difference

54. BD. OF GOVERNORS OF THE FED. RESERVE SYS., COMPREHENSIVE CAPITAL ANALYSIS AND REVIEW: OBJECTIVES AND OVERVIEW 1–2 (2011), *available at* <https://www.federalreserve.gov/newsevents/pressreleases/files/bcreg20110318a1.pdf> [<https://perma.cc/UJ7S-V3FY>].

55. Capital Plans, 12 C.F.R. § 225.8 (2012).

56. *Id.*; *see also* BD. OF GOVERNORS OF THE FED. RESERVE SYS., *supra* note 54, at 9–10.

57. BD. OF GOVERNORS OF THE FED. RESERVE SYS., *supra* note 54, at 12–15.

58. 12 C.F.R. § 225.8; BD. OF GOVERNORS OF THE FED. RESERVE SYS., *supra* note 54, at 1–9.

59. 12 C.F.R. § 225.8.

60. *See* TREAS. BANK’G REP., *supra* note 5, at 147.

61. *See id.*; *see also* OFFICE OF FIN. RESEARCH, FINANCIAL STABILITY REPORT 98 (2015), *available at* https://www.financialresearch.gov/financial-stability-reports/files/OFR_2015-Financial-Stability-Report_12-15-2015.pdf [<https://perma.cc/E7RY-EF3V>] (“The CCAR/DFAST process has not explicitly incorporated the risk of a funding run when a BHC is unable to roll over its short-term borrowing. This type of liquidity stress helped bring down Bear Stearns and Lehman Brothers in 2008 just before the severe economic downturn that followed, which raises the question of how liquidity shocks should be incorporated into stress testing.”).

62. *See* OFFICE OF FIN. RESEARCH, *supra* note 61, at 98.

is that, unlike the procedures which focus on bank capital, financial institutions do not pass or fail the liquidity stress tests. The CLAR is less prominent than the banking agencies' CCAR program, and often overlooked. Indeed, as with the IMF's Financial Sector Assessment Program, a review of the post-crisis legal literature provides little indication that it exists, but it has important implications that inform this Article's proposal on how to redesign stress testing going forward.

D. STRESS TESTS RESULTS FROM 2012–2018

Around the time of Dodd–Frank's passage, the merits of its stress testing rules could only be debated in the abstract. By now, there is nearly a decade of experience with how those procedures perform in practice. The following chart summarizes the outcomes of the Federal Reserve's CCAR program to date:⁶³

<u>Year</u>	<u>Institutions Tested</u>	<u>Quantitative Fails</u>	<u>Qualitative Fails</u>	<u>Qualitative Pass*</u> (conditional on resubmission)
2012	18	Ally Financial, Citigroup, SunTrust Banks, MetLife	None	None
2013	18	Ally Financial	Ally Financial BB&T Corporation	JPMorgan Chase Goldman Sachs
2014	30	Zions	Citigroup, RBS, HSBC, Santander	None
2015	31	None	Deutsche Bank, Santander	None
2016	33	None	Deutsche Bank, Santander	None
2017	34	None	None	Morgan Stanley
2018	35	None	Deutsche Bank	State Street, Morgan Stanley, Goldman Sachs

A few aspects of this track record are worth highlighting. First is the relatively low failure rate of banks based on the banking agencies' quantitative

63. Each year, the Federal Reserve publishes the results of the CCAR test along with guidance for next year. See, e.g., BD. OF GOVERNORS OF THE FED. RESERVE SYS., DODD–FRANK ACT STRESSTEST 2013: SUPERVISORY STRESS TEST METHODOLOGY AND RESULTS 1–4 (2013), available at https://www.federalreserve.gov/newsevents/press/bcreg/dfast_2013_results_20130314.pdf [<https://perma.cc/XGZ9-TSQU>].

criterion. Since 2015, stress tests have been passed based on banks' DFAST results. One obvious explanation is that the economic recovery has shored up bank balance sheets. But that stands in some tension with the fact that a few banks, namely Santander and Deutsche Bank, have suffered a number of financial setbacks in recent years and have also been cited as serial offenders on the qualitative tests during that same period.⁶⁴

Second, the basis for the qualitative failures have been wide-ranging and sometimes unexpected. The best example is Citigroup in 2014.⁶⁵ In explaining its negative determination, the Federal Reserve conceded that it had not identified any financial vulnerability per se in the bank's business model.⁶⁶ Rather, Citigroup's failure was justified on procedural grounds, and stemmed from a finding that its internal protocols for modeling the Fed's scenarios were not sufficiently rigorous to make the bank's quantitative results reliable.⁶⁷

A third significant feature of the Dodd-Frank stress tests is the highly informal manner with which they have sometimes been carried out. For example, although Bank of America and Goldman Sachs both passed in 2014, they only did so after modifying their Capital Plans in response to a preemptive notice from the Federal Reserve, which hinted that a negative assessment on the qualitative test might be forthcoming.⁶⁸ And in 2017, the Federal Reserve granted both Goldman Sachs and Morgan Stanley a conditional pass based on a determination that—even though their capital projections fell short of applicable requirements—the banks had nonetheless satisfied an otherwise unwritten safe harbor exception, which the Federal Reserve appeared to have crafted during the course of its review.⁶⁹

64. See David Henry, *Santander, Deutsche Bank: U.S. Stress Test Repeat Offenders*, REUTERS (June 30, 2016, 12:16 AM), <https://www.reuters.com/article/us-usa-banks-stress-failure-idUSKCN0ZGoDM> [<https://perma.cc/2E2A-D7GK>].

65. See Stephanie Armour et al., *Fed Kills Citi Plan to Pay Investors*, WALL ST. J., <https://www.wsj.com/articles/fed-rejects-citigroups-capital-plan-in-stress-test-approves-25-other-banks-1395864005> [<https://perma.cc/AJV6-K45D>] (last updated Mar. 26, 2014, 4:01 PM) (reporting on Citi executives' surprise at the negative result).

66. See BD. OF GOVERNORS OF THE FED. RESERVE SYS., COMPREHENSIVE CAPITAL ANALYSIS AND REVIEW 2014: ASSESSMENT FRAMEWORK AND RESULTS 6, 7 (2014), available at https://www.federalreserve.gov/newsevents/press/bcreg/ccar_20140326.pdf [<https://perma.cc/2UZM-qZ4V>] (noting "deficiencies" in the bank's "abilit[ies] to project revenue and losses under a stressful scenario for material parts of the firm's global operations" and "to develop scenarios for its internal stress testing that adequately reflect and stress its full range of business activities and exposures").

67. See *id.*

68. See Michael J. Moore, *Goldman Sachs's Stress-Test Stumble Leads to Cut in Repo Lending*, BLOOMBERG (July 15, 2014, 11:00 PM), <https://www.bloomberg.com/news/articles/2014-07-16/goldman-sachs-s-stress-test-stumble-leads-to-cut-in-repo-lending> [<https://perma.cc/A37X-PFFG>]; *Comprehensive Capital Analysis*, *supra* note 8, at 105–08.

69. The implicit rule appears to be that dividend plans are authorized so long as the anticipated distributions do not exceed the levels from the prior year. See Liz Hoffman & Lalita Clozel, *Morgan Stanley, Goldman Got Help From Fed on Stress Tests*, WALL ST. J., <https://www.wsj.com/articles/wall-street>

E. RECENT PROPOSALS TO ROLL BACK THE DODD–FRANK RULES

Although some of the anecdotes described above inspired criticism in policy circles and the business press, regulatory stress testing programs experienced a continuous expansion since the SCAP was unveiled in 2009. That trend, however, was dramatically reversed starting in 2017, due to a wave of reforms proposals that have come from both Congress and the Trump Administration. Many of those proposals represent a wholesale critique of the post-crisis policy framework installed pursuant to Dodd–Frank, and a major theme is to cut back reliance on stress tests as a regulatory tool.

1. Executive Branch Reforms

An initial impetus for reform came from the executive branch with a Trump Administration Executive Order, issued on February 8, 2017, which called for a comprehensive review of existing financial regulations.⁷⁰ To that end, the Treasury Department has begun to release a series of policy memoranda which outline the federal banking agencies' reform agenda for each sector of the financial system. Most relevant here is a report released by the Treasury in June of 2017 ("the Treasury Banking Report").⁷¹

The Treasury Banking Report contains a number of recommendations, all of which would result in the stress testing process becoming either more transparent or less onerous for banks. For example, the Treasury proposed that the Federal Reserve subject its stress testing and capital planning review frameworks to enhanced public scrutiny.⁷² The Treasury Banking Report also expresses support for giving banking regulators greater (downward) flexibility when implementing the stress tests, by tailoring those requirements based on the business model and complexity of a particular bank.⁷³ The same document also raises the possibility that mid-sized institutions, with assets between \$50 and \$250 billion, could be exempted from many of Dodd–Frank's current stress testing mandates.⁷⁴

The federal banking agencies have already started to take action on some of the Treasury Banking Report's proposals. Last December, the Federal

gets-the-friendlier-fed-its-been-waiting-for-1530558419 [https://perma.cc/876Z-PBM3] (last updated July 2, 2018, 9:29 PM); Matt Levine, *Not Stressing the Stress Tests*, BLOOMBERG: OPINION (July 3, 2018, 10:01 AM), <https://www.bloomberg.com/opinion/articles/2018-07-03/not-stressing-the-stress-tests> [https://perma.cc/S7EB-F8WN].

70. See generally Exec. Order No. 13,772, 82 Fed. Reg. 9965 (Feb. 8, 2017) (establishing Core Principles for Regulating the United States Financial System).

71. See generally TREAS. BANK'G REP., *supra* note 5 (reporting on the depository system, covering banks, savings associations, and credit unions of all sizes, types, and regulatory charters).

72. *Id.* at 12.

73. *Id.*

74. *Id.* at 12, 48–49 (recommending that, "[f]or the company-run stress tests, banks should be permitted to determine the appropriate number of models that are required to develop sufficient output results, based on the complexity of the banking organization and the nature of its assets").

Reserve Board put forward a package of three proposals in the Federal Register, which internalized the recommendations of the Treasury Department and aimed to improve transparency of CCAR.⁷⁵ As detailed in an accompanying press release, the general thrust is to provide more transparency about agencies' modeling methodologies, including the way their stress scenarios are developed and how their results are interpreted.⁷⁶

2. Statutory Reforms

Congress has been equally or more aggressive. The most ambitious proposal presented is the Financial CHOICE Act, which passed the House in 2017.⁷⁷ The Financial CHOICE Act envisions a complete remake of the Dodd–Frank Act structure, including its stress tests. The mechanism it uses to do so is included in a set of so-called Off-Ramp provisions, which apply to large financial institutions that meet the requirements for being considered a “Qualifying Banking Organization.”⁷⁸ To be a “Qualifying Banking Organization,” a bank must comply with a ten-percent leverage ratio rule that is introduced in other parts of the bill. Banks that maintain capital above that threshold are exempted from all stress testing imposed under Dodd–Frank.⁷⁹

While momentum behind the Financial CHOICE Act has petered out, the thinking it reflects inspired a piece of legislation that did get passed by Congress and was enacted on May 24, 2018, the Economic Growth, Regulatory Relief, and Consumer Protection Act (“EGRRCPA”).⁸⁰ One of the main thrusts of its reforms is to reduce the regulatory burden for mid-sized banks, defined as firms holding less than \$50 billion in consolidated assets.

75. See generally Enhanced Disclosure of the Models Used in the Federal Reserve's Supervisory Stress Test, 82 Fed. Reg. 59,547 (proposed Dec. 15, 2017) (proposing enhanced disclosure requirements); Policy Statement on the Scenario Design Framework for Stress Testing, 82 Fed. Reg. 59,533 (Dec. 15, 2017) (proposing modifications to the Federal Reserve Board's scenario design framework); Stress Testing Policy Statement, 82 Fed. Reg. 59,528 (proposed Dec. 15, 2017) (outlining the characteristics of the proposed supervisory stress test scenarios).

76. Press Release, Bd. of Governors of the Fed. Reserve Sys., Federal Reserve Board Requests Comment on Package of Proposals That Would Increase the Transparency of Its Stress Testing Program (Dec. 7, 2017), available at <https://www.federalreserve.gov/newsevents/pressreleases/bcreg20171207a.htm> [<https://perma.cc/5N86-A3MC>]; cf. Randal K. Quarles, Vice Chairman for Supervision, Bd. of Governors of the Fed. Reserve Sys., Speech at the American Bar Association Banking Law Committee Annual Meeting, Washington, D.C.: Early Observations on Improving the Effectiveness of Post-Crisis Regulation (Jan. 19, 2018), available at <https://www.federalreserve.gov/newsevents/speech/quarles20180119a.htm> [<https://perma.cc/7ARK-447U>] (emphasizing priority on transparency for stress tests and tailoring to institutional size).

77. See Financial CHOICE Act of 2017, H.R. 10, 115th Cong. (2017).

78. *Id.* § 602(a)(2).

79. See *id.* § 601(b)(1).

80. See Economic Growth, Regulatory Relief, and Consumer Protection Act, Pub. L. No. 115-74, 132 Stat. 1296 (2018).

Banks under that threshold have now been relieved of all their former stress testing requirements.⁸¹

The role of financial stress testing has followed a dramatic arc in the decade since the financial crisis. Starting in 2009, those requirements were transformed from a kind of legal window-dressing to a lynchpin of the policy regime. And within the past few years, the same procedures have been subject to a burst of proposed or enacted reforms which now put their continued viability in doubt. Financial regulation has therefore reached a critical juncture where the merits of stress testing must be clarified in order to chart an informed path forward. That task is undertaken in the following Part.

III. EVALUATING THE PROMISE AND LIMITS OF STRESS TESTS: AN UNFINISHED TASK

The question of whether the post-crisis stress tests should be scrapped in full or part cannot be answered without a clear understanding of what precisely it is they aim to do, as well as the practical barriers to achieving those goals. This Part walks through the principal arguments for and against the use of stress testing in its current form. Working through these claims at length is necessary not only as a matter of thorough analysis and critique, but also supplies much of the groundwork for understanding the logic behind this Article's reform proposal, which is laid out in the following Part.

A. *THE INCOMPLETE CASE AGAINST*

The discussion below reviews the three main objections against regulatory stress testing. They range from arguments that the basic concept is flawed, to more pragmatic concerns which suggest that the administration of the Dodd–Frank stress tests should be substantially modified or scaled back. Each of the three objections can be seen as animating proposals in the reform agenda surveyed above. As will be argued, however, those critiques are either misplaced or, at best, are of more limited significance than is assumed and do not support the kinds of reforms which are currently being pursued.

1. Quantitative Skepticism

A common threshold complaint against regulatory stress tests is that the premise of relying on complex quantitative models as a risk management device is inherently suspect. A major source of criticism along those lines comes from the failure of existing models to anticipate the financial crisis in the first place. As central banker Andrew Haldane put it, in a prominent policy speech delivered in 2009, “[R]isk management models have during this crisis proved themselves wrong in a . . . fundamental sense. . . . With

81. See *id.* § 401(c)(3), 401(e).

hindsight, these models were both very precise and very wrong.”⁸² In addition to that poor pre-crisis track record, the argument goes, the overwhelming complexity and deep uncertainties that characterize modern financial markets mean that stress testing is unlikely to perform any better going forward.

