

STRESS TESTING BANKS

Til Schuermann*

Oliver Wyman

Wharton Financial Institutions Center

First Draft: September 2011

This Draft: April 2012

This Print: April 17, 2012

Prepared for the *Committee on Capital Markets Regulation*

Abstract

How much capital and liquidity does a bank need – to support its risk taking activities? During the recent (and still ongoing) financial crisis, answers to this question using standard approaches, e.g. regulatory capital ratios, were no longer credible, and thus broad-based supervisory stress testing became the new tool. Bank balance sheets are notoriously opaque and are susceptible to asset substitution (easy swapping of high risk for low risk assets), so stress tests, tailored to the situation at hand, can provide clarity by openly disclosing details of the results and approaches taken, allowing trust to be regained. With that trust re-established, the cost-benefit of stress testing disclosures may tip away from bank-specific towards more aggregated information. This still provides the market with unique information (supervisors, after all, have access to proprietary bank data) without dis-incentivizing market participants from producing private information and trading on it – with all the downstream benefits of information-rich prices and market discipline.

Keywords: capital requirements, leverage, systemic risk.

JEL Codes: G21, G28, G20.

* Oliver Wyman and Wharton Financial Institutions Center; til.schuermann@oliverwyman.com. I would like to thank Mark Flannery, Itay Goldstein, Bengt Holmstrom, Bill Janeway, Umit Kaya, Ugur Koyluoglu, Andy Kuritzkes, John Lester, Clinton Lively, Hashem Pesaran, Brian Peters, Barry Schachter, Hal Scott, and members of the Committee for helpful discussions and suggestions, and Cary Lin for helpful research assistance. All errors remain mine, of course.

1. Introduction

There are three kinds of capital and liquidity: 1) the capital/liquidity you have; 2) the capital/liquidity you need (to support your business activities); and 3) the capital/liquidity the regulators *think* that you need.¹ Stress testing, regulatory capital/liquidity and bank-internal (so-called “economic capital/liquidity”) models all seek to do the same thing: to assess the amount of capital and liquidity needed to support the business activities of the financial institution. Capital adequacy addresses the right side of the balance sheet (net worth), and liquidity the left side (share of assets that are “liquid”, however defined). If all goes well, both economic and regulatory capital/liquidity are less than the required regulatory minimum, and their difference (between economic and regulatory) is small, namely that regulatory models do not deviate substantially from internal model results.

Prior to their failure or near-failure, financial institutions such as Bear Stearns, Washington Mutual, Fannie Mae, Freddie Mac, Lehman and Wachovia were adequately capitalized, at least according to regulatory capital rules disclosed in their public filings.² This set of institutions spans a broad range of regulatory capital regimes and regulators: the SEC and Basel 2 capital rules (Bear Stearns, Lehman), the OCC and the Federal Reserve and Basel 1 (Wachovia), the OTS (WaMu), and OFHEO (Fannie and Freddie) – the last actually based on a narrow stress scenario. All firms had broad exposure to residential real estate assets, either in the form of whole loans (mortgages) or securities (MBS) or both, and all had internal risk models which may or may not have deviated materially from the regulatory models (we don’t know as this is/was firm proprietary information). Yet to the question of what is the capital you need vs. the capital you have, in each case the answer came out wrong. To be sure, neither firm-internal (economic) or regulatory capital and liquidity models can guarantee failure prevention; indeed, that is not their purpose as every firm accepts some probability of failure, sized by its risk appetite. But the cascading of defaults, and the resulting deep skepticism of stated capital adequacy by the market, forced regulators to turn to other tools for assessing, in a credible way, the capital adequacy of banks. That tool turned out to be stress testing.³

This paper lays out a framework for the stress testing of banks: why is it useful and why has it become such a popular tool for the regulatory community in the course of the recent financial crisis; how is stress testing done – design and execution; and finally, with stress testing

¹ This pithy summary I owe to Peter Nakada.

² Kuritzkes and Scott (2009) make the case for a more market-oriented assessment of capital adequacy.

³ Flannery (2012) argues that stress tests should be evaluated on a fair value (rather than book capital) basis.

results in hand, how should one handle their disclosure, and should it be different in crisis vs. “normal” times. The framework is equally applicable to capital and liquidity adequacy, but for simplicity the bulk of the discussion will focus on capital.

A successful macro-prudential stress testing program, particularly in a crisis, has at least two components: first, a credible assessment of the capital strength of the tested institutions to size the capital “hole” that needs to be filled, and second a credible way of filling that hole. The U.S. bank stress test in 2009, the Supervisory Capital Assessment Program or SCAP, may serve as a useful example. The U.S. entered 2009 with enormous uncertainty about the health of its banking system. Absent more concrete and credible understanding of the problems on bank balance sheets, investors were reluctant to commit capital, especially given the looming threat of possible government dilution. With a credible assessment of losses under a sufficiently stressful macroeconomic scenario, the supervisors hoped to draw a line in the sand for the markets: fill this hole, and you won’t risk being diluted later because the scenario wasn’t tough enough. Moreover, if some institutions could not convince investors to fill the hole, a U.S. government program, namely Treasury’s Capital Assistance Program (CAP), stood ready to supply the needed capital. Importantly, the U.S. Treasury was a sufficiently credible debt issuer that the CAP promise was itself credible.⁴ All banks with assets greater than \$100bn (YE 2008) were included, accounting for two-thirds of total assets and about half of total loans in the U.S. banking system. In the end, ten of the 19 SCAP banks were required to raise a total of \$75bn in capital within six months, and indeed raised \$77bn of Tier 1 common equity in that period.⁵ None needed to draw on CAP funds.

The European experience in 2010 and 2011 stands in stark contrast to the 2009 SCAP. Against the background of a looming sovereign debt crisis in the peripheral euro-zone countries, the Committee of European Bank Supervisors (CEBS) conducted a stress test of 91 European banks in 2010 covering about two-thirds of total European bank assets and at least half in any given participating country. The stress test included imposing haircuts on the market value of sovereign bonds held in the trading book; the bulk of the sovereign exposure, however, was (and is) in the banking book. Of the 91 banks, only seven were required to raise a total of €3.5bn (< \$5bn at the time) in capital. The level of disclosure provided was rather less than in the SCAP. For instance, loss rates by firm were made available only for two sub-categories: overall

⁴ Note that the act of a sovereign recapitalizing its banks involves that sovereign issuing debt and then investing (“downstreaming”) it as equity in the bank(s).

⁵ <http://www.federalreserve.gov/bankinforeg/scap.htm>.

retail and overall corporate.⁶ By contrast, the SCAP results released loss rates by major asset class such as first-lien mortgages, credit cards, commercial real estate, and so on. Markets reacted benignly nonetheless – until a few months later when Ireland requested financial assistance from the EU and the IMF. Subsequent stress tests of just the Irish banks revealed a total capital need of €24bn; all had previously passed the CEBS stress test. Moreover, to help close the credibility gap, the extent and degree of disclosure was far greater than any of the stress testing exercises to date.⁷ The markets reacted favorably, with both bank and Irish sovereign credit spreads tightening. The stakes for the 2011 European stress test, now conducted by the successor to the CEBS – the European Banking Authority (EBA) – had risen substantially.

