Capital in the Financial Crisis

Timothy F. Geithner
Andrew Metrick
Chase Ross
Capital in the Financial Crisis

Timothy F. Geithner, Andrew Metrick, Chase P. Ross

July 30, 2020

Abstract

Why was the financial system stabilized with so little additional capital and very little additional public capital by mid-2009 despite substantial loss estimates? In early 2009, mark-to-market losses implied $800 billion of credit losses remaining, and consensus estimates stood at $340 billion. But, the stress test in May 2009 required banks to only raise an incremental $75 billion of capital. Why? We discuss three reasons. First, credit losses are not equivalent to capital needs because banks that remain a going concern can offset losses with revenues. Second, net income and actuarial loss estimates—rather than market-implied losses—were the proper basis for the stress test estimates of capital needs. Finally, the expected path of the economy improved with the macroeconomic policy actions. Covid-19 presents the first significant test of the reforms to bank capital and the structure of the U.S. financial system. Policymakers in some countries will confront similar choices and trade-offs to restore confidence in the financial system’s capital position as they face the second systemic financial crisis in little over a decade. They will undoubtedly see parallels with the 2009 experience.

JEL Codes: G01, G20, G21, G28
Keywords: stress tests, bank capital, bank regulation, financial crises

Introduction

In early 2009, with consensus estimates of credit losses remaining for the U.S. banking system at $340 billion and mark-to-market losses at $800 billion, the American banking system stood on the brink of a precipice. Compared to $340 billion of Tier 1 common capital—the highest quality loss-absorbing capital—implied losses in the banking system were sufficiently large to leave the system nearly insolvent. Despite these loss estimates, the May 2009 stress test required banks to raise only $75 billion of capital.

Debates about lessons learned from the financial crisis rightly focus on the role capital played as a cause and as a part of its resolution. While research has tried to understand the broader questions of why the stress test worked and to what extent, this paper focuses immediately upstream: why did the financial system stabilize using capital amounting to a seemingly tiny fraction of the credible loss estimates?

1 We are grateful for thoughtful discussions and comments from Lew Alexander, Tim Clark, Gary Gorton, Beverly Hirtle, Matt Kabaker, Meg McConnell, Sharon Ross, Lee Sachs, Til Schuermann, Kevin Stiroh and participants at Yale’s finance seminar.
2 President of Warburg Pincus and U.S. Treasury Secretary from 2009 to 2013.
3 Yale School of Management and NBER. Email: andrew.metrick@yale.edu
4 Yale School of Management. Email: chase.ross@yale.edu
We argue that mark-to-market loss estimates in early 2009 were the wrong yardstick to measure capital needs. As long as the system stayed a going concern, the financial system would earn income to offset losses. A myriad of fiscal, monetary, and financial policies, along with an improving macro backdrop, began taking hold in the first half of 2009 as the Fed conducted the stress test. The likelihood of the most extreme tail outcomes fell, resulting in the stress test’s small capital requirements compared to loss estimates.

Understanding how hundreds of billions of losses translated to just $75 billion of incremental capital is only part of the story. The success—or lack thereof—of the stress test depended on whether investors believed that $75 billion of capital was sufficient to stabilize the banking system. We argue markets believed the capital needs estimate because the test was credible.

We start by briefly examining the capital dimension of the build-up to the crisis, although we refer the reader to Bernanke, Geithner, and Paulson (2020), Chapter 1, for a detailed account. The capital regime applied to banks was poorly designed, though more conservative than what applied to government-sponsored enterprises (GSEs) and investment banks. The architects of the capital regime used the relatively benign Great Moderation period to set regulatory capital levels. These shortcomings were not evident because actual losses in recent recessions had been relatively modest. Permanent, loss-absorbing capital formed too small a share of regulatory capital and risk weights were too low. As credit losses mount across the financial system, regulatory capital ratios painted a false picture of strength, undermining investor confidence in the regulatory capital regime, as shown in Figure 1.

The magnitude of the post-crisis changes to the regulatory risk weights illustrates the poor design of the capital regime pre-crisis. Although it is impossible to backfill Basel III definitions to pre-crisis capital ratios, we approximate the difference between crisis-era capital ratios and Basel III capital ratios using filings from 2013, when banks reported both Basel I Tier 1 Common and Basel III Common Equity Tier 1 ratios. Table 1 shows the reported Basel I Tier 1 Common ratios as well as our estimated Basel III capital ratios for major banks in 2007. Had regulators applied the same definition to capital and risk-weights in 2007 as they do under Basel III, banks’ capital ratios would have been 200 basis points (bps) lower for the same mix of assets, assuming a constant ratio between the Basel I and Basel III capital ratios.

Ex-post outcomes also reveal the inadequacy of the pre-crisis capital regime. The amount of high-quality capital the largest banks lost to credit losses shows how much capital they should have held ex-ante. For the 25 largest banks, Strah, Hynes, and Shaffer (2013) find that losses ate away an average of 322bps of Tier 1 Common equity, as shown in Figure 2. Ten of the largest banks’ Tier 1 Common eroded by more than 400bps, with Washington Mutual losing more than 12% of Tier 1 Common. In 2007, the average large bank had about 5% CET1. An average loss of high-quality capital greater than 3% is a remarkable share of the highest-quality capital banks had going into the crisis.