In her article, “Regulation by Hypothetical,” Professor Mehrsa Baradaran suggests that the use of scenario-testing in particular is questionable because, “hypothetical models of any sort have serious and irremediable structural flaws.”⁸³ More specifically, she concludes that “the significant problem of the hypothetical regime [is] it can prepare firms for cyclical market problems, but it cannot prepare them for unprecedented market occurrences.”⁸⁴ According to Professor Robert Weber, the solution is for regulators to adopt an attitude of “quantitative skepticism,” which prioritizes qualitative considerations as part of a more “deliberation-oriented” approach to evaluating the safety and soundness of financial institutions.⁸⁵ Such a response can be seen in recent reform proposals. One example is the Financial CHOICE Act and its Off-Ramp provisions, which seek to eliminate stress testing requirements for Too Big to Fail banks that satisfy more traditional regulatory targets.⁸⁶

Although many of the shortcomings of stress testing which these commentators identify are well-taken, a problem with this line of critique is that it overlooks the same weaknesses in available alternatives. At a basic conceptual level, all regulation is “regulation by hypothetical”: Every legal rule functions as a forward-looking intervention in an uncertain world and is necessarily premised on some set of assumptions about what the future holds. The challenge of anticipating how banks might perform when faced with unprecedented economic disruptions is endemic to financial regulation, whatever form it takes. It is not an unrealistic ambition peculiar to stress testing.

Likewise, the presence of model-risk—the possibility that a model may provide a biased forecast of future events because it fails to capture relevant aspects of reality—is not unique to the use of complicated quantitative techniques. “Models,” properly understood, are simply assertions about a set

82. Haldane, *supra* note 9, at 1; *see also* Baradaran, *supra* note 8, at 1277, 1281, 1317 (echoing this point).

83. Baradaran, *supra* note 8, at 1324; *see also id.* at 1294 (“[T]here is no model that can accurately predict the broad market effects of thirty percent unemployment because it is historically unprecedented—there is no data for such an event.”).

84. *Id.* at 1295.

85. Weber, *A Theory for Deliberation-Oriented Stress Testing Regulation*, *supra* note 8, at 2273 (defining quantitative skepticism as the recognition that “risk [modeling] reflects [an] aspiration[] to control increasingly uncontrollable phenomena”); *cf.* Anette Mikes, *Risk Management and Calculative Cultures*, 20 MGMT. ACCT. RES. 18, 22 (2009) (coining the term “quantitative skeptics” as part of a similar critique).

86. *See* Financial CHOICE Act of 2017, H.R. 10, 115th Cong.; *see supra* Section II.E.2.

of logical relationships.⁸⁷ Those assertions can be expressed informally through verbal statements or in formal mathematical terms.⁸⁸ Over-confident models, by definition, warrant skepticism. But that is a truism which does not itself have any special implication for *quantitative* risk modeling. Leading up to the crisis, most qualitative, informal models of financial markets were just as optimistic as their mathematical counterparts, and proved equally over-confident.⁸⁹ Thus, the common post-crisis motif that “the models failed” carries less policy significance than is often thought. By selectively attributing the problem of model-risk to quantitative financial analysis, it fails to explain how marginalizing the role of stress testing will help prevent the regulatory framework from embodying overly optimistic predictions about the future.

Conceptual issues aside, the main reasons why a shift toward greater quantitative skepticism is unwarranted are historical and pragmatic. From a historical perspective, the lack of stress testing prior to the 1980s was no coincidence—that was when the technology it requires became possible for the first time, as a result of the revolutionary decades in mathematical finance and computing power that took place in the 1950s, 60s, and 70s.⁹⁰ Ever since,

87. See DANIEL M. HAUSMAN, *THE INEXACT AND SEPARATE SCIENCE OF ECONOMICS* 70–82 (1992) (on the nature of economic models). In one sense, the failure of any quantitative model is always the product of a prior failure in qualitative reasoning. At some point, there must have been a (mistaken) qualitative determination that the quantitative model of choice would be reliable.

88. Both informal verbal models and formal mathematical models can be simple or complex. If anything, verbal models tend toward greater complexity, because language is freighted with much more symbolic meaning and ambiguity than is mathematical notation. See generally Tyler Cowen, *Is a Novel a Model?* (Jan. 29, 2005) (unpublished manuscript), available at <http://www.sscnet.ucla.edu/polisci/faculty/chwe/austen/cowen2005.pdf> [<https://perma.cc/J9RL-LHAC>] (arguing that, in many ways, novels can be seen as very complicated verbal models).

89. An influential theme from the 1990s and early 2000s, which held that the economy had entered an era of “Great Moderation,” is an example of a simple verbal model of financial markets which did not stand the test of time. See Ben S. Bernanke, Governor, Remarks at the Meetings of the Eastern Economic Association, Washington, D.C.: The Great Moderation (Feb. 20, 2004), available at <https://www.federalreserve.gov/boarddocs/speeches/2004/20040220> [<https://perma.cc/KPC4-T9S2>]; see also Haldane, *supra* note 9, at 1 (discussing how the “over-reli[ance]” of the financial model partly contributed to its failure). Anecdotes about commentators who “called” the financial crisis do not say much about the accuracy of verbal models, as it is always possible to find vague prognostications that look prophetic in retrospect. See John Tamny, *If They Tell You They Predicted the Financial Crisis, They’re Lying*, FORBES (Oct. 8, 2013, 2:00 PM), <https://www.forbes.com/sites/johntamny/2013/10/08/if-they-tell-you-they-predicted-the-financial-crisis-theyre-lying> [<https://perma.cc/YE3V-7F69>].

90. Relevant landmarks in mathematical finance include the Markowitz portfolio selection model, published in 1952, and the Black-Scholes options-pricing formula from 1973. See Mark Rubinstein, *Guiding Force*, in *FROM BLACK-SCHOLES TO BLACK HOLES: NEW FRONTIERS IN OPTIONS* 39–48 (Risk Magazine 1992) (“[T]he Black & Scholes model is now widely viewed as one of the most successful in the social sciences and perhaps, including its binomial extension, the most widely used formula, with embedded probabilities, in human history.”). Those insights from theoretical finance would have been impossible to apply in practice if not for the fact that computing power improved by a factor of one million from 1955 to 1990. See generally Stavros A. Zenios, *High-*

banks and other sophisticated financial market participants have leveraged those innovations to guide every investment decision they make. Those practices cannot be interpreted as an idiosyncratic error in judgment on the part of the financial community. Instead, they are part of the larger trend toward the quantification of all human endeavor that has come to define the modern world.⁹¹ Within that broader context, it is hard to see how turning back to a policy regime based on banking regulators' instinct and intuition could be effective. As a practical matter, there is a certain inevitability to a heavy reliance on regulatory tools which seek to measure systemic risks based on complicated quantitative models that make forward-looking predictions.

To summarize, there is no fatal flaw in the idea behind scenario-based stress testing per se (or quantitative risk management in general) that can be used to justify the reformist push to strip those procedures from the post-crisis legal framework on vague epistemological grounds. The family of arguments surveyed above, which at times may suggest otherwise, are best interpreted as reminders of more fundamental dilemmas that apply to financial regulation across the board. The main policy takeaway from those arguments is therefore relatively mundane. The Dodd–Frank rules should be carefully scrutinized to ensure their modeling techniques have not failed to incorporate lessons from the financial crisis experience or otherwise defy common sense. To the extent they do so, the financial system can be made more secure by finding ways to improve those methodologies going forward.⁹²

2. Regulatory Discretion, Opacity, & Uncertainty

The Dodd–Frank stress tests have also been subject to a series of critiques on procedural grounds, which circle around three related objections about how they are administered. First, the agencies' testing methodologies lack transparency, and give banks little notice of the criteria they are expected to meet.⁹³ Second, the banking agencies exercise broad discretion in interpreting the results of a bank's test and selecting the remedy which accompanies a determination that it has failed.⁹⁴ Third, the penalties which regulators are able to impose in connection with a negative determination

Performance Computing in Finance: The Last 10 Years and the Next, 25 *PARALLEL COMPUTING* 2149 (1999) (documenting the parallel rise of mathematical finance and computing power).

91. The rise of advanced statistical analyses in professional sports is one of the best illustrations of the gains which have been achieved across nearly every domain that has embraced an aggressive form of "qualitative skepticism" over the past few decades. *See generally* MICHAEL LEWIS, *MONEYBALL: THE ART OF WINNING AN UNFAIR GAME* (2003) (demonstrating how an Oakland athletics manager used advanced statistical analyses to successfully compile a Major League Baseball team).

92. *See infra* Section IV.A.1 (proposing several reforms to that end).

93. *See, e.g.*, U.S. HOUSE COMM. ON FIN. SERVS., *THE FINANCIAL CHOICE ACT: CREATING HOPE AND OPPORTUNITY FOR INVESTORS, CONSUMERS, AND ENTREPRENEURS* 9–10 (2017) [hereinafter *HOUSE CHOICE ACT REP.*].

94. *Id.*

often take the form of heavy-handed intrusions on the bank's business model or the decision-making authority of its managers.⁹⁵

These concerns were summed up in recent congressional testimony by financial economist Charles Calomiris, who called the Dodd–Frank “stress tests . . . a Kafkaesque Kabuki drama in which regulators punish banks for failing to meet standards that are never stated (either in advance or after the fact).”⁹⁶ The result, he concludes, is that stress tests have become “a source of uncertainty rather than a helpful guide against unanticipated risks.”⁹⁷ The same view has motivated many of the reform proposals. Professor Calomiris' testimony was referenced in the House Report issued in support of the Financial CHOICE Act.⁹⁸ And a primary focus of the Federal Reserve's recent policy guidance involves efforts to add greater transparency to the stress testing process.⁹⁹

But, the complaints about regulatory discretion, opacity, and uncertainty have all been led astray by the surface-level exoticism of the post-crisis stress tests as a regulatory tool. When placed in broader legal context, it becomes clear that the defects they purport to identify in the Dodd–Frank stress testing rules merely replicate hallmark features of banking regulation. If credited, the same line of critique would also condemn much of the legal structure dating back decades, or even centuries.

For one, the Dodd–Frank stress tests are in fact relatively transparent compared to traditional forms of bank examination, the results of which are kept completely confidential.¹⁰⁰ Moreover, pursuant to those reviews, bank supervisors have nearly unfettered discretion to make a holistic assessment of a bank's overall “safety-and-soundness,” and apply any of a vast suite of penalties when they deem a bank has fallen short.¹⁰¹ Since the Civil War, the standard penalty that has been imposed under that exercise is a restriction on the ability of a bank to issue dividends—the exact same measure as the default

95. *Id.* at 11–13.

96. *What's Wrong with Prudential Bank Regulation and How to Fix It: Hearing Before the H. Comm. on Fin. Servs.*, 114th Cong. 6 (2015) (statement of Charles W. Calomiris).

97. *Id.*

98. HOUSE CHOICE ACT REP., *supra* note 93, at 72 (“The stress tests have become a kind of ‘cat-and-mouse’ exercise in which Fed staff and bank compliance officers attempt to outwit one another in a game without rules or transparency. The secrecy surrounding the stress tests makes it difficult for Congress and the public to assess either the effectiveness of the Fed's regulatory oversight or the integrity of the findings yielded by the tests.”).

99. *See supra* notes 75–76 and accompanying text (citing the relevant regulatory documents and providing a summary of their contents).

100. The results of bank examinations are not disclosed and enjoy special immunity from FOIA Requests under the Freedom of Information Act. *See* John Crawford, *Predicting Failure*, 7 VA. L. & BUS. REV. 171, 197–98 (2012) (explaining the traditional “CAMELS” (Capital, Assets, Management, Earnings, Liquidity) approach to bank examinations and noting that it “is highly confidential and is known only to the relevant examiners and bank managers”).

101. Federal Deposit Insurance Act § 8(b), 12 U.S.C. § 1818(b) (2012).

remedy under Dodd–Frank’s CCAR protocols.¹⁰² Thus, the regulatory uncertainty critique is really an indictment of banking regulation in general, not the post-crisis stress tests.

The transparency issue in particular is a source of confusion. The justification for complete confidentiality in traditional bank examinations is straightforward. Releasing the results can trigger a run on banks that are found to be weak.¹⁰³ An identical logic underlies the classic problem of “stigma” that accompanies central bank lending during a financial crisis, and explains why banks are allowed to remain anonymous when requesting emergency funding from the Federal Reserve’s discount window.¹⁰⁴ The common assumption that transparency is an unalloyed regulatory good—captured in Justice Brandeis’ famous dictum that “[s]unlight is said to be the best of disinfectants; electric light the most efficient policeman”—simply does not hold when it comes to banking regulation.¹⁰⁵

More fundamentally, the allegedly dysfunctional “cat-and-mouse” dynamic between banks and regulators that opaque stress testing procedures create is in one sense the entire point.¹⁰⁶ The need for that dynamic stems from a foundational principle of financial regulation, known as Goodhart’s Law: When a measure of risk becomes the regulatory target for banks, it ceases to be a good measure for regulators.¹⁰⁷ Disclosing the model so that it is predictable allows banks to game the model, by attempting to mimic the Federal Reserve’s results rather than focus on their own risk.¹⁰⁸ As Federal

102. *Comprehensive Capital Analysis*, *supra* note 8, at 44 (“Since the Civil War era, banking law in the United States has restricted the ability of banks to pay dividends and make other distributions of corporate property.”).

103. See Gary Gorton, *The Development of Opacity in U.S. Banking*, 31 YALE J. REG. 825, 836–39 (2014).

104. Carmen M. Reinhart & Christoph Trebesch, *The International Monetary Fund: 70 Years of Reinvention*, 30 J. ECON. PERSP. 3, 23 (2016) (“In a domestic setting, banks often shy away from approaching a central bank’s discount window for fear that temporary illiquidity will be mistaken for insolvency by fellow market participants . . .”).

105. LOUIS D. BRANDEIS, *OTHER PEOPLE’S MONEY AND HOW THE BANKERS USE IT* 92 (1914).

106. See HOUSE CHOICE ACT REP., *supra* note 93, at 72.

107. See Charles Goodhart, *Problems of Monetary Management: The U.K. Experience*, in INFLATION, DEPRESSION AND ECONOMIC POLICY IN THE WEST 111, 116 (Anthony S. Courakis ed., 1981). Goodhart’s Law is an application of a more fundamental insight from macroeconomics known as the Lucas Critique, which is an argument that the consequences of any policy change cannot be projected from prior trends without taking account the forward-looking response of individuals, firms, and other economic variables to the new policy rule. See generally Robert E. Lucas, Jr., *Econometric Policy Evaluation: A Critique*, 1 CARNEGIE-ROCHESTER CONF. SERIES ON PUB. POL’Y 19 (1976) (presenting the theoretical framework for the Lucas Critique).