The results of 2011 EBA stress test of 90 banks in 21 countries were at first blush similarly mild as the previous year's.⁸ Eight banks were required to raise a total of only €2.5bn. However, the degree of disclosure was much more extensive, nearly reaching the high bar set by the Central Bank of Ireland in March 2011, including information on exposure by asset class by geography. Importantly, all bank level results are available to download in spreadsheet form to enable market analysts to easily impose their own loss rate assumptions. In this way the “official” results were no longer so final: analysts could (and did) easily impose their own sovereign haircuts on all exposures and thus test the solvency of any of the 90 institutions themselves. A summary of the major macro-prudential stress tests to date is provided in Table 3, and a summary of their disclosures in Table 1.

⁶ <http://www.eba.europa.eu/EU-wide-stress-testing/2010/2010-EU-wide-stress-test-results.aspx>.

⁷ <http://www.centralbank.ie/regulation/industry-sectors/credit-institutions/Documents/The%20Financial%20Measures%20Programme%20Report.pdf>.

⁸ <http://www.eba.europa.eu/EU-wide-stress-testing/2011/2011-EU-wide-stress-test-results.aspx>.

	Base & Stress Scenario	Bank level results	Asset/Product level loss rates	Exposure detail (asset class, maturity, geography)	Bank vs. supervisory/3 rd party estimates
SCAP March 2009	Stress	✓	✓	--	--
CEBS July 2010	Both	✓	Retail, all corporate only	--	--
CCAR March 2011	--	--	--	--	--
Ireland March 2011	Both	✓	✓	Sovereign only	✓
EBA July 2011	Both	✓	Retail, corporate, CRE	High	--
CCAR March 2012	--	✓	✓	--	--

Table 1: Summary of disclosures across stress test exercises.

The SCAP was the first of the macro-prudential stress tests of this crisis. But the changes at the micro-prudential or bank-specific level were at least equally significant, and they are summarized in Table 2. With the SCAP, stress testing at banks went from mostly single (or a handful) factor shock to using a broad macro scenario with market-wide stresses; from product or business unit stress testing focusing mostly on losses to firm-wide and comprehensive, encompassing losses, revenues and costs; all tied to a post-stress capital ratio to ensure a going concern.

Pre-SCAP	Post-SCAP
<ul style="list-style-type: none"> • Mostly single shock 	<ul style="list-style-type: none"> • Broad macro scenario and market stress
<ul style="list-style-type: none"> • Product or business unit level 	<ul style="list-style-type: none"> • Comprehensive, firm-wide
<ul style="list-style-type: none"> • Static 	<ul style="list-style-type: none"> • Dynamic and path dependent
<ul style="list-style-type: none"> • Not usually tied to capital adequacy 	<ul style="list-style-type: none"> • Explicit post-stress common equity threshold
<ul style="list-style-type: none"> • Losses only 	<ul style="list-style-type: none"> • Losses, revenues and costs

Table 2: Features of stress testing, pre- and post-SCAP

The remainder of the paper proceeds as follows. Section 2 briefly reviews the scant literature, and Section 3 provides a discussion of how to design the stress scenario. Section 4 describes modeling approaches for the three components needed to implement stress testing: losses, net revenues (profitability), and balance sheet dynamics. Section 5 reviews the disclosure regimes across the different stress tests to date in more detail, presents a discussion of disclosure in “normal” times, and Section 6 provides some concluding remarks.

2. Stress testing in the literature

Stress testing has been part of the risk manager’s toolkit for a long time. It is perhaps the most basic of risk-based questions to want to know the resilience of an exposure to deteriorating conditions, be it a single position or loan or a whole portfolio. Typically the stresses take the form of sensitivities (spreads double, prices drop, volatilities rise) or scenarios (black Monday 1987, Fall of 1998, post-Lehman bankruptcy, severe recession, stagflation). These types of stresses lend themselves naturally to understanding financial risks, particularly in a data rich environment such as found in a trading operation. Nonfinancial risks like operational, reputational and other business risks are much harder to quantify and parameterize yet rely heavily on scenario analysis (earthquakes and other natural disasters, computer hacking, legal risks, and so on). While the original Basel I Accord of 1988 did not make formal mention of stress testing, with the Market Risk Amendment of 1995 stress testing merited its own section and thus became embedded in the regulatory codex. Indeed evidence of stress testing capabilities is a requirement for regulatory approval of internal models.

Risk management as a technical discipline came into its own with the publication of the RiskMetrics technical document in 1994, and stress testing (of both kinds, sensitivities and scenarios) is mentioned throughout. The first edition of Jorion’s standard-setting VaR book (1996) had a subsection devoted to the topic – it was elevated to a chapter in subsequent editions – and surely there are earlier examples. Stress testing as a risk management discipline was found largely in the relatively data rich environment of the trading room, with the closely related treasury function conducting interest rate scenarios and shocks.⁹ The Committee on Global Financial Systems (CGFS) of the BIS conducted a survey on stress testing in 2000, and it reinforces this view.¹⁰ In their summary of the CGFS report, Fender, Gibson and Mosser (2001)

⁹ See Kupiec (1999) and Berkowitz (2000) for more extensive discussions of VaR-based stress testing.

¹⁰ See CGFS (2000) and the summary of its principal findings in Fender, Gibson and Mosser (2001).

point out that most of the scenarios manifest in terms of shocks to market rates, prices or volatilities. Typical examples are equity market crashes such as October 1987, rates shocks such as 1994, credit spread widening such as during the fall of 1998, and so on. Such stress scenarios have the virtue of being unambiguously articulated and defined and are thus transparent and easy to implement and communicate – on assets that have themselves natural market prices or analogs, as is mostly the case in the trading book. More typical banking assets, such as corporate loans (especially to privately held firms) and consumer loans (e.g. auto loans), are less naturally amenable to this approach.

Formal stress testing of the banking book, which is dominated by credit risk, is more recent, in part because quantitative credit risk modeling is itself a newer discipline.¹¹ Perhaps stimulated by the success of RiskMetrics, the late 1990s saw a spurt of activity in the development of credit portfolio models, the two prominent examples being CreditMetrics (1997) and CreditRisk+ (Wilde, 1997).¹² Stress testing, however, did not feature in these papers. Yet as Koyluoglu and Hickman (1998) show quite clearly, all of these credit portfolio models share a common framework of mapping outcomes in the real economy, often represented by an abstract state vector, to the credit loss distribution, and thus should lend themselves naturally to stress testing. With that in mind, Bangia et al. (2002), following broadly the CreditMetrics framework, show how to use credit migration matrices to conduct macroeconomic stress tests on credit portfolios. Foglia (2008) provides a survey of the literature (at least through late 2008) of stress testing credit risk, both for individual banks or portfolios as well as banking systems. More recently, Rebonato (2010) with his suggestively titled book *Coherent Stress Testing* (we return to the problem of coherence below), argues for a Bayesian approach to financial stress testing, i.e. one which is able to formally include expert knowledge in the stress testing design, with an emphasis on exploring causal relations using Bayesian networks.

With few exceptions, regulatory requirements on stress testing were thin prior to the crisis, though considerable expectations about stress testing capabilities were voiced in supervisory guidance in the U.S. Examples include the Joint Policy Statement on Interest Rate Risk (SR 96-13), guidance on counterparty credit risk (SR 99-03¹³), as well as country risk management (SR 02-05). But banks had significant discretion with regard to specific design and implementation

¹¹ To be sure, the credit rating agencies, having been in the business of rating corporate bonds for nearly a century, likely employ stress testing in their bond rating methodology, but old documentation to this effect is hard to come by.

¹² For an excellent overview and comparison of these and related models, see Koyluoglu and Hickman (1998).