As the crisis intensified, the higher levels of capital and more stable funding of the banking system provided a measure of protection compared to nonbanks. In the end, the regulatory capital measures did not offer a particularly valuable picture of the banking system’s solvency.

The literature on stress testing and capital in the financial crisis is growing. Bernanke, Geithner, and Paulson (2020) provide a comprehensive, first-hand description of the many components of the policy response to the crisis. Most related to our paper, Jester, Nason, and Norton (2020) discuss the recapitalization efforts in 2008, and Clark, Kabaker, and Sachs (2020) discuss the capital dimensions in 2009, including the stress test and the related capital backstop. Additionally, Goldstein (2017) presents a thorough overview of bank capital and
stress testing from 2009 to 2015, along with comparisons of the many stress tests conducted around the world since the crisis.

One thread of the literature conducts event studies of stress tests to gauge their effectiveness and market perception of the results. The broad consensus is that the 2009 U.S. stress test was credible and successful, compared to early European stress tests. Flannery, Hirtle, and Kovner (2017) show that Federal Reserve stress tests produce information for stressed banks, and they provide a review of the even study literature with methodological suggestions. Comparing the market reactions of the U.S. and European stress tests, Candelon and Sy (2015) find that the 2009 U.S. stress test had a positive impact on bank returns. Fernandes, Igan, and Pinheiro (2020) analyze stress tests in the U.S. and find they produce new information, but the tests do not eliminate incentives to produce private information. Bayazitova and Shivdasani (2011) show a significant effect on banks’ valuations via a certification effect. By looking at stress test events, including announcements, methodology updates, and outcomes, Sahin, Haan, and Neretina (2020) show that stress tests have asset pricing implications, affect credit and equity markets, and change banks’ betas, suggesting that stress tests have a significant effect on systemic risk. A central goal of the 2009 test was to stabilize credit provision, and Cohen and Scatigna (2016) provides evidence that, post-crisis, banks with larger capital buffers were both more profitable and able to expand lending.

Another thread of literature offers lessons from stress testing experience and provides best practices for future tests. Hirtle, Schuermann, and Stiroh (2009) highlight the use of the capital backstop as a useful tool to ensure each bank would ultimately raise the necessary capital. Greenlaw et al. (2012) provide a framework for analyzing stress tests and emphasize a macroprudential focus, rather than microprudential focus on prompt corrective action and resolution. Tarullo (2010) discusses the lessons learned and focuses on the transparency of the test and how ongoing supervisory plans should approach transparency. Schuermann ((2014), (2016)) explains the post-crisis popularity of stress testing and the difference between stress tests conducted during a crisis and stress tests conducted during normal times. Bookstaber et al. (2013) discuss how stress tests should account for the interconnections of the financial system. Feldberg and Metrick (2019) provide best practices for supervisors conducting stress tests in peacetime, with the goal of uncovering growing risks in both the traditional and shadow banking system. Goldstein and Sapra (2013) ask when stress tests should disclose results and weigh the costs and benefits of disclosures, focusing on what makes the analysis and regulator credible. They suggest that the 2009 stress test was particularly successful not due to the information in the disclosures, but because the Fed pre-announced what would happen to failing firms.

This paper is organized as follows: section 2 discusses mark-to-market losses implied from subprime MBS market prices and shows how bank prices covaried strongly with these loss measures. Section 3 describes the US stress test. In section 4 we tackle the “stress test mystery”—the seeming disconnect between expected losses of $600 billion and the subsequent $75 billion in incremental capital the stress test required banks to raise. Section 5 concludes.

**Estimating Losses**

The market price of reference assets, such as subprime mortgage-backed securities (MBS), shaped public perceptions of bank solvency during the crisis. The collapse in funding and subsequent fire-sales distorted asset-backed security prices, and uncertainty about the size and distribution of potential losses across the
financial system exacerbated the panic. As the prices of MBS fell through 2008 and into early 2009, implied losses to the banking system grew and fears of insolvency increased.

Contemporaneous estimates of capital needs reflected rough estimates of gross losses implied by mark-to-market ABX reference tranches. Figure 3 shows that implied losses on all subprime MBS issued between 2005 and 2007 peaked at $840 billion in 2009Q1. Tangible common equity across all U.S. commercial banks amounted to roughly $700 billion at the end of 2006. If commercial banks held all subprime MBS, marking-to-market would have wiped out more than the entirety of high-quality loss-absorbing capital. Greenlaw et al. (2008) estimate the financial system’s total capital—comprised of banks, broker-dealers, GSEs, hedge funds, and others—to be $1.9 trillion; in this case, subprime MBS mark-to-market losses would have wiped away nearly half of all capital in the financial system.