108. See Paul Glasserman & Gowtham Tangirala, *Are the Federal Reserve’s Stress Test Results Predictable?*, 18 J. ALTERNATIVE INV. 82, 82 (2016) (“[U]sing results made public thus far across various stress tests, we find that projected losses by bank and loan category are fairly (and increasingly) predictable.”); Til Schuermann, *The Fed’s Stress Tests Add Risk to the Financial System*, WALL ST. J.: OPINION (Mar. 19, 2013, 7:08 PM), <https://www.wsj.com/articles/SB10001424127887324532004578362543899602754> [<https://perma.cc/638V-2RHN>] (“As the

Reserve Board Vice-Chairman Randy Quarles has acknowledged in a recent policy speech, with a fully transparent stress test, nothing really gets tested.¹⁰⁹ In short, the optimal level of opacity in regulatory stress testing is arguably quite high, and it is far from clear that the current rules should be substantially more transparent.¹¹⁰

The one area where the regulatory uncertainty critique does have force is with respect to the qualitative review which regulators perform as part of the CCAR stress test program. For whatever reason, the federal banking agencies have treated that procedure as an invitation to engage in what can fairly be described as unstructured regulatory meddling.¹¹¹ But the current reforms go far beyond the qualitative CCAR test, which has no statutory basis under Dodd–Frank and could easily be excised from the current rules.¹¹² At its core, stress testing is a quantitative exercise. There is nothing about the quantitative portions of the post-crisis stress tests that entails unbounded regulatory discretion, and it is possible to administer those procedures in a rule-bound, mechanical manner.¹¹³

3. Compliance Costs

A final common critique of the stress tests turns on compliance costs. The Dodd–Frank procedures are so elaborate, it is argued, that the financial burden banks incur as part of that process washes out any potential benefits they may have.¹¹⁴ The regulatory burden issue has been emphasized for mid-sized banks in particular, based on the perception that the high fixed-cost of running the stress tests disproportionately burdens those institutions.¹¹⁵ These concerns animate the Treasury Department’s recommendation that

Fed’s models have become more and more important in deciding the fate of the biggest banks, those banks have focused more and more on trying to mimic the Fed’s results rather than tracing out their own risk profiles.”).

109. See Quarles, *supra* note 76, at 8–9 (“I appreciate the risks to the financial system of the industry converging on the Federal Reserve’s stress testing model too completely, so I am hesitant to support complete disclosure of our models for that reason.”).

110. See Itay Goldstein & Yaron Leitner, *Stress Tests and Information Disclosure*, 177 J. ECON. THEORY 34, 35 (2018) (“A key question that occupies policymakers and bankers is whether such disclosure is indeed optimal and, if so, at what level of detail.”); Itay Goldstein & Haresh Sapra, *Should Banks’ Stress Test Results be Disclosed? An Analysis of the Costs and Benefits*, 8 FOUND. & TRENDS IN FIN., Dec. 2013, at 1, 3 (“Many proponents of disclosure of stress-test results have linked the severity of the recent financial crisis to bank opacity” but “by the time regulators intervened, it was too late as there was a widespread panic because the market could not distinguish a solvent bank from an insolvent bank and such panic brought the whole financial system to its knees.”).

111. See *supra* Section II.D (describing financial stress test results from 2012 to 2018).

112. See generally BD. OF GOVERNORS OF THE FED. RESERVE SYS., *supra* note 54 (explaining that the CCAR test has no statutory basis in Dodd–Frank, unlike the DFAST).

113. See *infra* Section IV.A.

114. See generally HOUSE CHOICE ACT REP., *supra* note 93 (discussing the burdens that banks may experience).

115. James F. Powers III, *Stress Testing Under Dodd–Frank: Easing the Regulatory Burden for Mid-sized Financial Companies*, 20 N.C. BANKING INST. 361, 367–71 (2016).

the banking agencies take a more “tailored” approach to the stress testing process for banks which have less than \$250 billion dollars in assets.¹¹⁶ The same logic also informed the 2018 Economic Growth, Regulatory Relief, and Consumer Protection Act, which has effected one of the most impactful reforms in the Dodd–Frank rules to-date by eliminating all stress test requirements for banks with under \$50 billion dollars in assets.¹¹⁷

The legal, accounting, and other fees banks must pay compliance professionals in order to navigate the post-crisis stress testing rules are certainly not trivial. But some context is helpful. According to a study by the Government Accountability Office, average annual compliance costs for the very largest Too Big to Fail (“TBTF”) banks range from \$15 to \$30 million; for mid-sized banks, the range is \$250,000 to \$2 million.¹¹⁸ This means that, for institutions approaching \$50 billion dollars in size, the current stress test procedures are far from a crippling expense.¹¹⁹ The fiscal cost from public expenditures associated with federal banking agencies’ design and implementation of the Dodd–Frank stress testing programs are not overwhelming, either.¹²⁰

In addition to direct paper compliance costs, of course, are the indirect costs from reduced financial intermediation when banks pull back their lending activity in order to pass the stress tests (or comply with the remedy imposed after a negative result).¹²¹ But a unique aspect of stress testing from a policy perspective is that any question of compliance costs is essentially moot, given the scale of the problem they address. Although the social cost of a financial crisis is notoriously hard to measure, some estimates reach into the trillions of dollars.¹²² That cost is imposed on society when banks take risks

116. See TREAS. BANK’G REP., *supra* note 5, at 12; *supra* Section II.E.2.

117. See Economic Growth, Regulatory Relief, and Consumer Protection Act, Pub. L. No. 115–174, § 401(e), 132 Stat. 1296, 1359 (2018); see also Financial CHOICE Act of 2017, H.R. 10, 115th Cong. (introducing a bill to “repeal[] the provisions of the Dodd–Frank Act”); cf. Cox, *supra* note 6 and accompanying text.

118. See U.S. GOV’T ACCOUNTABILITY OFFICE, GAO-17-48, FEDERAL RESERVE: ADDITIONAL ACTIONS COULD HELP ENSURE THE ACHIEVEMENT OF STRESS TEST GOALS 30 (2016) [hereinafter GAO STRESS TEST REP.] (estimating compliance costs for the stress tests).

119. See generally Glasserman & Tangirala, *supra* note 108 (discussing banks’ compliance costs).

120. See generally GAO STRESS TEST REP., *supra* note 118 (discussing the fiscal cost of administering the stress tests).

121. See Kristle Cortés et al., *Stress Tests and Small Business Lending* 33 (Nat’l Bureau of Econ. Research, Working Paper No. 24365, 2018), available at <https://www.nber.org/papers/w24365.pdf> [<https://perma.cc/2FUD-NEZA>].

122. Eric Posner & E. Glen Weyl, *Benefit-Cost Analysis for Financial Regulation*, 103 AM. ECON. REV.: PAPERS & PROC. 393, 393–94 (2014) (citing a range of estimates from \$150 billion to \$3 trillion); see also John C. Coates IV, *Cost-Benefit Analysis of Financial Regulation: Case Studies and Implications*, 124 YALE L.J. 882, 966–67 (2015) (discussing the difficulty of precisely accounting for the total social cost of a financial crisis); cf. Carmen M. Reinhart & Kenneth S. Rogoff, *Banking Crises: An Equal Opportunity Menace*, 37 J. BANKING & FIN. 4557, 4557–58 (2013) (detailing how recessions that involve banking crises tend to be especially severe).

which are externalized to the broader financial system and real economy as a whole. No matter how steep a burden banks bear in order to comply with stress testing procedures or other requirements which seek to internalize those risks, it is potentially well worth the cost. All that matters is whether or not those regulations work.

Because the preceding point is widely recognized on some level, the regulatory burden argument has gained the most traction when applied to mid-sized banks. In addition to the observation regarding fixed compliance costs, a further premise of the argument is usually that, although the social costs of a financial crisis are often astronomical, stress testing can safely be limited to the main source of those costs: the TBTF megabanks at epicenter of the 2008 crisis. That premise fails for a few reasons.

First, the prevailing assumption that mid-sized banks survived the financial crisis unscathed is more myth than fact. According to a recent study by the FDIC, 489 banks failed from 2007 to 2013 and, as of 2010, the FDIC considered 884 banks to be in financial distress.¹²³ The financial crisis provides no basis for excluding mid-sized banks from procedures that estimate how they would perform in an adverse recessionary scenario, because those banks were unprepared for the events of 2008 and subsequently wiped out en masse as a result.¹²⁴

Second, even taking the mythology of resilient mid-sized banks as true, the same conclusion still follows. Phasing out mid-sized banks runs contrary to the wisdom that any policy response to a high-profile episode of market failure which fixates on “fighting the last war” is misguided.¹²⁵ And it is at least plausible that the next financial sector collapse will originate outside of the very largest institutions. In general, financial theory suggests that bigger banks are more stable due to their greater diversification.¹²⁶ A narrow focus on the

123. FED. DEPOSIT INS. CORP., CRISIS AND RESPONSE: AN FDIC HISTORY, 2008–2013, at 119, 181–83 (2017).

124. See generally Jeremy C. Kress & Matthew C. Turk, *Too Many to Fail: Against Community Bank Deregulation*, 115 NW. U. L. REV. (forthcoming 2020) (arguing that post-crisis financial regulation has neglected the risks posed by smaller financial institutions).

125. In other words, policymaking guided by counterfactual history that imagines a set of rules which would have been effective at preventing the last crisis is error-prone, because future crises will inevitably be different in important ways. If the paradigm market failure associated with financial bubbles is a hope that “this time is different,” the analogous regulatory failure in the aftermath of a bubble’s collapse is an assumption that “next time will be the same.” Cf. Carmen M. Reinhart & Kenneth S. Rogoff, *This Time is Different: Eight Centuries of Financial Folly* xxv (2009) (“No matter how different the latest financial frenzy or crisis always appears, there are usually remarkable similarities with past experience from other countries and from history. Recognizing these analogies and precedents is an essential step toward improving our global financial system, both to reduce the risk of future crisis and to better handle catastrophes when they happen.”).

126. See Jonathan R. Macey & James P. Holdcroft, Jr., *Failure Is an Option: An Ersatz-Antitrust Approach to Financial Regulation*, 120 YALE L.J. 1368, 1411 (2011); Peter S. Rose, *Diversification of the Banking Firm*, 24 FIN. REV. 251, 251, 260 (1989); William K. Templeton & Jacobus T.

events of 2008 also overlooks that, historically, the most common pattern in banking crises is not TBTF, but Too Many to Fail: The Great Depression and 1980s savings-and-loan crisis, for example, were both due to the contemporaneous failure of many small institutions.¹²⁷

Third, the possibility that the next crisis may take place among mid-sized institutions cannot be dismissed based on a market-share argument which points to the disproportionate concentration of the financial sector's assets in the big banks.¹²⁸ That rationale overlooks the potential for regulatory arbitrage, which is perhaps the defining problem in all of finance.¹²⁹ If only a handful of the very largest banks are stress tested, there is an incentive for assets to flow out of those institutions and into the under-regulated sector in order to chase higher returns.¹³⁰ The doubly perverse result is that any stress test exemption based on a particular size threshold should be expected to expand the share of the financial sector managed by banks below that threshold, while also making those banks more likely to fail.

The weaknesses in the compliance cost critique of financial stress testing surveyed above highlight the critical juncture at which the regulatory regime has arrived. A reform agenda based on half-measures which chip away at the Dodd-Frank rules along various margins is inappropriate because, if those procedures do in fact serve the policy goals they were designed to meet, then they should be broadly and energetically applied.¹³¹ Given the extensive criticism which the Dodd-Frank stress tests have received, a notable feature

Severiens, *The Effect of Nonbank Diversification on Bank Holding Company Risk*, 31 Q.J. BUS. & ECON. 3, 9 (1992); cf. GARY B. GORTON, MISUNDERSTANDING FINANCIAL CRISES: WHY WE DON'T SEE THEM COMING 29-43 (2012) (chronicling the constant outbreak of banking panics during the nineteenth century).

127. See RICHARD S. GROSSMAN, UNSETTLED ACCOUNT: THE EVOLUTION OF BANKING IN THE INDUSTRIALIZED WORLD SINCE 1800, at 37 (2010); see also Viral V. Acharya et al., *Too Many to Fail—An Analysis of Time-Inconsistency in Bank Closure Policies*, 16 J. FIN. INTERMEDIATION 1, 4 (2007).

128. A common estimate is that roughly 50 percent of all financial sector assets are currently held by the five largest banks. See FED. RESERVE BANK OF CHI., TOP BANKS AND HOLDING COMPANIES: JUNE 30, 2015, at 1 (2015), available at http://www.chicagofed.org/webpages/banking/financial_institution_reports/top_banks_bhcs.cfm [<https://perma.cc/D6NK-Q9X8>].

129. The quicksilver-like adaptability of the financial system means there is no guarantee the current levels of concentration will remain stable over time, because the outsized market share which TBTF banks enjoy is in large part endogenous to the legal rules presently in place. See Matthew C. Turk, *The Convergence of Insurance with Banking and Securities Industries, and the Limits of Regulatory Arbitrage in Finance*, 2015 COLUM. BUS. L. REV. 967, 977-78 [hereinafter Turk, *Limits of Regulatory Arbitrage in Finance*] (discussing the financial arbitrage issue in the context of the insurance industry); see also Thomas W. Merrill, *A Comment on Metzger and Zaring: The Quicksilver Problem*, 78 LAW & CONTEMP. PROBS. 189, 189-90 (2015) (providing the quicksilver metaphor).

130. In fact, a recent study focusing on the market for small business loans found that this process is already underway based on the current stress test size thresholds. See generally Cortés et al., *supra* note 121 (finding that banks which are stress tested tend to lower credit allowances and increase interest rates for small business loans).

131. See *supra* Section III.A.3 (on compliance costs).

of that commentary is that it rarely examines whether the affirmative case which proponents have laid out for the post-crisis rules holds up. The next Section takes up that task.

B. THE INCOMPLETE CASE FOR

Advocates of the post-crisis stress testing regime identify a wide range of benefits which those programs are thought to provide. Among the leading rationales are that the stress tests disseminate valuable information about the overall stability of the financial system; reduce the ability of Too Big to Fail institutions to take out-sized risks; and, enhance the effectiveness of other rules in the post-crisis toolkit. As argued in the following discussion, a convincing policy justification for the Dodd–Frank stress tests has yet to be established. In some cases, the anticipated benefits are unlikely to be forthcoming, no matter how the process is designed. In others, it appears that stress testing could serve a useful function but does not do so as currently structured.

1. The Information-Producing Function

The most prominent justification for regulatory stress tests focuses on their role as an information-producing device. Whom, exactly, do they inform? A useful summary of the standard answer can be found in a report published by the Federal Reserve Board in 2013, where it states:

[S]tress tests are intended to provide BHC [Bank Holding Company] management and boards of directors, the public, and supervisors with forward-looking information to help identify downside risks and the potential effect of adverse conditions on capital adequacy of these large banking organizations.¹³²

Under this account, information generated through the stress testing process is consumed by three groups: (1) the public, which can be taken to mean financial market investors in general; (2) the banks being tested; and (3) the regulators who administer those tests. Each of these audiences is considered in turn.

i. Information for Markets

The informational benefits that stress tests provide investors potentially take two different forms, depending on the context in which they are performed. The key distinction is whether those tests are run during the

132. See BD. OF GOVERNORS OF THE FED. RES. SYS., DODD–FRANK ACT STRESS TEST 2013, *supra* note 63, at 3; see also Daniel K. Tarullo, Governor, Bd. of Governors of the Fed. Reserve Sys., Speech at the Federal Reserve Third Annual Stress Test Modeling Symposium, Boston, Massachusetts: Stress Testing After Five Years (June 25, 2014) [hereinafter Tarullo, Stress Testing After Five Years]; Tarullo, *supra* note 47.