¹³ The most recent guidance on counterparty credit risk, SR 11-10, has greatly expanded on stress testing expectations. All SR letter can be found at <http://www.federalreserve.gov/bankinforeg/srletters/srletters.htm>.

of their stress tests. Brian Peters, then head of risk in bank supervision at the New York Fed, observed in March 2007 at an industry conference that no firm had a fully-developed program of integrated stress testing that captured all major financial risks on a firm-wide basis.¹⁴ Market risk stress tests were most advanced, while corporate or enterprise-wide stress testing, whereby all businesses were subjected to a common set of stress scenarios, was at best in a developmental phase.

3. Designing the Stress Scenario

One of the principal challenges faced by both the supervisors and the firms in designing stress scenarios is coherence. The scenarios are inherently multi-factor: we seek to develop a rich description of adverse states of the world in the form of several risk factors, be they financial or real, taking on extreme yet coherent (or possible) values. It is not sufficient to specify only high unemployment or only significant widening of credit spreads or only a sudden drop in equity prices; when one risk factor moves significantly, the others don't stay fixed. The real difficulty is in specifying a coherent joint outcome of all the relevant risk factors. For instance, not all exchange rates can depreciate at once; some have to appreciate. A high inflation scenario needs to account for likely monetary policy responses, such as an increase in the policy interest rate. Every market shock scenario resulting in a flight from risky assets – “flight to quality” – must have a (usually small) set of assets that can be considered safe havens. These are typically government bonds from the safest sovereigns (e.g. U.S., Japan, Germany, Switzerland). To be sure, as sovereign government budgets are increasingly strained, questioning the low-risk assumption of those treasury instruments would certainly be a worthwhile stress scenario, but it would need to define an alternative “risk-free” asset class to which capital can flee.

While the problem of coherence is generic to scenario design, it is especially acute when considering stress scenarios for market risk, i.e. for portfolios of traded securities and derivatives. These portfolios are typically marked to market as a matter of course and risk managed in the context of a value-at-risk (VaR) system. Practically this means that the hundreds of thousands (or more) positions in the trading book are mapped to tens of thousands of risk factors, and those risk factors are tracked on a (usually) daily basis and form the “data” used to estimate risk parameters like volatilities and correlations. Finding coherent outcomes in such a high dimensional space, short of resorting to historical realizations, is daunting indeed.

¹⁴ Presentation delivered at Marcus Evans conference “Implementing Stress Tests into the Risk Management Process”, Washington DC, March 1-2, 2007.

Compounding the problem is the challenge of finding a scenario where the real and the financial factors are jointly coherent. The 2009 SCAP had a rather simple scenario specification. The state space had but three dimensions – GDP growth, unemployment, and house price index (HPI) – and the market risk scenario was based in historical experience: an instantaneous risk factor impact reflecting changes from June 30 to December 31, 2008. This period represented a massive flight to quality, the markets experienced the failure of at least one global financial institution (Lehman), and risk premia at the time arguably placed a significant probability on the kind of adverse real economic outcome painted by the tri-variate SCAP scenario. This solution achieved a loose coherence of the real and financial stress. The price one pays for choosing a historical scenario is the usual one: it does not test for something new. Figures 3 and 4 compare some of these risk factors (real GDP, unemployment, equity and home prices indices) across the three U.S. stress tests to date, both to each other as well as to actual realizations since 2008 Q4.

For the 2011 EBA test, the supervisors specified over 70 risk factors for the trading book, eight macro-factors for each of 21 countries (macro-factors such as GDP growth, inflation, unemployment, real estate price indices – residential and commercial, short and long term government rates, and stock prices), plus sovereign haircuts across seven maturity buckets. The macroeconomic stress scenario was generated by economists at the ECB with reference to the EU Commission baseline economic forecast.

All supervisory stress tests to date have imposed the same scenario on all banks. Naturally, any scenario may be especially severe for some banks and much less so for others, depending on the business mix and geographic footprint. This one-size-fits-all approach is analogous to the problem of regulatory vs. internal economic capital models: the former by design is the same for all banks, while the latter, being bespoke to a given bank, directly takes account of the particular business mix of that bank. This problem of same vs. bespoke stress scenario becomes especially acute when we move from crisis times, when there may be less debate about what a relevant adverse scenario might look like, to “normal” times. The 2011 and 2012 CCAR recognized this problem and asked banks to submit results using their own scenarios (baseline and stress) in addition to results under the common supervisory stress scenario. This was an important step forward from the 2009 SCAP: by asking banks to develop their own stress scenario(s), which was to reveal the particular sensitivities and vulnerabilities of their portfolio and business mix, supervisors could learn from the banks about what they thought to be the high risk scenarios. This is useful not just for micro-prudential supervision – learning about the risk of a given bank – but also for macro-prudential supervision by allowing for the possibility of learning about

common risks across banks hitherto undiscovered or under-emphasized. With this dual approach, supervisors could directly compare results across banks from the common scenario without sacrificing risk-discovery.

4. Executing on the stress scenario: losses and revenues

With the macro-scenario in hand, how does one arrive at the corresponding micro-outcomes: losses and revenues under adverse market and macroeconomic conditions? To date there is very little discussion in the public domain on how to solve this problem, except perhaps for stress testing the trading book. Indeed, one of the more important contributions of the supervisory stress tests in the U.S. and Europe has been the accompanying methodology documents disclosed by the supervisors which are, understandably, more heavily focused on the banking book.¹⁵

4.1. Modeling losses

For a firm active in many markets (product and geography), the first task is to map from the few macro-factors into the many intermediate risk factors that drive losses for particular products by geography. The EBA was forced to confront the geographic heterogeneity problem directly by virtue of spanning 21 sovereign nations with rather different economies. U.S. supervisors, stress testing an economic region just somewhat smaller than that of the EBA, left the task of accounting for the not inconsiderable geographic heterogeneity to individual firms. Regional differences are critical in modeling losses for real estate lending (residential and commercial) but is hardly limited to those products. Since the U.S. experiences regional business cycles – the national business cycle obscures considerable variation across states – nearly all lending has some geographic component. For example, credit card losses are especially sensitive to unemployment, and in July 2011, with the national rate at 9.1%, the state-level unemployment rate ranged from 3.3% in North Dakota to 12.9% in Nevada. Similar dynamics are at work in wholesale lending, particularly for SME (small and medium enterprise) lending whose performance has a strong geographic component.

The problem of mapping from macro to more intermediate risk factors is not limited to geography. An interesting example is auto lending and leasing where the collateral assets are used cars. While auto sales invariably decline in a recession, and the decline in 2008-2009 was unprecedented in the post-war period, used car sales typically suffer less. Yes, households buy

¹⁵ For SCAP, see <http://www.federalreserve.gov/bankinforeg/bcreg20090424a1.pdf>. For EBA, see <http://www.eba.europa.eu/EU-wide-stress-testing/2011/The-EBA-publishes-details-of-its-stress-test-scena.aspx>. For 2011 and 2012 CCAR, see <http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20110318a1.pdf> and <http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20120313a1.pdf> respectively.

fewer cars in a recession, but if they do need to purchase a car, it is relatively more likely to be a used car. So even if the default rate on auto loans increases significantly during a recession, the corresponding loss given default (LGD) or loss severity need not. A useful indicator of the health of the used car market, and thus the collateral of an auto lending portfolio, is the Manheim index. Over the course of the most recent recession (Dec. 2007 – June 2009), the index rose 4% while total new auto and light truck sales declined by 37%.