We illustrate the importance of subprime securities prices for bank valuations by calculating banks’ betas to subprime returns after controlling for the aggregate market. Figure 4 shows the 60-day rolling beta estimates—a measure of covariance—of banks’ CDS returns relative to ABX.HE.AAA.2006.1, an index that reflects the performance of AAA subprime securities issued in the second half of 2005. We create an index of short bank CDS returns (i.e., selling protection) with the value-weighted average of 5 year CDS contracts for JPM, C, WFC, BAC, MS, and GS. We value-weight using 2006 year-end market capitalization and translate CDS spreads to returns.

In early 2007, the covariance of banks’ CDS returns with the reference ABX tranche is unsurprisingly statistically insignificant from 0 and small. But, beginning in the second half of 2007, banks’ CDS returns become linked to the subprime index: they become so intertwined that this single tranche can help explain returns in bank stocks and CDS on a day-to-day basis. At peak covariance in October 2008, a 1% decline in the ABX was associated with a 0.75% decline in banks’ CDS returns on the same day holding the aggregate market return constant. Policymakers’ actions to quell the crisis centered around breaking the relationship between bank valuations and subprime mortgage prices.

Institution-specific equity return betas to subprime prices reveal substantial heterogeneity across the biggest banks and broker-dealers. Figure 5 shows the beta estimates from regressing daily firm returns on the return of the ABX.HE.AAA.2006.1 and the market return during the crisis period. During the crisis bank stocks betas to the ABX, after controlling for the aggregate market, were large, especially for the companies that ultimately ceased operations or were acquired. For example, a 1% fall in the ABX index corresponded to a fall of 0.7% in the stock price of Merrill Lynch, 0.6% for Citi and Lehman Brothers holding the aggregate market return constant. Betas for relatively better-capitalized banks like JP Morgan were smaller, below 0.1.

Breaking the relationship between bank valuations and ABS prices started with three stages of recapitalization. The first stage of recapitalization was the GSEs’ conservatorship and the Capital Purchase Plan (CPP), the second stage of recapitalization includes the series of subsequent capital injections for AIG, Citi, Bank of America, and the GSEs, and the last stage included the May 2009 stress tests and its accompanying capital.

---

5 Of course, commercial banks did not hold all subprime MBS. Greenlaw et al. (2008) estimate U.S. commercial banks held 18% of subprime MBS outstanding, behind only insurance companies at 23%, and ahead of estimated U.S. hedge fund holdings of 17%.

6 Specifically, we translate CDS spreads to returns using $CDS_t^{ret} = -(CDS_{t-1}/250 + \Delta CDS_t \times RVPPV01_{t-1})$ where $RVPPV01$ is the risky present value of 1 basis point calculated using a linearly interpolated Libor swap curve. $CDS_t$ is the CDS spread.
backstop. We summarize the total capital injections and their forms for each period in Figure 6. For a detailed discussion of the stages of recapitalization, refer to Jester et al. (2020), Jester, Nason, and Norton (2020), and Clark, Kabaker, and Sachs (2020) who discuss the GSEs, the CPP, and the early 2009 recapitalizations, respectively.

The Stress Test

Despite a series of capital injections, guarantees, and loss-sharing agreements, the real economy was in bad shape in January 2009. Market perception held that the banking system remained insolvent. Real GDP had contracted at an annualized 4% rate in 2008Q4, and subsequent GDP revisions would show the actual number exceeded 8%. Banks’ CDS spreads remained worryingly high. Initial claims continued to climb, more than doubling the pre-crisis average in January 2009, and the January payroll printed a 700,000 net decline in nonfarm payrolls. Estimated credit losses remaining were very high relative to capital in the system. In this context, the Federal Reserve designed and implemented the test, announcing it in February and releasing the results in early May.

The CPP and the previous series of capital injections had not lifted the cloud over the financial system. Investors were skeptical of the regulatory capital regime and couldn’t discern healthy from unhealthy banks with public information. Investors struggled to assess the risk of future policy actions that might be dilutive to existing equity holders. Stress testing, with a standing capital backstop, provided policymakers the ability to remove the worst-case tail from the distribution of outcomes.

The stress test estimated losses of $600 billion across the 19 banks tested in the “more adverse” scenario, implying $950 billion in losses from mid-2007 to the end of 2010. Despite the significant loss estimates, the incremental capital needs of the financial system that the stress test identified were modest: only $75 billion spread across ten banks. We reproduce the results for the aggregate system in Table [tab:scap-results]. Other than GMAC, the firms required to raise additional capital were able to do so privately. Markets received the test results well: on average, bank CDS spreads fell by a third in the days following the announcement.

The stress test helped markets distinguish between healthy and less healthy banks, which is apparent by the test’s estimated dispersion in losses and capital actions across banks. As a share of risk-weighted assets, losses for the tested banks ranged from 4% to 11.8%, post-stress Tier 1 Common ratios ranged from 2.3% to 15.5%, and estimated losses ranged from 4.0% to 11.8%. There was a corresponding difference in capital raised by banks from $0 to $33 billion.

