Testing Macroprudential Stress Tests: The Risk of Regulatory Risk Weights[☆]

Viral Acharya[*,a], Robert Engle[a], Diane Pierret[a,b]

[a]*NYU Stern School of Business, Volatility Institute, 44 West 4th Street, New York, NY 10012.*
[b]*Université catholique de Louvain, ISBA, 20 Voie du Roman Pays, B-1348 Louvain-La-Neuve, Belgium.*

Abstract

Macroprudential stress tests have been employed by regulators in the United States and Europe to assess and address the solvency condition of financial firms in adverse macroeconomic scenarios. We compare the risk assessments and outcomes implied by these stress tests with those from a simple methodology that relies on publicly available market data and forecasts the capital shortfall of financial firms in severe market-wide downturns. We find that: (i) the ranking of required capitalization of financial firms in stress tests is inadequate ex post compared to that implied by market data and by realized outcomes during a crisis; and (ii) this inadequacy of rankings arises due to the reliance on regulatory risk weights in calculating required capitalization; however, (iii) the ranking of financial firms by estimated losses in stress appears consistent with market data and realized outcomes. These findings highlight the role of regulatory discretion in stress tests, characterize its perverse manifestations, and suggest that stress tests would be more effective if capital requirements were measured differently from the current static risk-weights approach.

Key words: macroprudential regulation, stress test, systemic risk, risk-weighted assets.
JEL: G28, G21, G11, G01.

[☆]Prepared for the Carnegie Rochester Conference on Public Policy, November 15-16, 2013. The authors are grateful to Deborah Lucas and conference participants for helpful discussions. The authors would also like to thank seminar participants at the Federal Reserve Bank of New York (Research, Supervision), De Nederlandsche Bank, Federal Reserve Board of Governors, Banque de France, Office of Financial Research, NYU Stern Finance, Luxembourg School of Finance, PRMIA, Risk Dynamics, Federal Reserve of Cleveland, Duisenberg School of Finance Central Bankers' Conference and the Prudential Policy and Financial Stability team of the National Bank of Belgium and Katherine Waldock for helpful comments. The authors are grateful to supporters of the Volatility Institute of the Stern School at NYU for financial support that made this research possible. Thanks go to the Sloan Foundation, the Banque de France, the University of New South Wales, the Université de Lausanne, Deutsche Bank, BlackRock and the Michael Armellino Foundation.

[*]Corresponding author.

Email addresses: `vacharya@stern.nyu.edu` (Viral Acharya), `rengle@stern.nyu.edu` (Robert Engle), `dpierret@stern.nyu.edu` (Diane Pierret)

January 31, 2014

1. Introduction

Since the financial crisis of 2008, macroprudential stress tests have become a standard tool used by regulators to assess the resilience of financial systems. Macro stress tests have been designed to help macroprudential regulation, which essentially aims at preventing the costs of financial distress to spread to the real economy (Acharya et al. (2009, 2010a,b, 2012); Borio and Drehmann (2009); Hanson et al. (2011); Hirtle et al. (2009)). Therefore, macro stress tests focus on a group of financial institutions that, taken together, can have an impact on the economy (Borio et al. (2012)) and create systemic risk. Macroprudential regulation of these institutions reduces the probability and the cost of a financial crisis by forcing institutions to internalize their contribution to systemic risk.

However, concerns have been raised that macro stress tests do not serve the goal of macroprudential regulation as they should. Greenlaw et al. (2012) argue that macro stress tests are still microprudential in nature since they focus on the solvency of individual institutions. They also remain microprudential as they fail to characterize the 'endogenous' nature of systemic risk (Borio and Drehmann (2009); Galati and Moessner (2011)). According to Borio and Drehmann (2009), macro stress tests "risk lulling policymakers in a false sense of security" as they fail to provide real-time ex ante measurement of systemic risk.

A further shortcoming of the current approach to assessing capital requirements is that it is strongly dependent on Basel regulation defining measures (the capital ratios) of the financial performance of banks. Hanson et al. (2011) show that the capital ratios give incentives to banks to shrink their assets, which in turn leads to the amplification of financial distress. More recent concerns focus on the denominator of capital ratios (the risk-weighted assets) where multiple surveys point out the inconsistency of risk weights measurement across banks (Le Lesle and Avramova (2012); Mariathasan and Merrouche (2013); Basel Committee on Banking Supervision (2013a,b)). Haldane (2011, 2012) also raises fears over the complexity and the robustness of risk weights and calls for simpler market-based metrics.