The problem of loose coupling of loss severity to the business cycle is not limited to auto loans. Acharya et al. (2007) show that for corporate credit, an important determinant of LGD is whether the industry of the defaulted firm is in distress at the time of default. The authors make a compelling asset specificity argument: if the airline industry is in distress, and a bank is stuck with the collateral on defaulted aircraft loans or leases, it will be hard to sell those aircraft except at very depressed prices. The healthcare sector may be relatively robust at that time, as indeed it has been in the recent recession, but it is difficult to transform an airplane into a hospital.

The EBA disclosure on methodology is especially rich. In the March 2011 document, for example, detailed guidance is provided on stressed probabilities of default (PDs) and stressed LGDs. Note that such guidance presumes that a bank has implemented an internal credit rating system for its commercial loan portfolio. For a Basel II bank this may not be unreasonable since internal ratings, mapped to a common external scale such as those used by the rating agencies, are a cornerstone of the Accord. With a credit rating (internal or external) in hand, computing stressed default rates for the portfolio becomes a straightforward exercise, either by assigning higher PDs to a given rating, or by imposing a downward migration on the current portfolio.¹⁶ Since the EBA stress test was based on risk weighted assets (RWA) computed using Basel II risk weights which are ratings sensitive, banks were forced to make use of stress migration matrices to compute not only increased defaults (the last column of the matrix) but also the entire future ratings distribution to arrive at the correct RWA value. The U.S. stress tests were conducted under Basel I risk weights which are not obligor ratings sensitive. The fuss about RWA calculations matters since the denominator of capital ratios, used to determine whether or not a bank needs to raise capital, is RWA. To be sure, this complicates any comparison of U.S. and European stress test results.

¹⁶ Of the 90 participating banks, 59 were so called IRB (internal ratings based) banks, meaning their internal models were validated to the supervisor's satisfaction for at least one regulatory portfolio (e.g. corporate, commercial real estate, etc.). Non-IRB banks were given very non-specific guidance (EBA 2011a, Section 5.5.1.1).

Implementation in the trading book is more straightforward and has a rich discussion in the public domain; see inter alia Allen, Boudoukh, and Saunders (2004), Jorion (2007), or Rebbonato (2010). In a nutshell, existing positions are simply repriced using the stress scenario risk factors, subject to the proviso that the risk factor mapping problem, discussed in Section 3, has been solved. The corresponding problem of stressing the counterparty credit risk that comes with derivatives activities has received less attention.¹⁷ Counterparty credit risk arises when, in a derivative transaction revalued to the stress scenario, the bank finds itself in the money (i.e. enjoys a derivative receivable) yet cannot be sure that the counterparty to the transaction will be solvent to make good on the payment. Thus the value is discounted, where the discount is a function of the expected default likelihood of the counterparty *under the stress scenario*, which presumably is higher than today. This adjustment is called a credit value adjustment (CVA), and banks with significant derivative activities manage CVA as a matter of course. As Canabarro (2010) and Hopper (2010) point out, the modeling challenge to stress testing counterparty credit risk is considerable. Not only does the PD of the counterparty change in a stressful environment, but so does the exposure. Thus any CVA stress test involves two distinct simulation exercises. If the collateral posted by the counterparty is anything other than cash or cash equivalent, a revaluation of that collateral under the same stress scenario needs to be added to the process.¹⁸

4.2. Modeling revenues

Implementing stress scenarios on the revenue side of the equation remains largely a black box and seems far less well developed than stress testing for losses. Neither the 2009 SCAP nor the otherwise richly documented 2011 EBA disclosures devoted much space nor revealed much detail about the methods and approaches for computing revenues under stressful conditions. Total income in banks can be roughly divided into interest and non-interest income. Interest income is clearly a function of the yield curve and credit spreads posited under the stress scenario, but what the net impact of rising or falling rates are on bank profitability remains ambiguous, perhaps in part because of interest rate hedging strategies (English 2002, Purnanandam 2007). The impact of stress scenarios on noninterest income, which includes service charges, fiduciary, fees, and other income (e.g. from trading), is far harder to assess, and there is precious little discussion of its determinants in the literature. This is concerning since Stiroh (2004) shows that not only has the share of noninterest income been steadily rising in U.S.

¹⁷ For an excellent treatment, see Canabarro (2010) and Hopper (2010).

¹⁸ There is the added complication that major derivatives dealers actively manage CVA risk using a range of strategies and instruments that themselves vary in price and availability depending on market conditions.

banks, from 25% in 1985 to 43% in 2001, but it is associated with greater volatility and lower risk-adjusted returns. If we compare the 2009 SCAP, the 2011 EBA and the 2012 CCAR stress tests, the median bank in the U.S. was able to cover about 58% of its total projected losses with profits (including reserve releases, if any) in 2009 and 63% in 2012,¹⁹ compared with 66% in the European case. As Figure 1 shows, there is considerable variability across banks, especially in the EBA test, where in some cases profits even under the stress scenario are projected to outpace losses 4:1!