In part due to the reception of the stress test, the tested banks were able to raise $90 billion of Tier 1 common, and $125 billion across Tier 1 Common, preferred, and other Tier 1. Combined, the SCAP banks privately raised $314 billion and publicly raised $227 between 2007Q1 and 2010Q4. This cumulative $541 billion of capital, along with the $664 billion in capital levels at the beginning of 2007, absorbed total realized losses to U.S. banks and brokers of $806 billion, considerably better than the SCAP’s estimated $950 billion in losses in the “more adverse” scenario for the same time frame.

---

7 GMAC received capital from TARP via the Automotive Industry Financing Program.
8 The SCAP’s estimated losses include losses to only the 19 tested banks, whereas the $806 billion in losses from 2007 to 2010 contains losses to all banks and brokers.
The Stress Test Mystery

What explains the mystery of why the financial system was stabilized with so little additional capital and very little additional public capital by mid-2009 despite substantial loss estimates? This question is of first-order importance. Among the 19 banks tested in the stress test, total losses in 2009 and 2010 in the more adverse scenario totaled $600 billion, but the incremental required buffer was less than $75 billion. Why?

We address the mystery in three parts: the mechanical part, the conceptual part, and the macro part.

The Mechanical Part

Gross losses do not correspond one-for-one to capital needs. Banks offset losses with profits over time, and banks’ pre-tax income does not correspond one-for-one to regulatory capital. Forecasting bank profits are equally as crucial as forecasting losses to estimate capital needs. Directly comparing losses implied via mark-to-market losses on subprime ABS to capital levels, for instance, paints a bleak picture so long as banks remain a going concern and continue making profits. Estimates of capital needs can evolve as the forecaster changes expectations about either estimated losses or pre-tax revenues. Forecast updates help explain part of the difference between the Fed’s January 2009 capital shortfall estimate of $684 billion and the stress test’s $185 billion in May 2009: banks’ 2009Q1 pre-provision net revenue (PPNR) was more durable than expected, giving them more resources to absorb losses.

We find that both revenues and earnings per share (EPS) were consistently underestimated by analysts in the first half of 2009, as shown in the regression results presented in Table 3. The table shows that banks consistently outperformed expectations in the first two quarters of 2009. As markets calculated capital needs numbers, revenue and earnings assumptions were likely too pessimistic and could consequently lead to an overstatement of capital needs.

We analyze earnings and revenues surprises by comparing consensus forecasts for individual companies against the actual result for publicly traded U.S. companies with analyst coverage using I/B/E/S data. Table 3 presents results of regressing the surprise in earnings or revenues at the firm-quarter level on a indicator variable—equal to 1 if the firm was included in CPP, SCAP, a bank, or a financial company—and controls for dispersion in analyst forecasts, four lags of quarterly earnings surprise, and quarterly fixed-effects. Surprise is defined as \( \frac{(Actual - Consensus)}{Consensus} \times 100 \), and is winsorized at the 1 and 99% levels to reduce the influence of outliers.

The regression setup has the advantage that it allows us to compare analyst forecast accuracy of banks relative to all other companies. If, perhaps, forecasts were particularly inaccurate for both banks and nonbanks, then we would expect to see a statistically significant constant and a coefficient of zero on our financial category indicator variables. However, the table makes clear that on average forecasts were accurate for revenues across all non-financial companies because the constants in the first four columns are small and almost all are not different from zero. In the last four columns, we can see that on average analysts were too pessimistic on EPS (a positive constant means the average surprise was positive). The coefficient on the indicator variable is the average surprise for financial companies holding constant analysts’ forecast accuracy for all non-financial companies. The first four columns of the table show revenue surprises were 20 percentage points (pp) higher than the average consensus surprise for CPP firms, 16pp for SCAP firms, 2pp for all banks, and 3pp for financial companies. The last four columns show a similar measure for EPS: 119pp for CPP firms, 64pp for SCAP firms, 5pp for banks, and 9pp for financial companies.
Let’s consider a capital regime designed to allow capital to cover gross losses with no credit for earnings. In this view, the estimated losses for CPP banks was about $475 billion in 2009 and 2010. Excluding roughly $300 billion in PPNR, the losses would wipe out about 75% of 2008Q4 capital. Table 4 presents the capital impact of using gross instead of net losses for the largest banks included in the CPP assuming that losses translate one-for-one to capital erosion because public information does not allow us to make more realistic calculations. Using gross losses rather than net losses would have lowered stress capital ratios between 200 and 600 basis points.

The distribution of net income across firms and deferred tax asset related deductions drove a wedge between pre-tax income and incremental contributions to regulatory capital. The stress test itself only approximated the translation of pre-tax income to capital but improved on the practice of using pre-tax income as a proxy for contribution to regulatory capital.

Comparing estimates of capital needs also requires an apples-to-apples comparison. Some estimates—like the New York Fed’s January estimate of a capital shortfall totaling $684 billion in the worst-case scenario—examined the aggregate banking system. The stress test focused on 19 firms, each with more than $100 billion in assets.