“wartime” atmosphere of an economic downturn or conducted under non-recessionary “peacetime” conditions.¹³³

a. Stress Testing in Wartime

When performed during a recession, stress testing is primarily intended to serve as a crisis management tool which stabilizes markets by disclosing information to panicked investors that restores confidence in the financial system.¹³⁴ The view that regulatory stress tests are able “to calm” otherwise turbulent financial conditions has grown into conventional wisdom since the positive market response to the Federal Reserve’s emergency SCAP test in 2009.¹³⁵ Indeed, that assumption is so widely accepted that some commentators move on to the question of whether regulators’ ability to wield such influence over financial markets is a good thing.¹³⁶

From a theoretical perspective, however, the claim that stress tests can boost investor confidence by clarifying the underlying health of the financial system raises a puzzle. By definition, if the information which a stress test reveals allows investors to update their existing beliefs, it is as likely to be positive as negative.¹³⁷ Yet, a test that releases bad news should be expected to accelerate a run on the system, not restore market confidence.¹³⁸ That would appear to make stress testing a risky proposition as a crisis management device because, *ex ante*, the information those procedures generate is equally capable of destabilizing fragile financial markets as it is to calm them.¹³⁹

133. See Til Schuermann, *Stress Testing in Wartime and in Peacetime*, in *STRESS TESTING AND MACROPRUDENTIAL REGULATION: A TRANSATLANTIC ASSESSMENT* 125, 125–27 (Ronald W. Anderson ed., 2016).

134. See *id.* at 125–26.

135. See *Maintaining Financial Stability*, *supra* note 46; see also Baradaran, *supra* note 8, at 1287 (“[T]he [stress] tests and their results quickly became a way for the Federal Reserve to calm the markets and restore confidence in the banking system”); Gilson & Kraakman, *supra* note 8, at 358 (“From our perspective, the [SCAP] Stress Test resurrected the market in inter-bank lending by generating new information about the credit worthiness of the largest U.S. banks.”).

136. See, e.g., Baradaran, *supra* note 8, at 1299 (“There is something particularly troubling about regulators using stress tests to calm markets.”).

137. Cf. Eugene F. Fama, *Random Walks in Stock Market Prices*, 51 *FIN. ANALYSTS J.* 75, 79 (1995) (“Even if the analyst is no better than the random selection procedure, in any given comparison there is still a 50 percent chance that the security he has chosen will outperform the randomly selected security.”).

138. See Goldstein & Sapra, *supra* note 110, at 31–34.

139. The same conclusion also applies under more complicated models, where regulators and banks are assumed to hold private information which the stress tests are then used to signal to markets. See *id.* The issue with this theory is that Dodd–Frank mandates stress testing on an annual basis. *Id.* at 7. The stress tests must therefore be run during every future recession, and there is no reason to believe that the positive information that regulators were thought to possess in spring of 2009 was more than happenstance. In the next downturn, banks and regulators are just as likely to have a pessimistic outlook relative to markets and stress testing will force the confidential basis for that belief to be disclosed.

The lack of an obvious theoretical basis for stress testing during wartime is compounded by limited empirical evidence that they work well in practice. The only test that has ever been credited with reducing uncertainty during a financial crisis is the 2009 SCAP, and usually it is dangerous to make inferences from a sample-size of one.¹⁴⁰ There is also a counterexample. From 2009 to 2011, the European Central Bank ran a series of stress tests, closely modeled after the SCAP, which aimed to contain the Eurozone financial crisis.¹⁴¹ But those tests did nothing to stem the deterioration of European financial markets, and were immediately followed by the collapse of banking sectors in countries such as Ireland, Italy, Spain, Belgium, Portugal, and Greece.¹⁴²

The ECB's failure to replicate the Federal Reserve's success from 2009 is often reconciled with an argument that, despite a superficial resemblance between the U.S. and European stress tests, the latter were administered improperly. Specifically, while the SCAP's methodology "was publicly disclosed so that its credibility could be independently evaluated" (and then found to be credible),¹⁴³ the ECB tests were undermined by their unrealistic modeling assumptions.¹⁴⁴ The contrast between the two, however, is at best a matter of degree. When the Treasury Department first rolled out its stress testing plans in early 2009, the initiative was widely mocked as an act of political theater,¹⁴⁵ and there was more to that reception than public cynicism

140. Cf. GARY KING ET AL., DESIGNING SOCIAL INQUIRY: SCIENTIFIC INFERENCE IN QUALITATIVE RESEARCH 55–65 (1994) (providing some analytical ground rules for making "descriptive inferences" from small amounts of observational data).

141. See generally COMM. OF EUROPEAN BANKING SUPERVISORS, AGGREGATE OUTCOME OF THE 2010 EU WIDE STRESS TEST EXERCISE COORDINATED BY CEBS IN COOPERATION WITH THE ECB (2010), available at <https://eba.europa.eu/sites/default/documents/files/documents/10180/15938/95030af2-7b52-4530-afe1-f067a895d163/Summaryreport.pdf> [<https://perma.cc/BH4N-MFJ2>] (detailing the objectives, methodologies, and outcomes of EU-wide supervisory stress test); EUROPEAN BANKING AUTH., 2011 EU-WIDE STRESS TEST AGGREGATE REPORT (2011), available at https://eba.europa.eu/sites/default/documents/files/documents/10180/15935/54a9ec8e-3a44-449f-9a5f-e820cc2c2foa/EBA_ST_2011_Summary_Report_v6.pdf [<https://perma.cc/ZH85-WE3Z>] (same).

142. See generally Matthew C. Turk, *The Banking-Sovereign Nexus: Law, Economics & Policy*, 55 COLUM. J. TRANSNAT'L L. 592 (2017) (detailing the Eurozone financial crisis).

143. Gilson & Kraakman, *supra* note 8, at 358.

144. David Enrich, *New Doubts on EU Bank Stress Tests*, WALL ST. J., <https://www.wsj.com/articles/SB10001424052748704720004575377202517842246> [<https://perma.cc/TQ92-XX6W>] (last updated July 20, 2010, 12:01 AM); Patrick Jenkins & Brooke Masters, *Bank Watchdog Sets Out to Square the Circle*, FIN. TIMES (Feb. 14, 2011), <https://www.ft.com/content/d70dd886-3865-11e0-959c-00144feabdc0> [<https://perma.cc/FZK3-FPPA>].

145. For example, Saturday Night Live ran a skit in which Tim Giethner brainstormed ideas on how he intended to rig the tests. See *Saturday Night Live Season 34, Episode 21: A Special Address from the Secretary of the Treasury* (NBC television broadcast May 9, 2009); see also Gretchen Morgenson, *Stress Tests Are Over. The Stress Isn't.*, N.Y. TIMES (May 9, 2009), <https://www.nytimes.com/2009/05/10/business/10gret.html> [<https://perma.cc/56GQ-6H9N>]; David Wessel, *Bank Checkup Also Tests Regulators*, WALL ST. J., <https://www.wsj.com/>

during hard times. Although the Treasury did release some of its methodology, that disclosure was only partial, and was set forth in a twenty-page document which most financial professionals at the time considered too vague to evaluate in a meaningful way.¹⁴⁶ Moreover, the most crucial items to be disclosed—the “normal” and “adverse” economic scenarios which banks were asked to project—were overly optimistic on their face.¹⁴⁷ The SCAP’s reputation as a market-calming device therefore rests on a fair amount of post hoc reasoning: Since the stress test results reduced uncertainty, they must have contained credible information.

A further reason why it is difficult to interpret the market response to the SCAP is that it was at all times entangled with an accompanying bailout package, the Treasury Department’s Capital Assistance Plan. A telling incident in that respect came when both programs were announced in conjunction on February 25, 2009: Financial markets shot upwards at the news, before the relevant testing procedures had been substantively outlined or executed.¹⁴⁸ Even more telling is the reaction upon release of the Federal Reserve’s SCAP results in May of 2009. Although there was no movement in the stock price of the nine banks which passed, shares in all ten banks that *failed* the SCAP test rose significantly.¹⁴⁹ The simplest explanation is that the market had been informed of a pre-approved bailout for the weaker banks, rather than the stability of the institutions being tested.¹⁵⁰ Here the European

articles/SB123983475012122683 [https://perma.cc/9495-4E2T] (last updated Apr. 16, 2009, 12:01 AM).

146. See *supra* note 35 and accompanying text (citing to the relevant regulatory documents); see also Edmund L. Andrews & Eric Dash, *Government Offers Details of Bank Stress Test*, N.Y. TIMES (Feb. 25, 2009), <https://www.nytimes.com/2009/02/26/business/economy/26banks.html> [https://perma.cc/P3JC-XLBP]; J.V. Rizzi, *Stress Tests Failed the Public*, AM. BANKER (Mar. 20, 2013, 11:50 AM), <https://www.americanbanker.com/opinion/stress-tests-failed-the-public> [https://perma.cc/KR42-GYQB] (“The Federal Reserve has not disclosed the details of their model. Thus, it is difficult to determine if they ran the appropriate tests . . .”).

147. The unemployment rate projected under the more pessimistic “adverse” scenario was exceeded within a few months. See CONG. STRESS TEST REP., *supra* note 41, at 5; David Ellis, *Watchdog Wants Stress Test Do-Over*, CNN MONEY, https://money.cnn.com/2009/06/09/news/companies/tests_warren/index.htm [https://perma.cc/X6XN-WN2L] (last updated June 9, 2009, 6:48 AM).

148. See generally Paul Glasserman & Zhenyu Wang, *Valuing the Treasury’s Capital Assistance Program*, 57 MGMT. SCI. 1195 (2011) (estimating the option value of access to the Federal Reserve’s securities at the time the Capital Assistance Program was announced).

149. See CONG. STRESS TEST REP., *supra* note 41, at 4.

150. Peristian and co-authors provide the most prominent alternative explanation that focuses on the role of information. They argue that: (1) share prices rose for the failing banks because, although they failed, the underlying loss estimates were better than investors expected and (2) investors expected the passing banks to pass and were unsurprised by the loss estimates for those institutions. STAVROS PERISTIAN ET AL., FED. RESERVE BANK OF N.Y., STAFF REP. NO. 460, THE INFORMATION VALUE OF THE STRESS TEST AND BANK OPACITY 21–22 (2010), available at https://www.newyorkfed.org/medialibrary/media/research/staff_reports/sr460.pdf [https://perma.cc/T7XN-XEQS]; see also Gilson & Kraakman, *supra* note 8, at 358–59 (adopting interpretation of PERISTIAN ET AL., *supra*). It is hard to imagine a plausible *ex ante* distribution of investor beliefs that would make this result possible,

experience is once again instructive, because the ECB's stress tests were clearly decoupled from its bailout decision, and the impact of each measure was unambiguous. After three rounds of ECB stress tests from 2009 to 2011 failed to calm European financial markets, ECB Chairman Mario Draghi declared in July of 2012 that the central bank would "do whatever it takes to preserve the euro," and that pledge of open-ended public subsidies soon brought the Eurozone crisis to an end.¹⁵¹

Thus, to the extent stress tests can calm markets with information, that information likely concerns the competency and intentions of regulators rather than the economic fundamentals of the banking sector itself.¹⁵² In that sense, the SCAP could have been seen as demonstrating Treasury Secretary Geithner's managerial skill or, what is most likely, his commitment to extend the 2008 TARP bailouts indefinitely until the financial system had fully recovered.¹⁵³ From a policymaking standpoint, the fact that an emergency stress test revealed such information does not carry any obvious implications. Every action a government takes during a financial crisis signals markets to its expertise or willingness to issue subsidies.¹⁵⁴

b. Stress Testing in Peacetime

Regulatory stress testing under the peacetime conditions of a normal financial environment has a separate justification, regardless of whether those procedures are useful as a crisis management measure. When stress tests disclose negative information about the banking sector in an otherwise healthy economy, they are unlikely to have a destabilizing effect, because such

however. The theory requires that: (1) markets were perfectly informed as to the stability of nine banks; (2) imperfectly informed as to ten banks, in a way that was overly pessimistic in every case; and (3) the Federal Reserve's novel (and to a certain degree arbitrary) set of criteria perfectly sorted all 19 banks into these two groups.

151. Mario Draghi, President, European Cent. Bank, Address at the Global Investment Conference in London (July 26, 2012) ("Within our mandate, the ECB is ready to do whatever it takes to preserve the euro. And believe me, it will be enough."); *see also* Eshe Nelson, *Five Years Ago Today, Mario Draghi Saved the Euro*, QUARTZ (July 26, 2017), <https://qz.com/1038954/whatever-it-takes-five-years-ago-today-mario-draghi-saved-the-euro-with-a-momentous-speech> [<https://perma.cc/XZ8V-578S>].

152. *See* Bertrand Candelon & Amadou N.R. Sy, *How Did Markets React to Stress Tests?* 18–19 (Int'l Monetary Fund, Working Paper No. 15/75, 2015), *available at* <https://www.imf.org/external/pubs/ft/wp/2015/wp1575.pdf> [<https://perma.cc/SNL7-57PZ>].

153. *See* Gorton, *supra* note 45, at 983–84 (highlighting Tim Geithner's personal view that the essential crisis response strategy was to shore up bank balance sheets by tapping public funds with "overwhelming force").

154. For example, Tim Geithner's first speech as Treasury Secretary was panned by markets because his remarks were perceived to be vague and ill-prepared. Yet that episode clearly does not provide any lesson about the value of policy speeches as a crisis management tool. *See* Eric Dash & Jack Healy, *Stocks Slide as New Bailout Disappoints*, N.Y. TIMES (Feb. 10, 2009), <https://www.nytimes.com/2009/02/11/business/11markets.html> [<https://perma.cc/95YC-NS9T>] (reporting that "[i]nvestors said [Geithner's] failure to publicly detail the mechanics of the plan forced them to make wild assumptions about various players in the industry").

information will tend to reveal weaknesses that are idiosyncratic to an outlier handful of mismanaged firms rather than spark a run on the system as a whole. As a consequence, the release of either positive or negative stress test results is equally beneficial in financial peacetime. Both outcomes make the banking sector more efficient by facilitating the allocation of capital toward its most productive use.

The argument for stress testing during the normal course is therefore hard to fault in the abstract. It simply tracks the common intuition that information is the lifeblood of financial markets; the more that is available, the better they work.¹⁵⁵ This is where the past decade of experience with the Dodd–Frank Act becomes critical. As of 2010, it was entirely plausible to believe that some form of regulatory stress testing could provide valuable information to financial markets. At the same time, it was impossible to know with any certainty if the specific procedures devised under Dodd–Frank would meet that goal. With the benefit of hindsight, however, the policy analysis can now put theoretical speculation aside and become more concrete. The relevant question is whether the past several years of peacetime stress testing have in fact delivered their anticipated informational benefits in practice.

A review of the post-crisis stress tests suggests that they have not. According to a study by Professors Glasserman and Tangirala, there was no statistically significant market response when the banking agencies released their stress test results for 2014 or 2015.¹⁵⁶ Another event study, by Professor Ekaterina Neretina and co-authors, looked at the period from 2009 to 2015 and likewise found that stress tests had little effect on equity returns of large U.S. banks in most years.¹⁵⁷ From an efficient-market-hypothesis perspective (in which asset prices are assumed to reflect all publicly available information), the absence of any market movement upon the announcement of those results indicates that investors do not consider the Dodd–Frank tests informative.¹⁵⁸ In light of light of voluminous data and qualitative analyses which the Federal Reserve issues each year as part of its DFAST and CCAR programs—how can this be? There are two explanations.

First, while the Dodd–Frank stress tests disclose plenty of raw information, it is not new to sophisticated financial market participants. In substance, that information is redundant with what investors can already glean from the extensive disclosures that banks must make on a quarterly and

155. See Gilson & Kraakman, *supra* note 8, at 362; Maintaining Financial Stability, *supra* note 46 (“Even outside of a period of crisis, the disclosure of stress test results and assessments provides valuable information to market participants and the public, enhances transparency, and promotes market discipline.”).