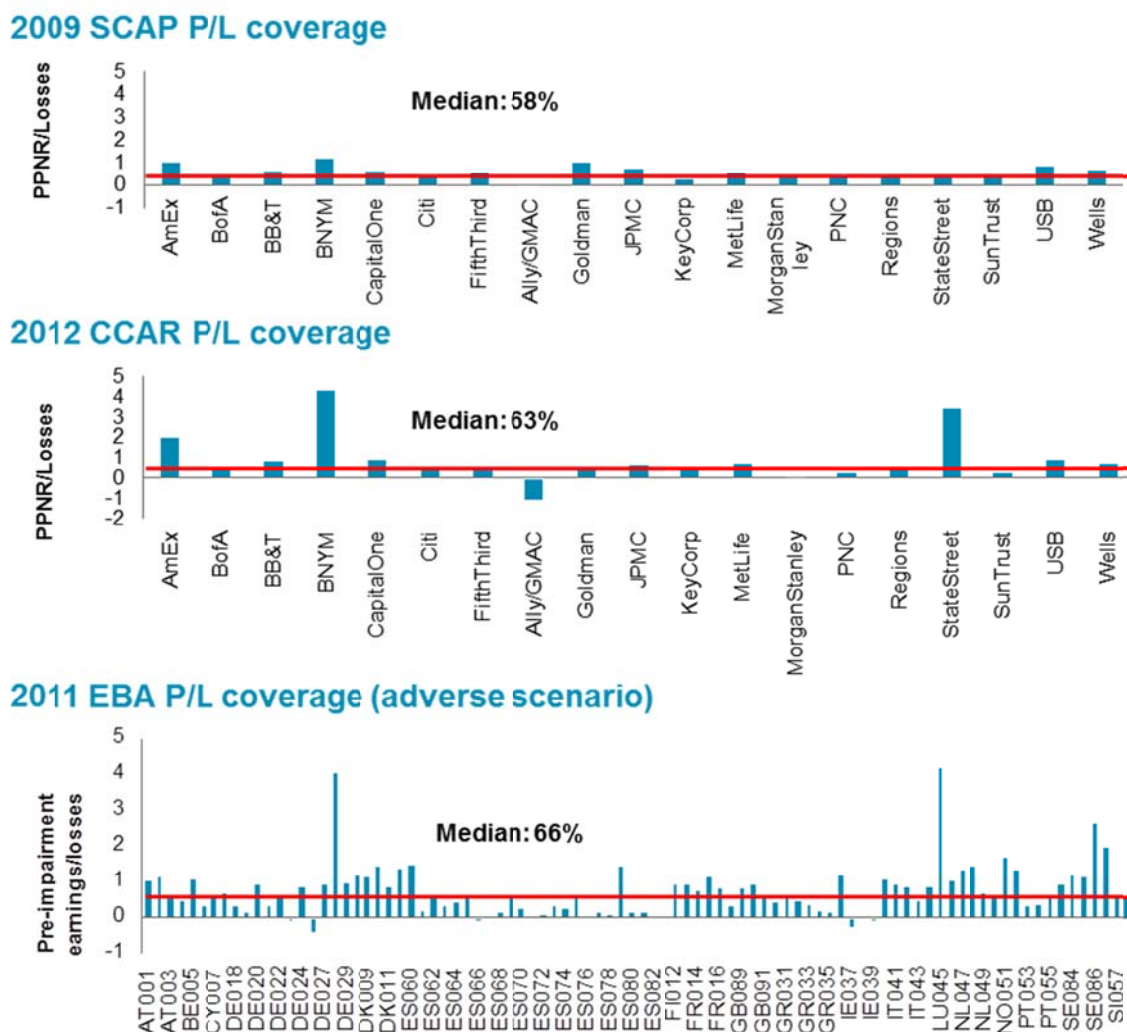


Figure 1: Projected coverage of losses with profits in 2009 SCAP and 2011 EBA stress tests.

¹⁹ PPNR calculations in the 2012 CCAR were net of operational risk related losses, OREO expenses, as well as mortgage repurchase and put-back costs, meaning these items were not reported separately (though they totaled \$115bn for all 19 banks) (Board of Governors, 2012).

4.3. Modeling the balance sheet

Recall that capital adequacy is defined in terms of a capital ratio, roughly capital over assets. Of course both the numerator and denominator are nuanced. All supervisory stress tests have insisted to varying degrees that the relevant form of capital be common equity. The 2010 CEBS test allowed for some forms of hybrid capital typical of state participations, but the requirements were tightened a year later. As discussed in Section 4.1, the denominator is typically risk-weighted assets (RWA), where the risk weights are determined by the prevailing regulatory capital regime, namely Basel I (in the U.S. cases of the SCAP and CCAR) and Basel II (in the two Europe-wide and the Irish stress tests). The many subtleties of what this implies is beyond the scope of this paper, but suffice it to say that a bank may be forced to raise capital under one regime but not the other, and without considerable detail about the portfolio, there is no way to know which regime will result in a more favorable treatment.

Regardless of the risk weight regime, determining post-stress capital adequacy requires modeling both the income statement and the balance sheet, both flows and stocks, over the course of the stress test horizon, which is typically two years.²⁰ This is illustrated in Figure 2 below. The point of departure is the current balance sheet, at which point the bank meets the required capital (and, if included, liquidity) ratios. The starting balance sheet generates the first quarter's income and loss, which in turn determines the quarter-end balance sheet. The modeler is then faced with the problem of considering the nature and amount of new assets originated and/or sold during the quarter, and any other capital depleting or conserving actions such as acquisitions or spin-offs, dividend changes or share (re-)purchase or issuance programs, including employee stock and stock option programs. The problem of balance sheet modeling exists under a static (be it in raw form, as in the 2011 EBA, or in risk weighted form, as in the 2009 SCAP) or dynamic balance sheet assumption. The bank should not drop below the required capital (and liquidity) ratios in any quarter. Moreover, at the end of the stress horizon, the bank needs to estimate the amount of reserves needed to cover expected losses on loans and leases for the following year. In this way the stress tests are really three years (or T+1 years for a T-year stress test).

²⁰ The horizon is 9 quarters for the CCAR as it is based on Q3, not Q4, balance sheets.

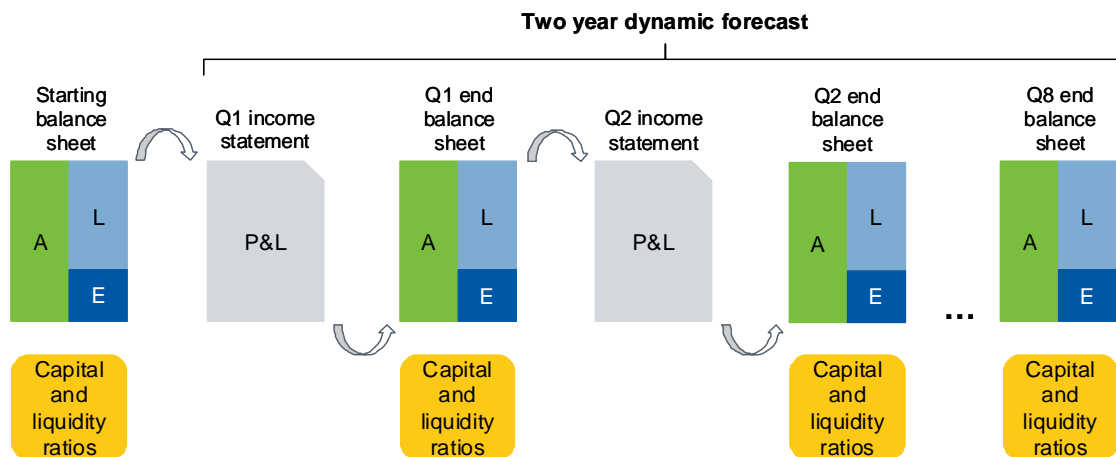


Figure 2: Stress testing balance sheet and income statement dynamics.

5. Stress testing disclosure

Stress testing is here to stay, whether because it is just good risk management practice, or because it is enshrined in legislation (through the Dodd-Frank Act). In the debate on the disclosure regime, it is not clear that more is always better. We divide the discussion into crisis and non-crisis or normal times, with the simple point that normal times may not require or even desire the same degree of transparency as is clearly needed in times of crisis.

We have seen very large differences in disclosure across the different supervisory stress tests, as summarized in Table 1. The SCAP in 2009 opened Pandora’s box by disclosing projected stress losses for each of the 19 participating banks, for eight different categories or asset classes, as well as resources other than capital to absorb losses (mostly pre-provision net revenue and reserve releases, if any). Until then, regulatory disclosures (e.g. Y-9C reports for U.S. bank holding companies) reported only realized losses (the past), not projected losses (a possible future). This allowed the market to easily check the severity of the stress test, not just in terms of the scenario, but much more importantly in terms of the resulting outcomes at the bank level. Given the crisis of confidence prevalent in the market at the time, this amount of transparency was crucial. Two years later, the CCAR displayed a radically different disclosure regime: only the macro-scenario was published, but no bank level results. The only indication of bank level outcomes were subsequent dividend and other capital actions announced by some banks: banks allowed to raise their dividends were interpreted to have “passed” the stress test. The market digested this meager information event without a hiccup.

Dodd-Frank, however, requires the Fed to disclose results of regular stress testing, and with the 2012 CCAR, and the accompanying rules (final and proposed²¹), we got a glimpse of what regular disclosure might look like. The 2012 CCAR disclosed nearly the same level of detail as the 2009 SCAP, namely bank-level loss rates and dollar losses by major regulatory asset class (following the categories of the FR Y-9C bank holding company reports): first and second lien mortgages, commercial and industrial (C&I) lending, CRE, credit cards, other consumer, and other loans. In addition, the Fed reported dollar PPNR, gains/losses on the AFS/HTM securities portfolio, as well as trading and counterparty losses for those firms who were required to conduct the trading book stress.²² And, as with the 2009 SCAP, the numbers reported were supervisory estimates, not the bank-own estimates of losses (and PPNR) under the stress scenario.