Moreover, the stress test did not use mark-to-market implied losses to calculate indicative loss rates and marked a conceptual divergence from comparing mark-to-market losses in distressed assets, typically sub-prime MBS, to banks’ capital. Recall Figure 3: the entire amount of tangible common in U.S. commercial banks at the beginning of 2007 was $712 billion, and mark-to-market implied losses in AAA to BB-rated subprime MBS peaked above $800 billion in March 2009—the banking system, some would argue, was insolvent.

The difference between market-implied losses and actuarial losses—as calculated with credit fundamentals—is the liquidity premium (Gorton 2008). It is no mystery that implied loss rates picked up as ABS market liquidity disappeared. Not until after the unveiling of programs explicitly aimed at restoring the ABS markets—TALF and PPIP, discussed in detail below—did ABS market liquidity return to a reasonable level and market-implied loss rates start falling.

Actuarial estimates of losses during the crises were consistently lower than mark-to-market estimates. In April 2008, for example, Bank of England (2008) calculated actuarial losses on all subprime MBS at $193 billion (in the worst-case scenario) and mark-to-market implied losses of $380 billion. Realized losses calculated by Ospina and Uhlig (2018) in all subprime MBS was $119 billion, amounting to a 10% loss rate. The stress test imposed an aggregate loss rate in the more adverse scenario for first-lien mortgages of 8.8%, and 13.8% for second or junior lien mortgages, while market-implied loss rates on all subprime MBS exceeded 50% in the first quarter of 2009. The realized loss rate across prime, Alt-A, and subprime mortgages was 8% (Ospina and Uhlig 2018). By reasonable measures, market-implied losses vastly overstated realized losses, and actuarial estimates—either using standard industry practice or via the Fed’s stress test

10 Realized loss estimates are as of December 2013, as a percent of the principal amount at issuance and includes bonds issued up to 2008.
methodology—produced loss estimates in the right ballpark. Estimates for capital needs based on marked-implied losses grossly overstated capital requirements.

Reflecting both the explanations above, the SCAP found the cumulative effect of PPNR, accounting adjustments related to mergers, and capital actions in Q1 2009 as $540 billion. The SCAP required only $75 billion incremental capital raising partly because the roughly half a trillion dollars in profits, plus the approximately $400 billion in Tier 1 Common at year-end 2008, absorbed the majority of losses suffered over the two years.

The Conceptual Part

So far we have discussed mechanical calculations: how regulators arrived at an estimated loss number, and why that number starkly differed from existing estimates. The stress test worked ex-post, evidenced by CDS spreads for large banks. Why did markets believe the $75 billion incremental capital was a sufficiently large buffer? The stress test was effective because it was both credible and coincided with improving macro results.

We summarize our view of the most critical features of the stress test which bolstered the test’s credibility: the capital backstop, the high loss rates and loss estimates, and the test’s transparency. In addition to the stress test, policies that attenuated the worst-case macroeconomic outcomes helped the credibility of the stress test by implicitly reducing capital needs. These policies included a broad mix of monetary and fiscal policies, including hardening bank liability guarantees (e.g., the FDIC’s Debt Guarantee Program), market funding backstops (e.g., PDCF, TAF, and the TSLF) and lending programs (e.g., TALF and PPIP).

Capital Backstop

The test calculated that if banks with inadequate capital were unable to raise sufficient capital privately, they could raise capital from the public backstop, the Capital Assistance Program (CAP). The risk that disclosures revealed a weaker bank or banking system than expected required an up-front solution which the capital backstop provided. Announced on February 10, 2009—nearly three months before the stress test results—the CAP allowed banks to convert their preferred shares to common at 90% of the average closing price of the stock in the 20 trading days before February 9. The provision of a public capital backstop with a price floor provided a credible signal that policymakers would recapitalize the system to a level sufficient to withstand Great Depression-like losses, even if banks were not able to raise the additional capital from private investors.

The CAP helped align regulators’ incentives as they conducted the stress test, knowing that banks ultimately had recourse to the CAP. However, regulators were aware of the risk that the remaining TARP capital could prove insufficient. In February 2009, the Administration’s budget proposal included an additional $250 billion request for incremental funding to support $750 billion in purchases in case conditions in the banking system eroded further and the stress test revealed a larger capital hole than expected.

11 The stress test also included implementation details that, while well-designed, were not primarily aimed at supporting the credibility of the test, and we will not discuss them in detail. These implementation details include a focus on levels of capital rather than ratios, hurdle rates, and treatment of pre-provision net revenues.

12 The budget proposal notes: “...the President’s Budget includes a $250 billion contingent reserve for further efforts to stabilize the financial system. (The reserve, which reflects a net cost to the Government, would support $750 billion in asset purchases.) The existence of this reserve in the Budget does not represent a specific request, rather as events warrant, the Administration will work with the Congress to determine the appropriate size and shape of such efforts, and as
Banks had six months to raise private capital, and the backstop provided funds in the form of convertible preferred securities if they could not raise enough capital privately. CAP preferred could convert to common equity if necessary, and conversion was mandatory after seven years. The Citi conversion in February 2009 provided credibility that the conversion of public capital to loss-absorbing equity was possible. Ultimately, no banks turned to the CAP, and the program concluded unused.