156. Glasserman & Tangirala, *supra* note 108, at 17–18.

157. Ekaterina Neretina et al., *Banking Stress Test Effects on Returns and Risks* 11 (De Nederlandsche Bank, Working Paper No. 419, 2015), available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2423720 [<https://perma.cc/WW64-JAQF>].

158. Glasserman & Tangirala, *supra* note 108, at 17–18.

annual basis pursuant to the federal securities laws and related banking regulations.¹⁵⁹ The banks' public accounting statements often span hundreds of pages and are sufficiently detailed, it turns out, that financial analysts have been able to run their own stress tests models from scratch based on the balance sheet data those documents contain.¹⁶⁰

Second, to the extent the Dodd–Frank tests disclose information that is genuinely new—in the sense that it departs from prevailing market estimates—those projections are perceived as unreliable and disregarded. There are a number of reasons why. For one, investors have cause to suspect that the Federal Reserve strategically biases its results. As found in a study by Professor Andrew Bird and co-authors, for example, the banking agencies calibrate their models in a way that systematically favors larger banks.¹⁶¹ The banks successfully bias the results too. Research by Professor Marcia Millon Cornett and co-authors concludes that banks are often able to frontload capital in the months leading up to their stress tests, and then deleverage in subsequent quarters.¹⁶² Markets therefore do not react to the snapshot view which the stress test results reflect, because investors know that, by the end of the year, the particular bank balance sheet which the Federal Reserve examined will no longer exist.¹⁶³

Taken together, the preceding observations indicate that the post-crisis stress tests have not provided financial market participants with a source of credible, new information about the banking sector, whether administered during a crisis or in a stable economic environment. As a result, there is limited justification for the current Dodd–Frank programs on grounds that they are able to calm or inform markets, at least as to fundamentals of the financial system. When those programs do produce valuable information, it is

159. See, e.g., The Goldman Sachs Group, Inc., Annual Report (Form 10-K) (Nov. 28, 2008); Citigroup Inc., Annual Report (Form 10-Q) (June 30, 2009).

160. Glasserman & Tangirala, *supra* note 108, at 83. In fact, there is some indication that publicly available information allows financial analysts to tailor those models so that they predict the Federal Reserve's own results before they are released. See Beverly Hirtle et al., *Becoming More Alike? Comparing Bank and Federal Reserve Stress Test Results*, FED. RES. BANK N.Y. (July 21, 2014, 7:00 AM), <https://libertystreeteconomics.newyorkfed.org/2014/07/becoming-more-alike-comparing-bank-and-federal-reserve-stress-test-results.html> [<https://perma.cc/QMW2-D6V8>].

161. Andrew Bird et al., *Bias and the Efficacy of Stress Test Disclosures*, TEPPER SCH. BUS. CARNEGIE MELLON U., July 3, 2019, at 1, 4.

162. Marcia Millon Cornett et al., *An Examination of Bank Behavior Around Federal Reserve Stress Tests*, J. FIN. INTERMEDIATION, May 6, 2018, at 1, 12. Capital One was arguably able to pass its 2017 stress test with this strategy. See Trefis Team, *Fed's Conditional Approval a Minor Hiccup for Capital One's Capital Return Plan*, FORBES (June 30, 2017, 10:10 AM), <https://www.forbes.com/sites/greatspeculations/2017/06/30/feds-conditional-approval-a-minor-hiccup-for-capital-ones-capital-return-plan> [<https://perma.cc/J6JL-GN9Q>].

163. Banks' near-perfect passage rate for the quantitative stress tests—no institution has failed a quantitative CCAR tests since 2013—may further reinforce investors' view that nothing is to be learned from their results because the banks are able to game the banking agencies' modeling assumptions in advance. See *supra* Section II.D (providing results from the first decade of Dodd–Frank stress tests).

mainly as a matter of accident, not design, as that information turns on what it implicitly reveals about the assumptions of regulators running the tests. This need not be the case, however, because as this Article's proposal will show, there are ways that stress testing can uncover meaningful insights into economic fundamentals of the financial system, and also signal markets about the working assumptions of banking regulators in a more explicit manner.¹⁶⁴

ii. Information to Banks

A second potential audience for the Dodd–Frank stress tests is the banks themselves. In theory, banks might benefit from the information those tests provide in two ways. Most directly, bank managers could use the stress test results to gain insight on how to better run their institutions.¹⁶⁵ And more indirectly, the same benefit may also be forthcoming through a form of market discipline: A negative market reaction to the results of a stress test sends a signal to managers that investors believe the bank is slipping into a weak financial position.¹⁶⁶ As argued by Professors Ronald J. Gilson and Reinier Kraakman, stress tests thereby function as a kind of “early warning” device for bank officers and directors, because “repeated stress tests would inform the market of the capital deficiencies of banks in time to allow management to address the capital gap well short of the point at which it might initiate a bank run.”¹⁶⁷

An immediate limitation of the theory that stress tests inform bank managers is that, taken on its face, it does not follow that those procedures need to be mandated and supervised by regulators. If stress tests provide information that managers find valuable—and the rise of stress testing as an industry practice since the late 1980s suggests they often do—managers would have an incentive to adopt them voluntarily as part of the bank's internal risk-management protocols. The missing conceptual distinction turns on information as a private versus public good. Whenever stress tests tend to uncover hidden risk factors that a bank manager would find actionable, they confer a private benefit upon the bank and do not need to be imposed as a legal requirement. For that reason, regulatory stress tests are often assumed to provide a public good, by remedying under-investment in information about systemic risks that no individual bank has an interest in reducing.¹⁶⁸

164. See *infra* Part IV.

165. See Tarullo, *Stress Testing After Five Years*, *supra* note 132.

166. See Goldstein & Sapra, *supra* note 110, at 12.

167. Gilson & Kraakman, *supra* note 8, at 362.

168. Haldane, *supra* note 9, at 6, 9 (characterizing regulatory stress tests as providing an “missing informational public good”). Nor is it plausible that managers benefit from the stress test results as uniform benchmark, because benchmarking their institution's performance relative to competing banks is most likely feasible based on publicly available information. See *supra* note 158 and accompanying text.

Banks are not the relevant audience for stress testing under a public goods rationale.

The early warning device theory also misconceives the institutional setting in which stress testing is situated, as well as the underlying market failures that banking regulation aims to address. As designed, the Dodd–Frank stress tests do not attempt to estimate the possibility of a run on the bank. Instead, they measure whether a bank has maintained a buffer of capital that exceeds the existing minimum requirements by such a wide margin that it will still be able to meet them in the event of an economic downturn. But a premise of those minimum requirements is that market forces, if left alone, will *encourage* banks to be overleveraged.¹⁶⁹ Shareholders therefore cannot be expected to discipline managers based on negative stress test results, because shareholders would prefer there were no capital requirements in the first place.¹⁷⁰ A similar logic also applies to the monitoring role of creditors, who may lack an incentive to discipline managers for negative stress test results because they are insured against the downside of bank failures in various ways.¹⁷¹

In short, the claim that stress tests provide banks with valuable information encounters considerable analytical hurdles which put the regulatory benefits anticipated by that theory in doubt. Moreover, setting

169. Shareholders prefer leverage because it provides a higher rate of return on corporate profits—for financial institution shareholders in particular, leveraged returns are essentially a free lunch because they can be achieved without compensating creditors for the bank's greater risk of default. See generally ANAT ADMATI & MARTIN HELLMIG, *THE BANKERS' NEW CLOTHES: WHAT'S WRONG WITH BANKING AND WHAT TO DO ABOUT IT* (2013) (explaining why financial institutions generally prefer highly levered capital structures); Abel Elizalde & Rafael Repullo, *Economic and Regulatory Capital in Banking: What is the Difference?*, 3 INT'L J. CENT. BANKING 87, 88 (2007) (exploring the distinction between the amount of "economic capital" that market forces encourage banks to maintain with the level of "regulatory capital" required under banking laws).

170. With a leveraged capital structure, shareholders are also less sensitive to a bank's exposure to remote downside risks, since creditors would absorb most losses in the event of a severe downturn. For these reasons, a more complicated agency cost theory—which assumes capital markets may rely on stress test results as a corporate governance device that helps to detect managerial misbehavior—is also misplaced. But see Gilson & Kraakman, *supra* note 8, at 362 n.133 (proposing such a theory).

171. This is the problem of moral hazard. Depositors are indifferent to the possibility of bank failure because their deposits are federally insured by the FDIC. Moreover, it is generally assumed that the FDIC is unable to effectively charge banks for access to the insurance fund in a way that accounts for their risk profile. See generally Yuk-Shee Chan et al., *Is Fairly Priced Deposit Insurance Possible?*, 47 J. FIN. 227, 227 (1992) (analyzing the barriers to the FDIC charging banks for deposit insurance based on institutions-specific risk profiles). Non-depository creditors also behave as if they are partially insured against a bank's default, because they anticipate a publicly subsidized bailout may be forthcoming from the Federal Reserve in its role as a lender of last resort. See generally John Crawford, *The Moral Hazard Paradox of Financial Safety Nets*, 25 CORNELL J.L. & PUB. POL'Y 95 (2015) (examining the role of moral hazard posed by bank bailouts in the recent financial crisis); Kathryn Judge, *The First Year: The Role of a Modern Lender of Last Resort*, 116 COLUM. L. REV. 843 (2016) (detailing how the Federal Reserve could leverage its lender of last resort function in light of the moral hazard problems that stem from its emergency lending activities).

those theoretical difficulties aside, there is once again the problem of limited empirical evidence that those benefits have materialized over the past decade of post-crisis stress testing. The pattern of market responses to results released pursuant to the Dodd–Frank tests does not include a clear instance where the early warning device mechanism those programs are thought to introduce has been triggered to-date.¹⁷² Accordingly, this Article’s proposal proceeds under the assumption that other regulatory interventions are better suited for informing bank managers, and focuses instead on the role of banks as producers (rather than consumers) of the informational public good that stress testing may provide.¹⁷³

iii. Information to Regulators

The federal banking agencies responsible for administering Dodd–Frank’s stress tests are widely considered to be another end-user of the information generated by those procedures. The claim that regulators obtain valuable information from stress testing is usually framed in terms of a related pair of benefits.¹⁷⁴ First, the stress test programs enable the agencies to engage in “macro-surveillance” of the banking sector, by alerting regulators to the emergence of systemic risks which banks or markets are otherwise unable to spot.¹⁷⁵ Second, stress tests serve a “regulatory learning” function because, as part of the macro-surveillance process, agencies learn ways to improve the supervisory practices and regulatory rulemakings that apply to the banking system.¹⁷⁶

a. Macro-Surveillance

One of the most under-appreciated aspects of the Dodd–Frank stress tests is that they cannot produce information that facilitates macro-surveillance because, as structured, they do not attempt to surveil any systemic risks. “Systemic risk,” as it is usually understood, refers to risks that arise from the

172. At least over the past decade, the market participants responsible for sounding an early warning appear to have been asleep at the wheel, since the announcement of results for the Dodd–Frank stress tests usually does not lead to any market reaction, positive or negative. *See supra* Section III.B.1.i.a. The most prominent exception to that indifference is the SCAP, where share prices rose sharply for the banks that failed. Neither outcome is consistent with market discipline of bank managers.

173. *See infra* Section IV.

174. There are two possibilities for how this could take place. First, even if the publicly available stress tests results do not inform bank managers or market investors, it may be that regulators have some unique expertise that gives them a comparative advantage in how to interpret that information. Second, and more plausibly, the process of administering the stress tests may reveal private information to regulators which can then be put to good use on a confidential basis.

175. *See, e.g.*, Gilson & Kraakman, *supra* note 8, at 362; Tarullo, *Stress Testing After Five Years*, *supra* note 132.

176. *See* Gilson & Kraakman, *supra* note 8, at 362.

network of interconnections among banking institutions which allow an episode of financial distress at one firm to destabilize others.¹⁷⁷ But Dodd–Frank’s DFAST and CCAR stress testing programs do not ask banks to measure their vulnerability to adverse conditions in other parts of the financial system. Instead, the hypothetical scenarios those procedures rely upon ask each bank to predict how it would be affected by a conventional recession involving a drop-off in economic activity at non-financial firms.¹⁷⁸ In effect, the post-crisis stress tests assume that each bank is the only bank in the world. As a result, they do not provide any systemic, macro-prudential view of the banking sector, and instead follow the same “micro-prudential” approach to financial supervision that the banking agencies have been applying for decades.¹⁷⁹

The limitations of that approach can be seen in the operation of the stress tests themselves. Consider the inaugural SCAP test from 2009. Pursuant to that exercise, the Federal Reserve determined that ten of the nineteen largest U.S. financial institutions were unprepared to withstand a scenario in which economic indicators relating to employment and GDP took a turn for the worse within the next 18 months.¹⁸⁰ It also determined that Goldman Sachs and eight other Too Big to Fail banks would stay well-capitalized under those conditions during the same period. When taken together, this pair of results raises a further question: Assuming the Federal Reserve was able to project subsequent economic developments with perfect foresight—how was it able to conclude that Goldman Sachs would remain stable amidst the contemporaneous failure of roughly half the U.S. banking system? With the SCAP, no such conclusion was ever reached because the question was never asked; the test for Goldman Sachs was run on the premise that the other ten

177. For more detailed explorations of the concept, see generally Steven L. Schwarcz, *Systemic Risk*, 97 GEO. L.J. 193 (2008) (offering a framework for systemic risks and how they should be regulated). See also Turk, *Limits of Regulatory Arbitrage in Finance*, *supra* note 129, at 989–1000 (explaining how insurance companies are closely interconnected with banks and now vulnerable to systemic risks).

178. In other words, the question the stress tests ask is not: How would AIG be affected by a bankruptcy at Lehman Brothers? Rather, it is: How would AIG be affected if consumers reduced their spending, producers sold fewer goods, and firms hired fewer workers? See Policy Statement on the Scenario Design Framework for Stress Testing, 12 C.F.R. § 252 app. A (2015); see also *supra* Section II.C (providing an overview of the stress testing criteria). In technical jargon, this means that the stress tests examine the impact of an exogenous shock to the real economy, rather than risks which are endogenous to the financial system itself.

179. See David Greenlaw et al., *Stressed Out: Macroprudential Principles for Stress Testing* 4–7 (Univ. of Chi. Booth Sch. of Bus., Working Paper No. 71, 2012), available at <https://poseidon01.ssrn.com/delivery.php?ID=g8o119000o72116064126091087065017098118002064083044031109124065112110092005101014025056030043025033007124027100122117082113112061055008035004077123088107006066071001010073004002087113103087118005006027095105011028119116067098100090031002011006065024&EXT=pdf> [<https://perma.cc/RS6F-gXFA>] (observing the micro-prudential nature of post-crisis stress testing); Claudio Borio et al., *Stress-Testing Macro Stress Testing: Does It Live up to Expectations?*, 12 J. FIN. STABILITY 3, 4 (2014) (same).

180. See *supra* Section II.B (providing an overview of the SCAP).

banks which failed did not exist. From a macro-prudential perspective focused on the potential domino effect of negative spillovers across financial firms, that premise sidesteps the essential inquiry. The same oversight explains the Federal Reserve's seemingly self-contradictory results, which are only possible to reconcile through a highly stylized micro-prudential lens that views each financial institution in isolation.