By contrast, the 2011 Irish and 2011 Europe-wide EBA stress tests, both of which disclosed after the CCAR, were considerable in their detail, including comparison of bank and third-party estimates of losses in the Irish case (revealing the bias any bank is likely to have when estimating its own potential losses), and data in electronic, downloadable form in the EBA case. Ireland especially was suffering from an acute credibility problem, having emerged in July 2010 from the CEBS stress test with flying colors only to require massive external aid four months later.

This divergent experience between Europe and the U.S. provides some hints on how to design a disclosure regime during “normal” times. The discussion on the benefits *and* costs of stress test disclosures in Goldstein and Sapra (2012) will help us. They argue persuasively that in a world with frictions and strategic environments, the benefits (better market discipline) may not outweigh the costs: banks may make poor portfolio choices designed to maximize the chance of passing the test (window dressing) and thereby give up longer term value; traders may place too much weight on the public information of stress test disclosure and be disincentivized to produce private information about the banks; and finally, with information content of market prices now damaged, market discipline is harmed, and supervisors will find market prices less useful for policy decisions (micro- as well as macro-prudential).

To be sure, some disclosure is still preferable to no disclosure, and Goldstein and Sapra propose disclosing aggregated but not necessarily bank-specific results, with sufficient information about category outcomes (loss rates by major asset class, for instance). Aggregation has the advantage of being less wrong since idiosyncratic errors in estimating bank conditions under hypothesized stress scenarios are averaged away. In this way supervisors can still provide

²¹ <http://www.gpo.gov/fdsys/pkg/FR-2011-12-01/pdf/2011-30665.pdf>

²² In 2012, these were the six institutions with the largest trading portfolios.

useful macro-prudential information which only they can provide – loss rates by asset class, total capital decline in the system (or significant fraction of the banking system) – without drowning out signals about individual banks from the market participants themselves.

During times of crisis, with enormous uncertainty about the health of the banking system, the ability of supervisors to correctly assess the health of individual firms, and the resulting inability of the market to be able to tell a good bank from a bad, then the benefit of detailed bank-specific stress test disclosure may very well be significant. Indeed Goldstein and Sapra argue that stress test disclosures, when more disaggregated, ought to be accompanied by detailed descriptions of the exposures of the banks. This is precisely what was done in the Irish bank stress test of 2011, an acute case of loss of confidence (and subsequent regaining), as well as the 2011 EBA stress test. Because credibility of European supervisors was rather low by that point, only with very detailed disclosure, bank by bank, of their exposures by asset class, by country, by maturity bucket, could the market do its own math and arrive at its own conclusions.

Between March 2009 and March 2011, the 19 SCAP banks had raised about \$300bn in capital, the S&P500 had increased by 65%, the economy was no longer in recession, and arguably the supervisory agencies had regained credibility. The non-event of the non-disclosure of the 2011 CCAR suggests that the market seemed content to live in a state of “symmetric ignorance,” to borrow a term from Dang, Gorton and Holmstrom (2010). Of course this might change should the economy receive another adverse shock, but until it does, it is not clear that an EBA-like disclosure regime is necessarily desirable nor stability enhancing. Europe, by contrast, is not out of the weeds yet (as of this writing). Yet even the EBA is not limitless with its disclosure of the 2011 stress test results. It is worth noting that funding liquidity was also stressed at banks, but without disclosing the results. Because liquidity positions are highly dynamic and thus subject to rapid change, snapshot disclosure, especially with delay (the as-of date for the 2011 EBA stress test was YE 2010), is unlikely to be informative at the time of disclosure.²³

Recall the discussion in the introduction: regulatory capital models (risk weighting), internal economic capital models and stress testing all have the same goal, namely to determine the amount of capital needed to support the business (risk taking) of the bank. Both regulatory and economic capital models (and especially the former) evolve very slowly and thus have difficulty adapting to financial innovation and rapidly changing macro conditions. Indeed, some of the

²³ Reuters, Sept. 2, 2011, “EBA won’t seek disclosure of bank liquidity.” Available at <http://www.reuters.com/article/2011/09/02/idUSL5E7K23PI20110902>.

innovation is motivated by those slowly evolving, one-size-fits-all regulatory capital rules. Moreover, bank balance sheets are notoriously opaque and subject to easy-to-hide asset substitution (higher risk for lower risk assets); Morgan 2002. Stress tests, especially macro-prudential supervisory stress tests, are by construction adapted to the then current environment and bank portfolios. Between balance sheet opacity, asset substitution and regulatory arbitrage, it is easy to see the value of a “pop quiz” in the form of bespoke stress testing (Acharya et al. 2011).

6. Conclusion

The problem of sizing the amount of capital needed to support the risk taking of a bank is not new; but the use of broad-based supervisory stress tests for an entire banking system is. The first use was in 2009 in the U.S., and its success there has made it the supervisory and risk management hammer to deal with all nails. A critical component of the exercise is the disclosure of the results. The reason stress testing became an imperative was precisely because existing approaches that were publicly disclosed, such as regulatory capital ratios, were no longer informative and heavily (if not entirely) discounted by the market. To regain credibility, supervisory authorities needed to disclose enough to allow the market to “check the math.”

But broad-based supervisory stress testing has not been universally successful, as the 2010 and 2011 European experience has shown. Nor is it clear how useful such broad supervisory stress testing with concomitant disclosure will be as a matter of routine. Its value in the crisis was undoubtedly its “pop quiz” nature. It was sprung on the banks at short notice, and thus very difficult to manipulate through careful pre-positioning, and it was tailored to the situation at hand, genuinely revealing new information to all participants and the public. As a result, trust was regained. Once trust has been re-established, the cost-benefit of stress testing disclosures may tip away from bank-specific towards more aggregated information. This still provides the market with unique information (supervisors, after all, have access to proprietary bank data) without dis-incentivizing market participants from producing private information and trading on it – with all the downstream benefits of information-rich prices and market discipline.

References

- Acharya, Viral V., Sreedhar T. Bharath, and Anand Srinivasan, 2007. "Does Industry-wide Distress Affect Defaulted Firms? Evidence from Creditor recoveries." *Journal of Financial Economics* 85, 787-821.
- Acharya, Viral, Hamid Mehran, Til Schuermann, and Anjan Thakor, 2011. "Robust Capital Regulation." Federal Reserve Bank of New York Staff Report No. 490.
- Allen, Linda, Jacob Boudoukh, and Anthony Saunders, 2004. *Understanding Market, Credit and Operational Risk: The Value at Risk Approach*. Blackwell: New York, NY.
- Bangia, Anil, Francis X. Diebold, André Kronimus, Christian Schagen and Til Schuermann, 2002. "Ratings Migration and the Business Cycle, With Applications to Credit Portfolio Stress Testing." *Journal of Banking & Finance* 26:2/3, 235-264.
- Berkowitz, Jeremy, 2000, "A Coherent Framework for Stress Testing." *Journal of Risk* 2, 1-11.
- Board of Governors of the Federal Reserve System, 2009a. "The Supervisory Capital Assessment Program: Design and Implementation." April 24, 2009. <http://www.federalreserve.gov/bankinforeg/bcreg20090424a1.pdf>.
- _____, 2009b. "The Supervisory Capital Assessment Program: Overview of Results." May 7, 2009. <http://www.federalreserve.gov/bankinforeg/bcreg20090507a1.pdf>.
- _____, 2011. "Comprehensive Capital Analysis and Review: Objectives and Overview." March 18, 2011. <http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20110318a1.pdf>.
- _____, 2012. "Comprehensive Capital Analysis and Review 2012: Methodology and Results for Stress Scenario Projections. March 13, 2012. <http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20120313a1.pdf>.
- Basel Committee on Banking Supervision, 2009, "Principles for Sound Stress Testing Practices and Supervision." Available at <http://www.bis.org/publ/bcbs155.htm>.
- Canabarro, Eduardo, 2010. "Pricing and Hedging Counterparty Risk: Lessons Re-Learned?" ch. 6 in Eduardo Canabarro (ed.), *Counterparty Credit Risk*, Risk Books, London, UK.
- Committee on the Global Financial System, 2001, "A Survey of Stress Tests and Current Practice at Major Financial Institutions." Available at <http://www.bis.org/publ/cgfs18.htm>.
- Dang, Tri V., Gary Gorton and Bengt Holmstrom, 2010. "Financial Crises and the Optimality of Debt for Liquidity Provision." Working paper, available at <http://mfi.uchicago.edu/publications/papers/ignorance-crisis-and-the-optimality-of-debt-for-liquidity-provision.pdf>.
- English, William B., 2002. "Interest Rate Risk and Bank Net Interest Margins." *BIS Quarterly Review*, December, 67-82.
- European Banking Authority, 2011a. "2011 EU-Wide Stress Test: Methodological Note." 18 March 2011. Available at <http://www.eba.europa.eu/EU-wide-stress-testing/2011/The-EBA-publishes-details-of-its-stress-test-scena.aspx>.
- European Banking Authority, 2011b. "European Banking Authority 2011 EU-Wide Stress Test Aggregate Report." 15 July 2011. Available at <http://www.eba.europa.eu/EU-wide-stress-testing/2011/2011-EU-wide-stress-test-results.aspx>.