**High Loss Rates**

The stress test imposed particularly severe loss estimates relative to both consensus and historical experience. The adverse scenario’s two-year commercial loan loss rate peaked at 9.1%, higher than any loss rate since the 1920s. The adverse scenario’s unemployment peaked at 10.3%, implying the worst labor market in the post-war period, with one exception in late 1982. Comparing the adverse scenario outcomes with measures of consensus forecasts for GDP and unemployment illustrates the severity of the stress test, as shown in Table 2.

**Large Loss Estimates and Capital Needs**

The stress test estimated gross credit losses of $600 billion, exceeding consensus estimates of $336 billion and the Federal Reserve’s January 2009 estimate of $400 billion. Figure 7 shows the stress test’s estimated losses compared with other estimates.

After the Fed released the stress test results, Bridgewater Associates (2009)—titled “We Agree!”—said:

> The Stress Test numbers and ours are nearly the same!!!…The regulators did an excellent job of explaining exactly what they did for this stress test, and showing the numbers that produced the results. They did virtually exactly what we did since putting out our loss estimates nearly two years ago, and their numbers are essentially the same as ours…

The stress test’s two-year loss rates were somewhat higher than Bridgewater’s, while the net capital needs were nearly the same. Other market commentary found similar results. For example, Morgan Stanley’s bank analysts titled their reaction note “Stress Test Framework Shows It is A Tough Test.” (Graseck, Pate, and Kwong 2009).

**Transparency**

The test was transparent in two different ways: its design and its results. Before announcing the results, regulators were open about the design of the test, including indicative loss rates used as assumptions in the capital needs calculations (see Federal Reserve (2009b), (2009a)). The results, released on May 7, provided bank and business-segment specific details. The focus on transparency aimed to lend credibility to the test and allow markets to make direct comparisons across banks. Such a high level of transparency was necessary to dispel extreme concerns.

The information provided in the stress test results allowed private investors to gauge the credibility of the government’s loss estimates for themselves. The concern for private investors was whether they would have their capital wiped out either because the environment deteriorates and the bank is forced to raise additional

more information becomes available the Administration will define an estimate of potential costs.” (Office of Management and Budget 2009).
equity at a lower valuation, or that the government nationalizes the bank and equityholders are wiped out.

For these reasons, transparency at both the aggregate level and the institution-specific level was critical: markets needed to know the incidence of losses across specific firms to make capital injections.

The unusual amount of transparency was not without criticism, which largely fell into three groups. First, the Fed’s release of bank-specific results broke from the standard level of confidentiality involved in bank supervision. Second, there was fear that the upfront commitment to publicly disclosing the results encouraged supervisors to white-wash the results to protect weak banks. Immediately after the announcement of the stress test several media articles cited this concern. Andrews and Dash (2009) noted:

> Analysts say the administration’s worst projections, which it describes as unlikely, are not much more dire than what many private forecasters already expected…‘It sure sounds to me like they are designing this to make it sound like the banking system is in great shape,’ [said a bank analyst].

The third type of criticism—and in our view, the most well-founded—was the risk that full disclosure of banks’ capital positions would reveal firms as materially weaker than expected. Policymakers had to live with the real possibility that bank-specific revelations, with bank-specific weaknesses, could undermine confidence in the banking system. Bernanke (2015) noted that disclosing banks’ weaknesses could “possibly [lead] to new runs and further sharp declines in bank stock prices.”

The three highlighted features—and the general perception that the firms would not be allowed to fail—helped stabilize banks’ funding base and made their liabilities less information sensitive. The features also made it easier for the banks to raise capital from private investors to meet the stress test’s incremental capital needs and replace earlier injections of public capital.

**The Macro Part**

The macro context of the stress test reflects the broad mix of fiscal policies, financial policies, and good macro news in 2009. Capital needs estimates fell because the macro environment improved through 2009, while the cumulative force of the many policy measures improved confidence and truncated the most extreme tail of adverse outcomes. As fears of the collapse of the financial system faded and the pace of deterioration in the economy slowed, asset prices increased, aided by the backstop put in place by the Federal Reserve, the Treasury, and the FDIC. The Fed, Treasury, FDIC, FHA, and Congress ultimately committed $12 trillion to respond to the financial crisis, as calculated by Blinder and Zandi (2010). The subsequent rise in asset prices helped mitigate fears of insolvency, even as they had exacerbated concerns about insolvency earlier in the crisis.