The stress test programs subsequently established under Dodd–Frank depart from the SCAP in a number of ways while nonetheless retaining its essentially micro-prudential posture. The most notable exception primarily serves to prove the rule. Under the qualitative CCAR test, each bank is required to identify one other financial institution that is a business counterparty and provide a narrative account of the steps it might take in response to the failure of that firm.¹⁸¹ Thus, within the bounds of that short verbal description, Dodd–Frank extends the SCAP's recessionary scenarios to a hypothetical economy where there are two banks rather than one. As should be clear, that modest step toward greater realism still does not present regulators with a meaningful basis to survey the complex network of interconnections among banking institutions that is the source of systemic risk. It also implies the post-crisis stress tests yield minimal information that is relevant for macro-surveillance of the financial system.

b. Regulatory Learning

The absence of a macro-prudential component to the Dodd–Frank stress tests largely undercuts the regulatory learning hypothesis, which is generally premised on the idea that those procedures can inform the banking agencies' efforts to mitigate systemic risks within the financial sector. It is also doubtful that the post-crisis stress tests are able to facilitate regulatory learning from a more narrow, micro-prudential perspective.

One reason is that it does not appear that the federal banking agencies have developed the institutional capacity to ensure that the quantitative models which they rely on when administering the stress tests are methodologically sound. Notably, an audit conducted by the Office of the Inspector General in 2015 found that the Federal Reserve did not have any formal process in place for validating the modeling techniques which it applied.¹⁸² Nor was it plausible that those kinds of checks were run on a more informal basis: The same audit also found that the Federal Reserve did not

181. See Capital Planning, 12 C.F.R. § 225.8 (2017).

182. OFFICE OF INSPECTOR GEN., THE BOARD IDENTIFIED AREAS OF IMPROVEMENT FOR ITS SUPERVISORY STRESS TESTING MODEL VALIDATION ACTIVITIES, AND OPPORTUNITIES EXIST FOR FURTHER ENHANCEMENT 9–11 (2015), *available at* <https://oig.federalreserve.gov/reports/board-supervisory-stress-testing-model-validation-reissue-oct2015.pdf> [<https://perma.cc/67LW-Q2JD>].

employ any financial analysts on staff who would have the necessary training to perform such a task.¹⁸³

The fact that minimal technical resources have been committed to designing a rigorous stress testing process is especially disconcerting once the specific methodological shortcomings of those models are fully appreciated. As demonstrated in an important study by financial economist Viral Acharya and co-authors, the Federal Reserve's hypothetical stress scenarios are constructed around a circular set of parameters, which makes it inevitable that their results do no more than restate the risk-weights which regulators have previously specified in the capital requirements being tested.¹⁸⁴ The consequence, they explain, is that estimates from the "stress tests have no [logical] link with the realized risk of banks during a crisis."¹⁸⁵ Another study by economists Paul Glasserman and Gowtham Tangirala identifies an analogous dynamic at work across the three recessionary scenarios that each bank is required to run.¹⁸⁶ They conclude that any variance in outcomes among those scenarios "appear[s] to be an artifact of the stress testing process rather than an accurate reflection of potential bank losses."¹⁸⁷

To be clear, these problems are more fundamental than the modeling-error cliché of "garbage-in, garbage out"—where unrealistic assumptions or imperfect data inputs are thought to make a model's projections unreliable. Rather, the Dodd–Frank stress tests are best understood as a tautological mathematical exercise that produces no meaningful outputs at all.¹⁸⁸ This

183. *Id.* at 11. The Office of Inspector General's (OIG) 2015 report therefore lends considerable irony to Federal Reserve's determination that Bank of America failed the qualitative portion of its CCAR test in 2014 (on the grounds that the bank's internal stress testing protocols were unreliable). See Armour et al., *supra* note 65 and accompanying text. By the same standard, the OIG's report implies that the financial regulators themselves failed every stress test that was administered during the same timeframe.

184. See Viral V. Acharya et al., *Testing Macroprudential Stress Tests: The Risk of Regulatory Risk Weights* 16–17 (Nat'l Bureau of Econ. Research, Working Paper No. 18968, 2013), available at <https://www.nber.org/papers/w18968.pdf> [<https://perma.cc/KTD8-ZSD4>].

185. *Id.* at 22.

186. Glasserman & Tangirala, *supra* note 108, at 85. Glasserman and Tangirala's analysis of the 2015 and 2016 tests finds that estimated capital levels under the "severely adverse" scenario can be perfectly predicted for every bank by simply adjusting result of "adverse" scenario capital levels by 31 percent. Thus, despite the fact the scenarios being modeled, in Glasserman and Tangirala's words, "252 dimensional" mathematical objects, the variation in their outputs can always be anticipated with one step of elementary school level arithmetic. *Id.* at 86.

187. *Id.* at 88.

188. A simplified version of the Acharya (2017) analysis can illustrate this point. Imagine a bank with \$110 in assets and \$100 debt. The difference in value between the bank's assets and liabilities means that it has regulatory capital of \$10, and a leverage ratio of ten percent. Now stipulate that a recession occurs in which all assets in the economy drop by one percent. As a result, the bank's assets are worth \$109, regulatory capital drops to \$9, and it has a new leverage ratio of nine percent. It is unclear what new information about the real world is discovered through this kind of calculation.

forecloses the possibility of regulatory learning by financial authorities since, from the banking agencies' vantage point, nothing new has been learned.

Given the consensus that a self-reinforcing run on Too Big to Fail Banks was the *sine qua non* of the 2008 crisis,¹⁸⁹ it is startling to conclude that the post-crisis stress tests assume away the logical possibility of systemic risk from the outset. But that is more or less what they do. Once the methodological flaws in the Dodd–Frank procedures are laid bare, it is implausible that regulators could obtain information which would allow them to monitor systemic risks through macro-surveillance or engage in regulatory learning. If the stress test results did inform the federal banking agencies, that would be even more concerning, since there is little evidence the underlying models are rigorously validated or designed. An important lesson is that in order for regulators to learn anything of value from stress testing, the sophistication of the current Dodd–Frank programs must be substantially increased.¹⁹⁰

2. Safeguards & Stabilizes Banks

Regardless of whether the Dodd–Frank stress tests are able to disseminate valuable information, they may still serve an important function by contributing to the efficiency or stability of the financial system in other respects. The dividend and stock repurchase restrictions which the federal banking agencies are authorized to impose when firms fail the CCAR tests, for example, represent a form of command-and-control regulation that goes far beyond the disclosure of information. Accordingly, an alternative justification for those procedures turns on their ability to safeguard the banking sector by directly forcing financial institutions to raise more capital or otherwise forego risky lending strategies.¹⁹¹

Yet, if the goal is for banks to be better capitalized, it remains an open question why stress tests in particular are a helpful means to that end. The SCAP illustrates this point: As of February 2009, the Federal Reserve could have skipped its stress testing process altogether and immediately issued an order requiring the ten most leveraged Too Big to Fail banks to raise \$75 billion in additional capital.¹⁹² The relevant policy consideration is therefore, whether stress testing should be expected to stabilize the banking sector in a

189. See, e.g., Gary Gorton & Andrew Metrick, *Securitized Banking and the Run on Repo*, 104 J. FIN. ECON. 425, 447–48 (2012) (explaining that the 2008 recession was a result of “a run in the repo market”); Gorton, *supra* note 45, at 976.

190. See *infra* Section IV.A.1 (presenting a proposal on how this might be done).

191. And, even if no regulatory sanction is ever applied, the mere threat of failing a stress test may encourage banks to preemptively shore up their balance sheets by taking these measures as well. See Cortés et al., *supra* note 121, at 13; Levine, *supra* note 69.

192. That is the mandate ultimately imposed by the Federal Reserve upon its announcement of the SCAP results in March of 2009. See *supra* Section II.C. The Federal Reserve never asserted that the SCAP program established any special legal basis for issuing such a requirement, which it considered to be authorized pursuant to an emergency Congressional statute from 2008. See *supra* Section II.C.

new and useful way that is not already duplicated by other regulatory instruments in the post-crisis legal framework. There are three common rationales for why that may be the case.

First, it is often argued that, compared to pre-crisis forms of bank supervision, stress tests are able to deliver capital requirements or related legal restrictions in a manner that specifically targets financial institutions' exposure to systemic risks.¹⁹³ As explained in the discussion above, however, the Dodd–Frank stress tests have not been designed to fill that role. A consequence is that they replicate the same micro-prudential function that is performed by existing assessment procedures, and are only able to make particularized determinations that specific institutions are not safe-and-sound.¹⁹⁴ Given the modeling anomalies also detailed above, it is unlikely that the Dodd–Frank stress tests materially improve upon the precision of more traditional supervisory practices in that respect.¹⁹⁵

A second claim is that the Dodd–Frank stress tests are distinguished by their uniquely forward-looking orientation.¹⁹⁶ While existing rules establish a minimum regulatory floor that applies at all times, stress tests solve a separate problem by imposing an additional capital buffer on banks that are in full compliance with those minimums but may fall short of them in the near future. This argument also overstates the novelty of stress testing. A longstanding position of the federal banking agencies, which can be found in a number of legal areas, is that they reserve discretion to impose requirements above the stated regulatory baselines when higher standards are deemed appropriate for a particular institution.¹⁹⁷ One prominent mechanism for doing so, incorporated in statutory reforms in the early 1990s, is a set of supervisory protocols known as Prompt Corrective Action.¹⁹⁸ Under Prompt Corrective Action, if a bank's supervisory rating is considered passable, yet appears to be drifting downward, regulators may—and in some cases, must—order the bank to restrict dividends or take similar measures to boost its capital levels.¹⁹⁹ Those rules therefore reflect the same mode of preemptive, forward-looking intervention that is contemplated in the Dodd–Frank stress test programs.²⁰⁰

193. See, e.g., Tarullo, *Stress Testing After Five Years*, *supra* note 132.

194. See *supra* Section III.B.1.iii.a.

195. See *supra* Section III.B.1.iii.b.

196. See, e.g., Tarullo, *Stress Testing After Five Years*, *supra* note 132.

197. See *Comprehensive Capital Analysis*, *supra* note 8, at 73–86; Jeremy C. Kress, *Solving Banking's "Too Big to Manage" Problem*, 104 MINN. L. REV. 171, 229–32 (2019).

198. Federal Deposit Insurance Corporation Improvement Act of 1991 (FDICIA), Pub. L. No. 102–242, 105 Stat. 2253 (codified at 12 U.S.C. § 1831o (2012)).

199. 12 U.S.C. § 1831n.

200. Consistent with the animating spirit behind the stress tests, the explicit goal of Prompt Corrective Action is to preempt bank supervisors' incentive to engage in "regulatory forbearance," by committing them to issue new capital mandates before it is too late. Cf. Craig O. Brown & I. Serdar Dinç, *Too Many to Fail? Evidence of Regulatory Forbearance When the Banking Sector*

The third rationale is that stress tests represent a distinctively “counter-cyclical” form of capital regulation.²⁰¹ The need for a more counter-cyclical approach emerged as a scholarly consensus following the crisis, in order to address the dysfunctional pro-cyclical quality of minimum capital requirements that are held constant over the business cycle: Any fixed capital ratio becomes easier for banks to satisfy when asset prices are rising, but harder to meet when they are falling.²⁰² Counter-cyclical capital rules are thought to remedy both aspects of that dysfunction. During good economic times, they adjust for the implicit mismeasurement of bank capital, which is biased upward by asset bubbles.²⁰³ They also mitigate the downward pressure on capital levels that occurs during a financial crisis, which may otherwise force banks to engage in destructive “fire sales” of assets in order to meet minimum capital requirements.²⁰⁴

The counter-cyclical rationale for stress testing encounters two issues, however. First, the identical problem is specifically addressed by another novel post-crisis regulation, known as the Counter-Cyclical Capital Buffer rule, which sets forth a procedure for the federal banking agencies to raise and lower the relevant minimum capital ratios over time.²⁰⁵ Second, and more importantly, the conventional assumption that the stress tests are counter-cyclical is itself questionable. The SCAP, for example, could not have been more pro-cyclical. In the midst of the deepest economic downturn since the Great Depression, the Federal Reserve’s stress test resulted in a mandate that most Too Big to Fail banks maintain *more* capital.²⁰⁶ Dodd–Frank’s stress tests do not modify the SCAP in any respect that would eliminate that feature. The DFAST and CCAR programs are run on an annual basis, and cannot be suspended during a recession.²⁰⁷ Moreover, banks must always be in compliance with the capital levels those tests estimate under the most adverse

is Weak, 24 REV. FIN. STUD. 1378, 1400–02 (2011) (reviewing the evidence that bank supervisors engage in regulatory forbearance in deciding when to intervene in weak or failing banks).

201. With counter-cyclical rules, the relative regulatory burden is calibrated to move opposite the business cycle, by ratcheting the stringency of requirements upwards in healthy economic periods and then dialing them down during a recession. See generally Jonathan S. Masur & Eric A. Posner, *Should Regulation Be Countercyclical?*, 34 YALE J. ON REG. 857 (2017) (discussing the potential implications of counter-cyclical government regulations). On the counter-cyclical quality of post-crisis stress testing, see, for example, BEVERLY HIRTLE ET AL., FED. RESERVE BANK OF N.Y., STAFF REP. NO. 409, MACROPRUDENTIAL SUPERVISION OF FINANCIAL INSTITUTIONS: LESSONS FROM THE SCAP 1 (2009).

202. See Charles Goodhart, *Is a Less Pro-Cyclical Financial System an Achievable Goal?*, 211 NAT’L INST. ECON. REV. R17, R17–R18 (2010).

203. *Id.*

204. See *id.*

205. Countercyclical Capital Buffer, 12 C.F.R. pt. 217 app. A (2016). See generally Brett H. McDonnell, *Designing Countercyclical Capital Buffers*, 18 N.C. BANKING INST. 123 (2013) (discussing the Counter-Cyclical Buffer Rule and its likelihood of success).

206. See *supra* Section III.B.2.

207. See *supra* Section II.C.

hypothetical scenario.²⁰⁸ Because the stress tests require banks to assume a decline in economic conditions in which asset prices are falling, their most likely result during a recession will be to magnify, rather than mitigate, the pro-cyclical pressures that minimum capital ratios introduce.

It is certainly true that the Dodd–Frank stress tests impose additional restrictions on the leverage of bank balance sheets, and furthermore that measures to that effect can make the financial system safer. But to the extent there are benefits from securing a well-capitalized banking sector, it does not follow that stress tests should be the tool that policymakers use to achieve that result. A careful look at existing forms of capital requirements once again shows that the traditional banking regulation toolkit is impressively broad. It also shows that the novelty of stress testing within that scheme is more superficial than it may appear at first glance. Separating the case for stress tests from the case for “more capital,” therefore, leaves the current set of procedures in search of a coherent justification.

3. Complements other Regulations

A final justification for the post-crisis stress tests is that, direct benefits aside, those procedures are worthwhile because they indirectly make other components of the regulatory structure work better. According to a policy statement issued by the Federal Reserve, the Dodd–Frank stress tests serve as a complement to the standard capital requirement rules.²⁰⁹ Economists at the Federal Reserve have elsewhere argued that the DFAST and CCAR programs share a complementarity with one another as well.²¹⁰

Although several aspects of the stress testing process are cited in support of these claims, they each suffer from a common analytical looseness about what is being described. Technically, overlapping legal rules are complementary when they perform highly differentiated functions that intersect in a synergistic manner which yields increasing benefits from their joint use.²¹¹ By contrast, when regulations address the same policy problem by serving identical or only slightly different roles, they act as perfect or partial substitutes, respectively.²¹² Regulatory substitutes provide decreasing benefits

208. See *supra* Section II.C.

209. Stress Testing Policy Statement, 82 Fed. Reg. 59,528, 59,529 (proposed Dec. 15, 2017) (to be codified at 12 C.F.R. pt. 252) (“By assessing the capital adequacy of a covered company under severe projected economic and financial stress, the supervisory stress test complements minimum regulatory capital ratios, which reflect the covered company’s current condition.”).