- Fender, Ingo, Michael S. Gibson and Patricia C. Mosser, 2001. "An International Survey of Stress Tests." Federal Reserve Bank of New York *Current Issues in Economics and Finance* 7:10.
- Flannery, Mark J., 2012, "Measuring Equity Capital for Stress-Testing Large Financial Institutions." Working paper
- Foglia, Antonella, 2008. "Stress Testing Credit Risk: A Survey of Authorities' Approaches." Banca d'Italia Occasional Paper No. 37.
- Goldstein, Itay and Haresh Sapra, 2012, "Should Banks' Stress Test Results be Disclosed? An Analysis of the Costs and Benefits." Working paper.....
- Gupton, Gregory M, Christopher Finger and Mickey Bhatia, 1997, "CreditMetricsTM – Technical Document," this version: April 2. J.P. Morgan. Available at http://www.defaultrisk.com/pdf6j4/creditmetrics_techdoc.pdf.
- Hirtle, Beverly, Til Schuermann and Kevin Stiroh, 2009. "Macroprudential Supervision of Financial Institutions: Lessons from the SCAP." Federal Reserve Bank of New York Staff Report No. 409.
- Hopper, Gregory, 2010. "Stress Testing and Scenario Analysis: Some Second Generation Approaches" ch. 11 in Eduardo Canabarro (ed.), *Counterparty Credit Risk*, Risk Books, London, UK.
- Jorion, Philippe, 2007. *Value at Risk: The New Benchmark for Managing Financial Risk*, 3rd Edition. McGraw Hill: New York, NY.
- Koyluoglu, H.Ugur and Andrew Hickman, 1998. "Credit risk: Reconcilable differences." *Risk* 11 (10), 56–62.
- Kupiec, Paul H. (1998). "Stress Testing in a Value at Risk Framework," *Journal of Derivatives* 6(1): 7–24.
- Kuritzkes, Andrew and Hal Scott, 2009. "Markets are the best judge of bank capital." *Financial Times*, 23 September.
- Kuritzkes, Andrew, Til Schuermann and Scott M. Weiner, 2005, "Deposit Insurance and Risk Management of the U.S. Banking System: What is the Loss Distribution Faced by the FDIC?" *Journal of Financial Services Research* 27(3), pp. 217-243.
- Morgan, Donald P., 2002, "Rating Banks: Risk and Uncertainty in an Opaque Industry", *American Economic Review* 92(4), 874-888.
- Pesaran, M. Hashem, Björn-Jakob Treutler, Til Schuermann and Scott M. Weiner, 2006. "Macroeconomic Dynamics and Credit Risk: A Global Perspective." *Journal of Money, Credit and Banking* 38:5, 1211-1261.
- Purnanandam, Amiyatosh, 2007. "Interest Rate Risk Management at Commercial Banks: An Empirical Investigation." *Journal of Monetary Economics* 54, 1769-1808.
- Rebonato, Riccardo, 2010. *Coherent Stress Testing: A Bayesian Approach to the Analysis of Financial Stress*. John Wiley & Sons: New York.
- Stiroh, Kevin, 2004, "Diversification in Banking: Is Noninterest Income the Answer?" *Journal of Money, Credit and Banking* 36:5, 83-882.

Wilde, Tom, 1997, "CreditRisk+ – A Credit Risk Management Framework". Available at <http://www.csfb.com/institutional/research/assets/creditrisk.pdf>.

	Target capital ratio*	# of participating banks	Participation criteria (total coverage)	Balance sheet assumptions	Total required capital raise (for # of banks)	Risk types included: Market, Credit, Liquidity (funding), Operational
SCAP March 2009	<ul style="list-style-type: none"> • 4% T1C • 6% T1 	19	<ul style="list-style-type: none"> • All bank holding companies with at least \$100 bn total assets • (~2/3 of total banking assets) 	Constant RWA	\$75 bn (19)	M, ** C
CEBS July 2010	<ul style="list-style-type: none"> • 6% T1 	91 (20 countries)	<ul style="list-style-type: none"> • Largest banks in country until at least 50% of total assets are included • (~2/3 of total banking assets) 	Constant total assets	€3.5 bn (7)	M, C
CCAR March 2011	<ul style="list-style-type: none"> • 5% T1C 	19	<ul style="list-style-type: none"> • Original SCAP-19 	none	--	M, C
Ireland March 2011	<ul style="list-style-type: none"> • 6% T1C • 10.5% T1C (in base) 	4	<ul style="list-style-type: none"> • Largest banks not in wind-down mode 	Allowed for balance sheet shrinkage	€24bn (4)	M, C, L, O
EBA July 2011	<ul style="list-style-type: none"> • 5% T1C 	90 (21 countries)	<ul style="list-style-type: none"> • Largest banks in country until at least 50% of total assets are included • (~2/3 of total banking assets) 	Constant total assets	€2.5 bn (8)	M, C, L ***, O
CCAR March 2012	<ul style="list-style-type: none"> • 5% T1C • 4% T1; 8% Total; 3-4% leverage 	19	<ul style="list-style-type: none"> • SCAP-19 • An additional 11 BHCs with assets >\$50bn 	none	-- ****	M, C, O