**Fiscal Stimulus**

Fiscal expansion buffered the initial shock to the economy. Congress passed the $170 billion Economic Stimulus Act of 2008 and the later $787 billion American Recovery and Reinvestment Act. TARP provided the

---

13 The $12 trillion commitment excludes facilities that were, in principle, unlimited; this includes Federal Reserve programs like the primary, secondary, and seasonal credit facilities; the PDCF; the AMLF; and currency swap lines; the Treasury’s Fannie Mae and Freddie Mac intervention; the FDIC’s bank resolution authority (although costs to the FDIC are ultimately recouped via subsequent assessments of the banking system); and the FHA’s expanded mortgage lending. See Table 1 of Blinder and Zandi (2010) for additional program-specific details.
$250 billion used in the first round of capital injections (the CPP) and provided funding for a mix of other programs including the Bank of America and Citi programs, TALF, PPIP, the SBA loan purchase program, and several automotive-related programs—combined totaling $600 billion. Additional unemployment insurance benefits and other stimulus programs added about $110 billion more. Blinder and Zandi (2010) calculates the total outlays from Congress at $1.7 trillion.

Although actual outlays from these programs took time to roll out—roughly 7% of the ARRA’s total spend-out occurred before the stress tests ended in May 2009—the cumulative impact of fiscal spending stopped the economy from running off a cliff. Blinder and Zandi (2010) project GDP with and without the fiscal stimulus. They find the economy, without the fiscal response, would have been smaller by roughly 3.4% by 2009Q4 (the forecast horizon for the stress tests), amounting to about $500 billion. They find the policy response cumulatively increased payroll employment by 2.9 million by year-end 2010. Furman (2020) finds the Recovery Act alone provided a fiscal expansion of about 2% of GDP, boosting 2009 GDP by about 0.2 percentage points. Combined with automatic stabilizers amounting to 1.4% of GDP, the total magnitude of the countercyclical measures amounted to 3.4%. As Furman (2020) notes, the actions amounted to “the largest fiscal response to a recession in U.S. history.”

Financial Policy

Several Federal Reserve, Treasury, and FDIC programs worked in concert to fight the crisis. Beyond the incremental steps in early 2009 to secure the survival of the GSEs, AIG, and other nonbank institutions like GMAC, we highlight four programs that played outsized roles in supporting the banking system and therefore played a significant—if implicit—role in the stress tests: the FDIC’s Debt Guarantee Program (DGP), the Term Asset-Backed Securities Loan Facility (TALF) and the Public-Private Investment Program for Legacy Assets (PPIP), and the Federal Reserve’s quantitative easing (QE1) program.

The FDIC’s debt guarantee program, one of two components of the Temporary Liquidity Guarantee Program (TGLP), guaranteed new wholesale bank debt. Initially unveiled during the Columbus Day Weekend capital injection, the DGP was only available to banks that also took CPP equity—the equity and guarantees were a joint package, banks could not choose one or the other. Through late 2008 and 2009, the implicit guarantee of bank liabilities helped lend additional credibility to the stress tests by implicitly saying “no more Lehman.” Additionally, bank liability guarantees made pre-provision net revenue estimates more credible.

The TALF provided $200 billion of financing to a broad set of investors to buy new ABS issuances of auto, student, credit card, and Small Business Administration loans. The Treasury provided $20 billion in first-loss protection, which, in effect, magnified the economic impact of the TARP funds used to support the program via leverage. The Fed and Treasury expanded the program to $1 trillion in March 2009 and expanded the set of eligible securities, including commercial mortgage securities. (Clark, Kabaker, and Sachs 2020).

The PPIP expanded on the spirit of TALF by providing funds towards existing securities, rather than newly issued loans. Clark, Kabaker, and Sachs (2020) discuss the difficulties in creating the program—both

---

14 However, the government’s total contribution to GDP growth turned negative in 2010; the average contribution to real GDP growth averaged -0.07% per quarter from 2010 to 2011 and accelerated downwards through 2012 and beyond.
practical and conceptual—but the program ultimately matched private capital with government capital amounting to $24 billion within nine public-private investment funds.

The take-up of TALF and PPIP fell short of the initially announced $1 trillion; the two programs peaked at about $75 billion. However, the programs accomplished their goals: ABS spreads fell, lifting asset prices of legacy loans still on banks’ balance sheets, and encouraging additional new issuance. The programs’ announcements marked the turning point in the residential and commercial mortgage-backed security markets, and by early May, the fire-sales in ABS markets moderated.

The Federal Reserve’s large scale asset purchasing program QE1 also provided critical support during and after the stress tests in 2009. Before the announcement in November 2008, mortgage spreads peaked at approximately 3.4%. After the test concluded in May 2009, they fell below 2%. Kohn and Sack (2020) find QE1’s positive effect on market functioning baked in by mid-2009, even though the Fed had only completed a third of its planned purchases.

The joint effect of fiscal and financial policies improved the macro backdrop of the stress tests. Blinder and Zandi (2010) find GDP growth in 2009 and 2010 with no policy response, either fiscal or financial would have been -7.4 and -3.7%.

Macro Outcomes

Macro policy helped reduce the likelihood—or, at least, the perception—of the most adverse outcomes, and that perception brightened as incoming data improved through the first half of 2009. Of course, macro results are endogenous to policy, and regulators conducted the stress test as part of a broad policy response designed to reduce the likelihood of the most adverse outcomes.