210. Tim P. Clark & Lisa H. Ryu, *CCAR and Stress Testing as Complementary Supervisory Tools*, BD. GOVERNORS FED. RES. SYS., <https://www.federalreserve.gov/bankinforeg/ccar-and-stress-testing-as-complementary-supervisory-tools.htm> [<https://perma.cc/58W3-P2NH>] (last updated June 24, 2015).

211. See generally Matthew C. Turk, *Overlapping Legal Rules in Financial Regulation and the Administrative State*, 54 GA. L. REV. (forthcoming 2020) (on file with the Indiana University Kelley School of Business).

212. *Id.*

when applied in conjunction, due to a crowding-out effect where the value-added from using one rule is either partially or perfectly redundant with that of the other rule.²¹³

Strictly speaking, the Dodd–Frank stress tests do not complement any other financial regulation as currently constituted. The DFAST and CCAR programs are best interpreted as either sub-parts of a single regulatory intervention, or near-perfect substitutes. Once results from the quantitative CCAR test have been obtained, for example, performing the qualitative CCAR test becomes less important, not more: Qualitative review merely provides an alternative format for reassessing a question that has previously been answered in other terms.²¹⁴ The same logic likewise applies to Dodd–Frank’s mandate that, in addition to the DFAST process overseen by the federal regulators, banks must also run their own variation on the agency-supervised version of the test in-house.²¹⁵ More generally, the law of diminishing returns means that, almost by definition, information produced by stress testing is at best a partial substitute for related disclosures which banks must otherwise make to markets and regulators.²¹⁶

Neither do the Dodd–Frank stress tests complement other capital rules. For the reasons sketched directly above, stress testing is largely duplicative of those requirements as a practical matter. A forward-looking test of bank capital does not make existing capital ratios more effective per se, instead it is a partial substitute that attempts to fill a gap in those rules by extrapolating their requirements into subsequent periods. Because Prompt Correct Action protocols do that as well, post-crisis stress test rules are a close substitute to the statutory provisions which established those procedures in the 1990s.²¹⁷ Lastly, the Dodd–Frank stress tests do not complement the Counter-Cyclical Capital Buffer since, as of now, the buffer rule operates without any legal connection to the CCAR and DFAST programs and is administered independent of their results.²¹⁸

The notion that stress testing indirectly enhances the effectiveness of other financial regulations is often taken for granted in the policy commentary. Yet the underlying mechanism for those complementarities is

213. *Id.*

214. *But see* Clark & Ryu, *supra* note 210 (discussing the results of qualitative analyses in stress tests).

215. *But see id.* (discussing the benefits of the new regulations from the Dodd–Frank Act).

216. *But see id.* (discussing how stress testing has “fundamentally changed” the regulation of banks and the financial industry).

217. In a perceptive article on the post-crisis reforms, Professor Robert Weber explains at length why regulatory stress tests are best understood as a complicated refashioning of the old Prompt Corrective Action protocols. *See* Weber, *A Theory for Deliberation-Oriented Stress Testing Regulation*, *supra* note 8, at 2293–94 (discussing how regulatory powers have changed from “the prompt corrective action regime” under new stress testing measures).

218. *See* McDonnell, *supra* note 205, at 124, 126 (discussing the impact of the Dodd–Frank Act on counter-cyclical capital buffers).

rarely well-defined. Once that task is taken up, it is apparent that the stress tests primarily function as substitutes for other rules in the post-crisis legal framework. The basic point is summed up well in a column by financial journalist Matt Levine, where he observes that “the stress test is just a complicated form of capital regulation.”²¹⁹ This raises the possibility that, if capital requirements have been set too low in certain respects, increasing the relevant minimum ratios by a few percentage points may be a perfect substitute for stress testing that strictly dominates the use of those procedures in their present form.

This Section has reviewed the last decade of debate over the promise and limits of post-crisis stress testing and arrived at some surprising conclusions. Despite the recent push to roll back the Dodd–Frank programs along several fronts, the line of critiques which have animated those efforts are not wholly persuasive. Notably, skeptics have avoided engagement with the rationales that were offered for introducing those rules in the first place.

Revisiting those rationales proves revealing for two reasons. First, it shows that the sheer span of justifications cited by advocates of the Dodd–Frank stress test rules has obscured the fact that, when examined in detail, none of those justifications can be substantiated in a convincing manner. Second, it highlights how the post-crisis stress tests themselves were a form “regulation by hypothetical”: As of 2010, the value of those procedures was impossible to know with certainty and necessarily a matter of theoretical speculation. The problem is that—although many of those theories were at least plausible enough to provide a defensible basis for moving forward at the time—they receive little confirmation when measured against the subsequent performance of stress testing in practice.

These findings would appear to give a more radical edge to arguments that the Dodd–Frank stress tests should be reformed at the margin. If none of their anticipated benefits have been realized, a sensible case could be made that the stress tests programs can be eliminated wholesale at no loss to the overall efficacy of financial regulation. As explained below, however, accelerating a rollback of post-crisis stress testing is not ideal. The limited benefits of the current rules mainly reflect a missed opportunity, rather than any fatal flaw, and a robust role for regulatory stress testing can be salvaged. What is really needed are reforms that make the federal banking agencies’ administration of those procedures more effective.

IV. A PROPOSAL: STRESS TESTING THE FEDERAL BANKING AGENCIES

In light of the preceding analysis, this Article proposes a simple but fundamental reform to the way that post-crisis stress testing is currently performed: Those procedures should be reoriented so that they test the effectiveness of rules promulgated by the federal banking agencies, rather

219. Levine, *supra* note 69.

than banks' compliance with those rules. This would mean the stress tests are used to identify a clear set of pre-announced benchmarks that will make the financial system safe and efficient, rather than identify downturn scenarios where existing regulations might fail to meet that goal at particular financial institutions. By reimagining the post-crisis stress tests as a tool for detecting regulatory failure, it is more likely they will satisfy their original aim of preventing the kind of market failures which may lead to a systemic crisis in the banking sector. In addition, reforming the stress tests along these lines would also mitigate many of the concerns that critics of the Dodd–Frank rules have voiced. The discussion below surveys the high-level considerations relevant to implementing this proposal and explains the benefits they would provide.

A. SUMMARY OF PROPOSAL

1. Methodological Changes

A critical starting point is to remedy methodological weaknesses in the current process that were previously identified above.²²⁰ First, the tests should be focused on estimating risk of default (or temporary period of financial distress) for a hypothetical bank that is in full compliance with minimum requirements, not on the bank's ability to satisfying those requirements on a forward-looking basis. From a financial stability perspective, whether a bank is meeting its regulatory requirements during a crisis is not itself a variable of any inherent interest. The variable of interest is what those requirements should be. When minimum standards are set sufficiently high, a bank can fall out of compliance to some degree at no social cost, because non-compliance per se does not imply a bank will default on counter-party financial institutions or engage in fire sale tactics that pose a systemic risk.²²¹

Second, the banking agencies' models should measure the possibility of default or distress in light of systemic risks that are endogenous to the financial sector. While the micro-prudential posture of Dodd–Frank's stress tests leads them far astray from that goal, other precedents reflect the core principles for how it might be done. One is the stress testing protocols from the IMF's Financial Sector Assessment Program, which take an explicitly macro-prudential approach and seek to model dynamics of a country's financial system as a whole.²²² The other is the Federal Reserve's presently-neglected CLAR test, which focuses on bank liquidity rather than capital, and

220. See *supra* Section III.A.

221. Traffic regulation provides a parallel example. A person who drives 68 MPH in a 65 MPH zone has violated the legal speed limit. But that non-compliance does not imply they have endangered other drivers in a meaningful way. If highways become substantially less safe any time an individual driver exceeds the speed limit by some marginal amount, that would not reflect well on the regulators who set the speed limit.

222. See *supra* note 30 and accompanying text (describing the IMF's stress tests).

attempts to model how disruptions in short-term funding markets for bank credit may destabilize a financial institution that otherwise appears fully solvent.²²³ A rigorously-designed combination of the IMF and CLAR tests is essentially what is needed.²²⁴ This is admittedly an ambitious task, but recent work by financial economists has made progress on modeling techniques which suggest it may be practicable.²²⁵

A further implication of this approach is that the stress testing process should examine a much broader range of scenarios than the three hypothetical recessions which Dodd–Frank requires the banking agencies to apply. That is because the diagnostic challenge is not to comprehensively anticipate the most likely economic downturn that can be foreseen, and then see how each bank would perform under those conditions. Instead, it is to determine whether the regulatory framework is sufficiently robust that the financial sector will remain stable when exposed to any plausible (or even slightly implausible) set of circumstances where systemic risks may arise. This means that the banking agencies formulate a dozen or so scenarios of varying likelihood—a few of which possibly isolate an extreme collapse in one specific corner of financial markets or are otherwise somewhat far-fetched—and then look to see if any patterns emerge that reflect a systemic vulnerability to those events which existing regulatory requirements have failed to account for.²²⁶

Although implementing these changes would require a considerable expansion of the banking agencies' quantitative stress tests, they would also allow for the current procedures to be streamlined in another respect. The qualitative stress tests which are administered through the Federal Reserve's CCAR program could be done away with altogether. For one, to the extent that the complex web of institutional interactions that give rise to systemic risk can be estimated with any reliability, quantitative models with transparent assumptions and objective, verifiable outputs are the only realistic mechanism for doing so. Armchair judgments—even those of astute, informed experts—do not represent a credible substitute. In addition, the particular regulatory requirements which the stress tests are best suited to examine are all expressed in terms of quantitative ratios. And lastly, on-site qualitative supervision of the Too Big to Fail banks already takes place pursuant to a

223. See *supra* notes 60–61 and accompanying text (describing the CLAR stress test).

224. The Dodd–Frank DFAST test's estimates of bank-specific asset losses due to fluctuations in employment, GDP, and inflation are relevant variables as well, but only one small piece of the puzzle. See Greenlaw et al., *supra* note 179, at 1–4.

225. For example, models have finally been developed that are able to predict the 2008 financial crisis when back-fitted to pre-2008 market data. See generally David Aikman et al., *Would Macroprudential Regulation Have Prevented the Last Crisis?*, 33 J. ECON. PERSP. 107 (2019) (answering the question posed in the title in the affirmative). Being able to “predict” the last crisis is of course a bare minimum requirement. The fact that models which meet that threshold could be considered a breakthrough highlights the scope of the challenge.

226. See generally Glasserman & Tangirala, *supra* note 108 (arguing for the need for more stress scenarios).

number of statutory provisions which are unrelated to Dodd-Frank's qualitative CCAR stress tests and, by comparison, is the more appropriate micro-prudential tool for identifying institution-specific organizational failures.²²⁷

2. A New Role for Banks

Banks would play a much different role under this proposal. Most importantly, the federal banking agencies would not make any determination that a particular financial institution has “passed” or “failed” the stress tests. The main point is not to avoid those particular labels. Rather, it is that the banks should not face any potential legal consequences from the results of the stress tests, whatever format they take.

This is not a deregulatory measure. If a bank is out of compliance with the standard suite of supervisory requirements, then the banking agencies remain free to impose dividend restrictions, lending caps, or any other remedy that is presently available pursuant to those rules.²²⁸ And the agencies may always exercise their authority to increase the stringency of those requirements as necessary. The lack of any possible sanction associated with participation in the stress tests is essential in this regard, because it removes any incentive for banks to strategically tailor their balance sheets in order to game the agencies' modeling assumptions.²²⁹ This in turn will mean that regulators obtain more accurate and reliable information about the risk profiles of the institutions they test.

In addition, there is no need for regulators to disclose the results for individual banks, regardless of whether any legal restrictions are triggered by their outcomes. Nor must the agencies generate a formal set of institution-specific results, even if they are held on a confidential basis as most supervisory examinations traditionally are. Any snapshot view of a particular bank's balance sheet is itself a hypothetical scenario, because the risk exposures it reflects at the time will inevitably evolve in the near future and no longer exist after the stress test has been performed. As such, the results for each bank are primarily useful as one data point for calculating a broader range of exposures that arise in the financial system going forward.²³⁰

The benefits of this approach are twofold. First, non-disclosure of bank-specific outcomes eliminates the possibility that negative results may trigger a run by markets on the financial institution at issue. Second, it makes the subset of information the banking agencies do disclose more reliable for

227. See Kress, *supra* note 197, 213–15 (detailing the role of on-site supervision at large banks).

228. See *supra* note 102 and accompanying text.

229. See *supra* Section III.B.1.iii.b.

230. See *supra* Section III.B.1.i.a. As argued, institution-specific disclosure carries few benefits as a form of market discipline, and also imposes costs by potentially triggering runs. See *supra* Sections III.A.2, III.B.1.ii.

investors, because the agencies have less reason to bias the stress tests outcomes in an attempt to manipulate market perceptions about a particular institution than is currently the case.²³¹

As a result of these reforms, the primary role of banks in the stress testing process would change in two ways. First, the banks would serve as the source of the agencies' modeling inputs, rather than as an institutional partner in designing and running the models. Dodd–Frank's mandate that banks undertake a parallel stress testing process in-house is ill-conceived, because it is especially burdensome on mid-sized banks and, more importantly, there is no market failure which prevents the largest banks from voluntarily establishing their own internal stress testing programs if it happens that that exercise provides a valuable risk management function.²³²

Second, banks would serve as funders of the remaining supervisory stress programs run by the agencies and be responsible for covering the public cost of administering the tests. This function is a critical part of the proposal. In order to design and validate the more sophisticated modeling techniques outlined above, the banking agencies must call upon substantially greater resources for research-and-development purposes than are now dedicated to the current stress testing programs.²³³ Those resources can and should be raised via a special tax on the banks which are subject to the stress testing, rather than allocated from the federal budget at a cost to U.S. taxpayers.²³⁴

Funding the stress testing programs with such a measure is potentially costless from a social welfare perspective, because it represents an economically efficient Pigouvian tax on the negative externality that is borne

231. See *supra* Section III.B.1.i.b.

232. See *supra* Section III.B.1.ii.

233. Cf. *supra* note 37 and accompanying text (noting the vast personnel resource initially dedicated to designing the 2009 SCAP test).