Table 3: Summary of macroprudential stress tests to date

*: T1: Tier 1 capital ratio; T1C: Tier 1 Common (or Core) capital ratio

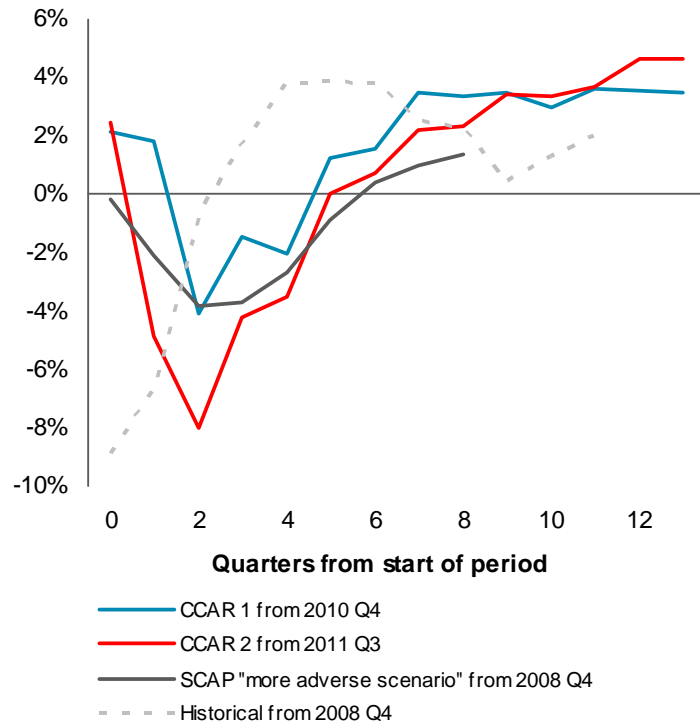
**: Only banks with at least \$100 bn in trading assets were required to conduct the market risk stress test

***: Liquidity risk was not directly assessed, though funding stresses were taken into account, especially as related to sovereign stress impacting funding costs for financial institutions.

****: 4 of the 19 did not pass in the sense of having not gaining non-objection to their submitted capital plans.

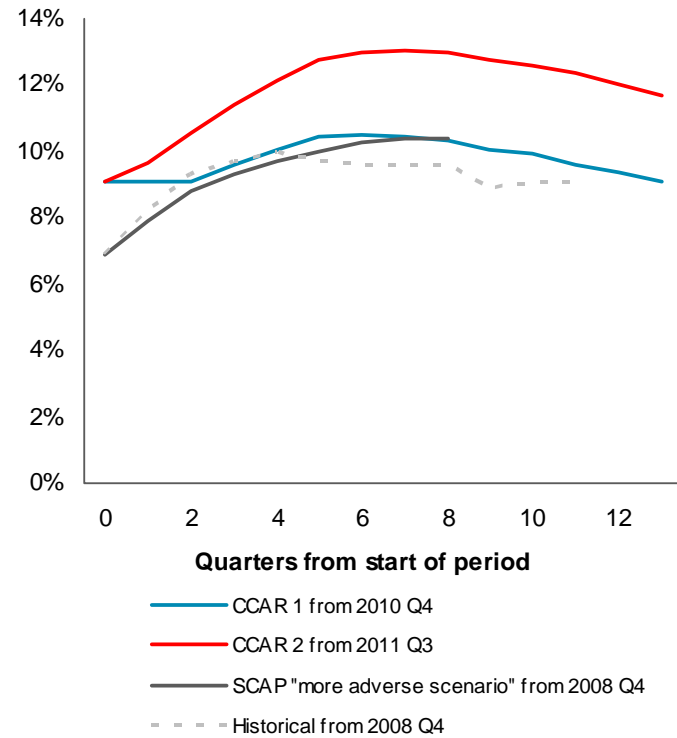
Real GDP growth

Stress-test scenarios vs. recent historical observations



Unemployment rate

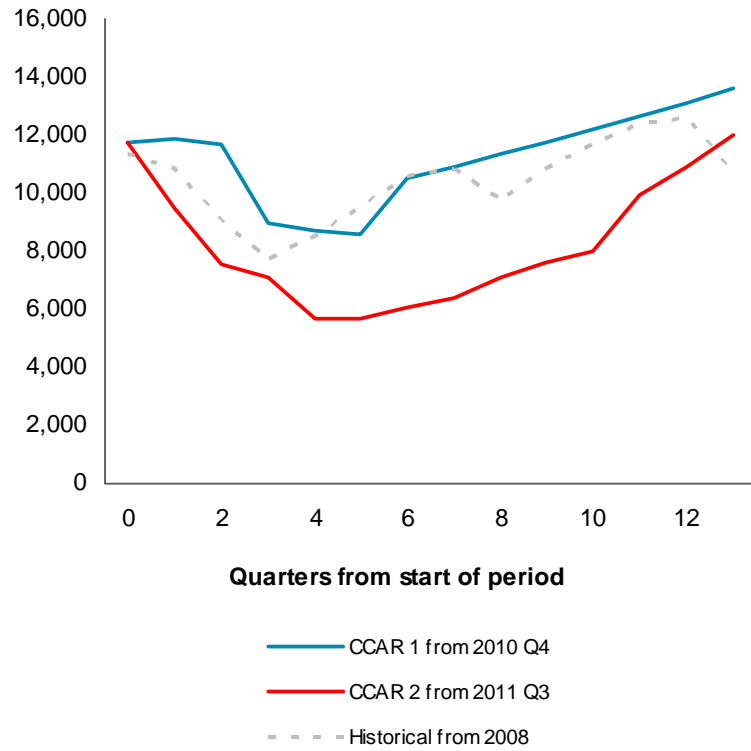
Stress-test scenarios vs. recent historical observations



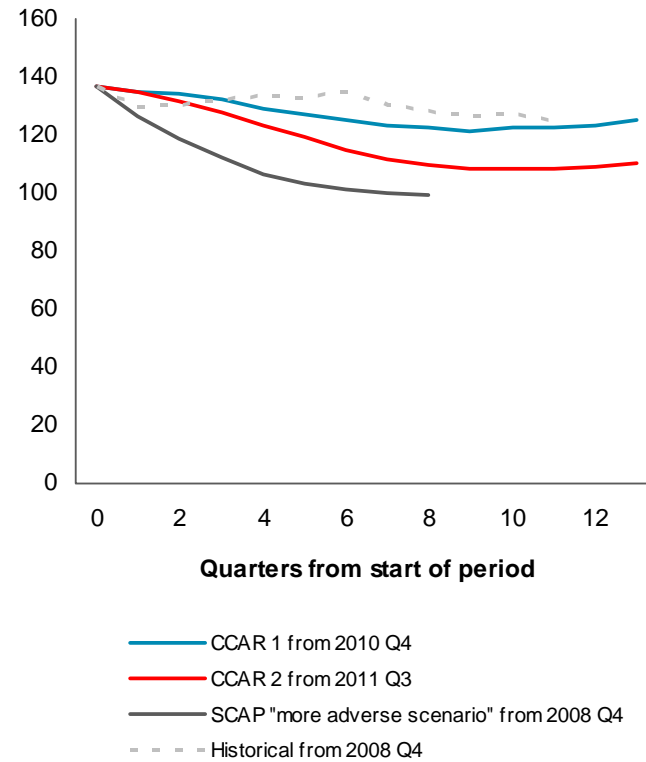
Source: Fed, The Supervisory Capital Assessment Program: Design and Implementation, 24 April 2009; Fed, Comprehensive Capital Analysis and Review: Objectives and Overview, 18 March, 2011; Fed, "Comprehensive Capital Review" document and "Capital Plan review" 22 November 2011; Datastream

Figure 3: U.S. real GDP and unemployment scenarios compared

Dow Jones total stock market index level
Stress-test scenarios vs. recent historical observations



House Price Index
Stress-test scenarios vs. recent historical observations



Source: Fed, The Supervisory Capital Assessment Program: Design and Implementation, 24 April 2009; Fed, Comprehensive Capital Analysis and Review: Objectives and Overview, 18 March, 2011; Fed, "Comprehensive Capital Review" document and "Capital Plan review" 22 November 2011; Datastream

Figure 4: U.S. equity and house price indices compared