The macro background turned for the better in 2009Q2. What did markets know at the time the Fed released the stress tests on May 7, 2009? Initial claims peaked in late March at 674,000, falling to 605,000 in the week before the Fed released the test results. Auto sales bottomed in February but increased 8% in March (as reported on April 1). The decline in nonfarm payrolls peaked in January 2009 at -741,000, and the pace of losses moderated in April—reported the day after the stress test results—to a still dismal -539,000. The broad equity market, too, rebounded: the S&P 500 reached its nadir in March, after which it began a slow climb. The numbers still reflected a weak and sagging economy, but a hodgepodge of economic indicators suggested the worst might have been behind us by May 2009. The data in 2009Q1 allowed markets to guess the confines of the trough better.

Measures of bank risk turned in 2009Q2, as well. Banks’ stock prices stopped their rapid fall in January 2009—partly due to the price floor imposed by CAP—and bottomed in March 2009, before starting a steady climb leading into the stress test results’ announcement. Banks’ Q1 earnings between April 13 and 17 were better than expected, although the beats mostly came from one-off factors. Other measures of bank risk started improving by April 2009, including banks’ PE ratios, implied volatility for bank stocks, the implied

15 The numbers quoted here reflect the initial as-reported data, which at times differs materially from subsequent revisions. We include the as-reported data as it more accurately captures what markets knew about the economy in real-time.
volatility of banks relative to the market, and the delta on one-year maturity at-the-money puts. Moderating bank risk created a modestly positive sentiment leading into the release of the test results.

The Wall Street Journal leaked loss rates and economic forecasts in the more adverse scenario on April 22, a few weeks before the Fed announced the test results. The assumptions reported in the leaked article implied higher losses than many analysts expected and underscored the test’s credibility. Critically, the lack of bank-specific information in the leak allowed markets to digest the test’s severity separately from the results for specific banks.

We perform an event study to find which events or policies were most effective and report the results in Table [tab:event-study-market]. We calculate banks’ three-day cumulative abnormal returns around several policy events (either announcement or implementation), macro news, and other events. The goal is to identify which events had the largest impact on banks’ valuations as a proxy for which events best helped restore confidence in the banking system. We find cumulative abnormal returns (CAR) by estimating each company’s “normal” beta to the CRSP value-weighted market return using daily data from January 2002 through June 2007:

\[ R_{it} = \alpha_i + \beta_{im} R_{mt} + \epsilon_{it} \]

where \( i \) denotes company \( i \) and \( m \) denotes the market return. We estimate the predicted return using the company’s estimated market beta and the market return. The difference between the predicted return and actual return is the abnormal return, \( AR_{it} \). The cumulative abnormal return is the sum of the daily differences between the company’s actual return and the predicted return over the following three trading days (inclusive of the event day).

We test whether the cumulative abnormal return is statistically different from zero on average across all the firms in the sample. We perform the estimation on three groups of financial companies: “banks and non-banks,” which includes depository and non-depository financial institutions, mortgage brokers and investment banks; “banks,” which consists of all publicly traded commercial banks in the U.S.; and “CPP banks,” which includes the banks that first received TARP money via the CPP in October 2008. Unfortunately, the reality of a fast-moving crisis and response is that some events overlap.

The CPP had the largest impact on bank stocks, with a 4.3% CAR for the broad industry and a large 13.2% CAR for the firms which received CPP funds. Unsurprisingly, the bankruptcy of Lehman in September 2008 dragged returns down for financial companies, and especially for the largest banks: CPP bank stocks’ CAR was -12.6%, however, the effect is not significantly different from zero reflecting considerable heterogeneity in the largest banks’ stock returns response to the bankruptcy. Scanning down the third column of the table, it is clear that both good news and policy boosted banks’ returns: the AIG intervention in September 2008, the April 2009 employment report, and the stress test release all had CAR above 2%. PPIP and PDCF had particularly large effects on CPP banks’ returns, with a 9.9% and 8.0% CAR, respectively.

The stress test occurred in the context of an economy slowly turning for the better. The combined force of the many interventions put a floor under the economy and limited the recession’s severity.
Conclusion

All crises are different, even if they stem from fundamentally similar fears about liquidity and solvency. They all pose a challenge in motivating action before falling into the abyss—before the crisis is so severe, the run so overpowering, that dramatic action is the only choice left.

Stress tests provide a recapitalization strategy if the solvency of the system is in question, and—if credible—it increases the likelihood that the recapitalization comes from private sources. The U.S. experience offers some ideas for how to use the stress test tool, combined with a set of emergency funding protections, in ways that might limit the need for public capital injections or expansions of the safety net that might expose the taxpayer to excessive risk.

Covid-19 has again brought the banks’ capital positions to the fore. Faced with a second systemic financial crisis in little over a decade, policymakers in some countries will confront similar choices and trade-offs in how to restore confidence in the financial system’s capital position. They will undoubtedly see parallels with the 2009 experience.

References