234. The basic idea of a sized-based bank tax has received thorough and favorable treatment in the finance literature. A number of legal mechanisms have also been incorporated in Basel III and the like, which contemplate surcharges of various forms being levied on the Too Big to Fail banks. See generally Olivier Jeanne & Anton Korinek, *Managing Credit Booms and Busts: A Pigouvian Taxation Approach* (Nat'l Bureau of Econ. Research, Working Paper No. 16377, 2010), available at <https://www.nber.org/papers/w16377.pdf> [<https://perma.cc/74Q3-UCXX>] (modeling the effects of a Pigouvian tax on borrowing); Thornton Matheson, *Taxing Financial Transactions: Issues and Evidence* (Int'l Monetary Fund, Working Paper No. 11/54, 2011), available at <https://www.imf.org/en/Publications/WP/Issues/2016/12/31/Taxing-Financial-Transactions-Issues-and-Evidence-24702> [<https://perma.cc/UG8K-YE92>] (reviewing theories and effectiveness of financial transaction taxes); Douglas A. Shackelford et al., *Taxation and the Financial Sector* (NYU Sch. of Law, Pub. Law Working Paper No. 10-30, 2010), available at <https://poseidon01.ssrn.com/delivery.php?ID=2511110651111011500508809507007308602704200504903003110912109010212301002300812111302504306005201002901501609709408600612008701101505703706712407201500112009807104608203706710911309502011120087115006124087126123097096089084002112005116090086099031&EXT=pdf> [<https://perma.cc/2UWQ-E49N>] (reviewing and evaluating efficacy of proposed corrective taxation schemes).

by society due to the systemic risks generated by large financial institutions.²³⁵ Tailoring a stress test excise based on the size of each particular bank is appropriate, as a way to more precisely align the Pigouvian properties of the tax at the margin and ensure that mid-sized banks are not disproportionately burdened.²³⁶

3. A New Role for Regulators

The federal banking agencies are responsible for two outputs under this proposal. The first is to design and administer the stress test process so that it produces a set of results which: (1) map the range of systemic risks that may potentially destabilize the financial sector; and (2) provide a measure of how well those systemic risks are addressed by existing regulatory requirements. The latter step need not cover every existing regulation. Three specific rules in Dodd–Frank are both necessary and sufficient: the risk-weighted capital ratios;²³⁷ the Counter-Cyclical Capital Buffer;²³⁸ and, a pair of liquidity requirements, known as the Liquidity Coverage Ratio and Net Stable Funding Ratio.²³⁹

The second output is a formal disclosure from the banking agencies which summarizes those results and provides guidance on how they will impact the administration of the three rules being tested going forward.

235. See generally A.C. PIGOU, *THE ECONOMICS OF WELFARE* (4th ed. 1932) (presenting the seminal framework for Pigouvian taxation of externality costs). For a canonical exposition of the advantages of Pigouvian taxation as a policy instrument, see Louis Kaplow & Steven Shavell, *On the Superiority of Corrective Taxes to Quantity Regulation*, 4 AM. L. & ECON. REV. 1, 1–14 (2002). See also generally Jonathan S. Masur & Eric A. Posner, *Toward a Pigouvian State*, 164 U. PA. L. REV. 93 (2015) (arguing for a Pigouvian tax).

236. Weighting a stress test tax in proportion to a financial institution's assets-under-management matches the private cost of compliance with the magnitude of systemic risk externalities, since smaller banks arguably pose less risk. See Amy G. Lorenc & Jeffery Y. Zhang, *The Differential Impact of Bank Size on Systemic Risk 2* (Fed. Reserve Bd., Finance and Economics Discussion Series No. 2018–066, 2018) (“[O]ur empirical results show that financial stress at large banks has a statistically significant and negative impact on the real economy. This impact increases with bank size.”).

237. The final agency rules which sets forth the currently applicable capital ratios was promulgated in 2013, and is known as Regulation Q. See Regulatory Capital Rules, 12 C.F.R. pts. 208, 217, 225 (2013).

238. See Countercyclical Capital Buffer, 12 C.F.R. pt. 217 app. A (2016).

239. The Liquidity Coverage Ratio (LCR), requires banks to hold a certain amount of “high-quality liquid assets,” a category meant to capture investments that can be sold at a reliable price on short notice. Liquidity Coverage Ratio, 78 Fed. Reg. 71,818, 71,819, 71,820 (proposed Nov. 29, 2013) (to be codified at 12 C.F.R. pt. 50). Specifically, the LCR is premised on the sale of assets during a hypothetical 30-day period in which a bank experiences disruption to its expected cash flows. Dodd–Frank also mandates regulatory rulemaking on a separate liquidity requirement, the Net Stable Funding Ratio (NSFR), which has yet to be finalized by the agencies. Net Stable Funding Ratio, 81 Fed. Reg. 35,124 (proposed June 1, 2016) (to be codified at 12 C.F.R. pt. 249). As envisioned in a notice of proposed rulemaking from 2016, the NSFR would focus on bank liabilities rather than assets, by requiring a threshold level of funding sources that are likely to remain available in the event of a market-wide credit freeze. *Id.*

Ideally, that guidance would be required to respond to some pre-stated criteria for whether the regulatory framework had “passed” or “failed” the test. A negative determination would trigger further guidance that either announces the agencies’ plan for increasing the stringency at least one of the three rules under review, or else articulates a specific rationale for why such an increase is unwarranted.

A positive determination should trigger critical evaluation of the regulatory requirements as well. Under Dodd–Frank, stress testing is a one-way ratchet: The only potential consequence is for existing minimum standards to be raised. But optimal financial regulation implies tradeoff between stability and efficiency. When restrictions on bank balance sheets are excessively onerous, financial institutions are unable to fulfill their fundamental role as intermediaries that bridge society’s savings with its investments, and the economy as a whole becomes less productive. Thus, if the banking agencies’ stress testing results indicate that regulatory requirements have been set too high, the accompanying guidance must identify how they can be safely decreased or explain why not. While crisis prevention was rightly considered the main policymaking imperative in response to the events in 2008, it is now a decade later, and the regulatory structure should incorporate flexibility to consider other long-term goals.

At a conceptual level, the changes outlined above reflect a radical inversion of the post-crisis stress testing process along a few basic dimensions. First, of course, is that the primary subject of the quantitative tests becomes the agencies and their rulemakings rather than bank balance sheets. Second, while the qualitative component of stress testing is discarded in the first instance, it reappears on the backend, with regulatory guidance in which the agencies must qualitatively justify their own quantitative results. Third, unlike the institution-specific remedies imposed under the current stress test procedures, the only legal response available to regulators would be to modify existing requirements on a wholesale basis for every bank to which they apply. Once stress testing is redesigned to serve a genuine macro-prudential function, particularized interventions at individual banks no longer make sense, because a premise of detecting systemic risk is that it exposes vulnerabilities which run across the banking sector as a whole. And fourth, to the extent it is possible to craft a set of non-discretionary criteria for determining positive or negative results, stress testing can function as a form of Prompt Corrective Action for the banking agencies, by mandating that regulators take action in response to a forward-looking assessment of the legal framework itself.

Despite these substantive departures from the Dodd–Frank status quo, the legal mechanics for adopting this Article’s proposed reforms are already largely in place. Pursuant to the CCAR program, the Federal Reserve already issues a lengthy policy statement on an annual basis, which summarizes and

interprets the stress tests results for each year.²⁴⁰ The banking agencies also enjoy authority to modify the three rules in question over time. The Counter-Cyclical Capital Buffer has been designed so that it is adjustable by regulators on an ad hoc basis.²⁴¹ Changes to the minimum capital and liquidity ratios could be implemented through the notice-and-comment process, or by the issuance of more informal interpretative rules whenever formal rulemaking is not required as a matter of administrative procedure.²⁴²

B. ADVANTAGES OF THE PROPOSAL

The previous Section of this Article argued that the leading criticisms of post-crisis stress testing are consistently misplaced. In some cases, those critiques fit awkwardly with the broader legal structure of financial regulation or the underlying policy challenges those rules encounter. In other respects, they reflect legitimate concerns, but ones that are unlikely to be resolved by the recent raft of reforms which have begun to sweep through the Trump Administration and Congress. The discussion below explains why the alternative reforms proposed in this Article reflect a more promising response to the same underlying critiques. It then briefly summarizes how those reforms would help secure the benefits that Dodd–Frank’s rules were originally intended to deliver.

1. Addressing Critiques of Stress Testing

While the methodological changes suggested above all place a premium on the use of quantitative risk-management techniques, they also alleviate many of the concerns raised by commentators who have cautioned for greater quantitative skepticism in connection with financial stress testing.²⁴³ For instance, substantially expanding the number of downturn scenarios the banking agencies must model is consistent with claims that the post-crisis stress tests reflect an overly ambitious attempt at “regulation by hypothetical.”²⁴⁴ Testing a wider array of economic scenarios reflects the reality that regulators will never perfectly anticipate how financial markets may evolve, and that no particular scenario can provide a complete picture of the different forms of systemic risk that might appear over time. By contrast, the Dodd–Frank tests stipulate a single “severely adverse” recession and assume that it will capture the full universe of downside risks in the financial system. That modeling decision involves considerably more trust in technocratic foresight and epitomizes the epistemological hubris which quantitative skeptics have in mind.

240. See *supra* Section II.D.

241. See 12 C.F.R. pt. 217 app. A.

242. See TREAS. BANK’G REP., *supra* note 5, at 140–47.

243. See *supra* Section III.A.1.

244. See Baradaran, *supra* note 8, at 1282.

For related reasons, this Article's proposal is also consistent with commentary that calls for a more "deliberative" approach to stress testing than is currently found in the post-crisis rules.²⁴⁵ The use of multiple scenarios allows regulators the opportunity to weigh the strengths and weakness of any given modeling approach. Requiring the banking agencies to articulate a connection between their stress test results and rulemaking outputs also introduces greater deliberation into the regulatory process. Under the current Dodd–Frank procedures, the banking agencies take outputs from the stress testing process at face value and mechanically translate them into new restrictions on bank balance sheets.

This Section's proposal also goes a long way toward addressing critiques of the Dodd–Frank stress tests relating to regulatory discretion, transparency, and uncertainty.²⁴⁶ Under the suggested framework, stress testing cannot devolve into a dysfunctional "Kabuki dance" by regulators. The banking agencies only have discretion to modify a generally applicable set of objective quantitative benchmarks. Banks will therefore never be subject, as they are now, to unexpected legal penalties that are improvised ad hoc by regulators on obscure qualitative grounds. Moreover, because banks cannot fail the stress tests, they have nothing to gain from anticipating the agencies' modeling methodologies in advance or attempting to manipulate their results. This means an end to the spurious battles presently waged over the transparency of those methodologies, which are best kept entirely opaque to the banks being tested.

Lastly, the proposal is also sensitive to the compliance cost arguments raised by Dodd–Frank's critics.²⁴⁷ The most compelling of those arguments is that the high-fixed cost of complying with the current procedures makes them disproportionately burdensome for mid-sized banks. By concentrating responsibility for designing and administering the stress testing process with the federal banking agencies, this Article's proposal eliminates the main source of those fixed costs: the DFAST program's mandate that each bank must develop its own customized stress test and run it in-house. With that requirement removed, the only fixed cost involves the relatively manageable task of compiling the internal balance sheet data requested by the banking agencies. Accordingly, the overall public and private cost of stress testing becomes almost completely variable based on institutional size: JPMorgan, Inc., with roughly \$2 trillion in assets, would bear 40 times the cost that a \$50 billion bank does.²⁴⁸ Staggering the cost of administering the stress tests would also have one other positive side effect: It would enable the number of banks participating in the process to expand, rather than contract as is happening

245. See Weber, *A Theory for Deliberation-Oriented Stress Testing Regulation*, *supra* note 8, at 2239.

246. See *supra* Section III.A.2.

247. See *supra* Section III.A.3.

248. See *supra* Section III.A.3.

with recent reforms, and thereby improves the potential for stress testing to provide a more complete macro-prudential view of systemic risk.²⁴⁹

2. Making Good on the Purported Benefits of Stress Testing

This Article's proposal would make progress at securing the informational benefits of stress testing, which are often attributed to the Dodd–Frank programs in various ways but have not yet been realized. The two essential steps for reaching that goal are to shore up methodological deficiencies in the current tests so that they can potentially produce meaningful results, and to eliminate (or at least reduce) the incentives which banks and regulators presently have to bias those results. By addressing both problems, the reforms outlined above make it possible for stress testing to facilitate macro surveillance of the financial system by the federal banking agencies.²⁵⁰ In addition, a requirement that the agencies release forward-looking guidance on how the stress test results will inform their rulemaking activities gives investors access to the category information they value most of all—the otherwise undisclosed assumptions which *regulators* are making.²⁵¹

Once the stress testing process produces reliable information, it also opens up an opportunity for the federal banking agencies to genuinely engage in regulatory learning.²⁵² Most directly, this would take place through the decision to modify the relevant minimum capital and liquidity ratios up or down over time based on what has been learned. Stress testing could also reveal better ways to structure the ratios, themselves. For example, Dodd–Frank's two liquidity requirements—the Liquidity Coverage Ratio and Net Stable Funding Ratio—are entirely novel regulatory instruments which did not exist prior to the crisis. As such, there is great uncertainty as to how they operate, as well as which of the pair might be more effective.²⁵³ Accordingly, regulators could gain clarity on how to proceed by designing a subset of the

249. The least compelling compliance cost argument is that mid-sized banks should bear no cost, and therefore be exempted from stress testing process altogether. For reasons explained above, however, that would be a mistake. *See supra* Section III.A.3.

250. *See supra* Section III.B.1.i. The same reforms also make it more likely that financial market investors will find that stress testing results disclose valuable information about economic fundamentals of the banking system. *See* Section III.B.1.i.

251. *See* Section III.B.1.i.

252. *See supra* Section III.B.1.ii.

253. In fact, as a consequence of that uncertainty, formal rulemaking on the Net Stable Funding Ratio ("NSFR") is still pending, nine years after Dodd–Frank was passed, and the banking agencies have indicated that they have suspended any plans to finalize the rule. There is some evidence from the finance literature, however, that the NSFR is the superior benchmark of the two. *See* Antoine Lallour & Hitoshi Mio, *Do We Need a Stable Funding Ratio? Banks' Funding in the Global Financial Crisis* 21–22 (Bank of Eng., Working Paper No. 602, 2016), available at <https://www.bankofengland.co.uk/-/media/boe/files/working-paper/2016/do-we-need-a-stable-funding-ratio-banks-funding-in-the-global-crisis.pdf?la=en&hash=925455009706E7D907FCF132D7F152AD2654CFC2> [<https://perma.cc/HU> U8-ZZZF].

stress scenarios to explicitly compare the relative performance of those rules by simulating the kinds of financial disruptions they are thought to prevent or contain.

Lastly, the reforms suggested by this Article would mean the stress tests actually improve regulators' ability to safeguard the banking sector with capital requirements and related balance sheet restrictions. After the methodological flaws in the Dodd–Frank tests are remedied, the banking agencies will be able to tailor capital requirement so that they track systemic risks in a forward-looking manner for the first time.²⁵⁴ And once the stress test results are tied to adjustments in the Counter-Cyclical Capital Buffer, the as-now unfulfilled potential of that rule could also be achieved.²⁵⁵ In the process, stress testing itself would be transformed into a legitimate regulatory complement to other post-crisis rules, by functioning as a calibration device that refines their content, rather than serving as an awkward substitute mechanism that simply reinforces their present shortcomings.²⁵⁶

V. CONCLUSION

This Article has presented a comprehensive analysis of the promise and limits of financial stress tests. That analysis reveals that both Dodd–Frank's architects as well as its reformist skeptics have misconceived the vices and virtues of the post-crisis stress testing rules. As has been argued, the current procedures bear surprisingly little relation to the systemic risks they were designed to address. At the same time, claims that those rules represent a harmful escalation of regulatory burdens, discretion or uncertainty are overstated. The Article shows a way out of this intellectual impasse, and charts a better path forward which can be achieved with a simple yet fundamental twist to the administration of stress tests. As it explains, what is needed is a set of reforms which transform stress tests into tools for diagnosing weaknesses in the regulatory requirements promulgated by federal banking agencies, rather than in the banks themselves. By stress testing the banking agencies for regulatory failure, the kinds of market failures which may lead to financial crisis are more likely to be prevented.

254. See *supra* Section III.B.2.

255. See *supra* Section III.B.2.

256. Turk, *supra* note 211 (manuscript at 25–26); see *supra* Section III.B.3.