In this paper, we compare the outcomes of macroprudential stress tests (including the most recent ones) to those from a simple methodology that relies on publicly available market data (see Acharya et al. (2010a,b, 2012); Brownlees and Engle (2011)).[1] The proposed measure (SRISK) represents the expected amount of capital an institution would need to raise in a crisis to restore a target capital ratio. This estimate of the expected capital shortfall of a financial firm in a crisis is weekly updated on the NYU Volatility Laboratory website (Vlab).[2] Vlab methodology is viewed in this paper as a benchmark to macro stress tests that addresses many of the above concerns.

The "test"[3] of macroprudential stress tests comprises three steps, namely: (i) We compare required capital shortfalls of stress tests to the market-implied capital shortfall SRISK; (ii) we consider the efficacy of regulatory risk weights in determining the allocation of required capitalization to banks; and (iii) we compare stress tests projected losses to market-implied losses and actual realized losses.

Our assessment of regulatory stress tests reveals (i) differences in required capitalization levels, and differences in the ranking of required capital shortfalls across banks. The difference in required capitalization levels reflect different severity levels that are inherent to government policies. By applying the same stress scenario and the same capital requirement rule in different states of the world, Vlab provides additional transparency and discipline to stress tests. This comparison highlights the role of discretionary rules in stress testing as well as their perverse manifestations.

Aside from the severity of a stress test, our assessment reveals (ii) a misallocation of capital shortfalls across banks due to the reliance on regulatory risk weights in determining

[1] Other surveys on macroprudential stress tests include Alfaro and Drehmann (2009); Borio et al. (2012); Greenlaw et al. (2012); Hirtle et al. (2009); Schuermann (2012).

[2] http://Vlab.stern.nyu.edu/

[3] The verb "testing" in the context of this paper has a very general meaning of comparing/assessing different outcomes.

the capital requirements in a stress test. In Europe, the milder stress applied on banks' capital makes the impact of the stress scenario only apparent on the denominator (risk-weighted assets) of capital ratios. For these stress tests, several results tend to show that if the too mild stress on the numerator (capital) may have been responsible for the low aggregate capital shortfall estimate, the stressed risk weights were responsible for the incorrect risk ranking and therefore, for the misallocation of capital shortfall across banks.

The stressed risk weights derived under Basel II appear as the main driver of lower capital ratios in Europe but these higher risk weights are shown to be uncorrelated to the actual risk of banks during a crisis. Furthermore, Basel risk standards provide no incentives for banks to diversify as risk weights ignore the subadditivity feature of portfolio risk. As a result, firms are encouraged to concentrate their entire portfolio on one asset category or exposure,[4] and the underestimation of risk weights automatically leads to excess leverage.

Our assessment further indicates that (iii) the ranking of financial firms by estimated losses in stress appears consistent with market data and realized outcomes. We find that the projected losses of stress tests and Vlab correlate well in the US. In Europe too, despite the weak stress applied to banks, the ranking of the stress test capital returns is well correlated with banks' realized returns during the European sovereign debt crisis.[5]

Overall, the findings suggest that stress tests would be more effective if capital requirements were measured differently. We argue that a risk-based capital requirement is not sufficient because (a) the increase of risk over time is not captured ex ante with risk weights, and (b) risk weights are flawed cross-sectionally as banks game their risk-weighted assets, i.e. take advantage of (a) to reduce their capital requirements. The paper therefore recommends the use of backtop measures next to the risk-based capital shortfalls of stress tests in order

[4]Empirical evidence for European banks concentrating on sovereign debt exposures can be found in Acharya and Steffen (2013).

[5]However, the too mild stress produces positive returns estimates for some banks and, amplified by the size, the correlation between the stress test net losses and the realized crisis losses becomes negative.

to detect the most important vulnerabilities of banks under stress.

The rest of the paper is structured as follows: Section 2 introduces macroprudential stress tests and Vlab. We "test" the stressed capital shortfalls in Section 3 and the efficacy of regulatory risk weights in Section 4. The stress tests losses are assessed in Section 5.

2. Stress tests and Vlab

2.1. Macroprudential stress tests sample

We consider stress tests conducted on a US and EU-wide level. These stress tests can be qualified as macroprudential stress tests as opposed to microprudential stress tests conducted on a bank-level as a requirement under the Pillar II of Basel II (Internal Capital Adequacy Assessment Process (ICAAP)). Other macroprudential stress tests, not discussed here, were undertaken by national authorities (e.g. Ireland, UK, Spain) and by the International Monetary Fund.

The Board of Governors of the Federal Reserve has been responsible for conducting macroprudential stress tests in the US. A first stress test exercise called the Supervisory Capital Assessment Program (SCAP) was launched in 2009 as a response to the 2008 financial crisis. With the Dodd-Frank Act of 2010, an annual supervisory stress test of the US financial system became a requirement, and the Fed's capital plans rule of 2011 required all US bank holding companies (BHC) with consolidated assets of $50 billion or more to develop and submit capital plans to the Federal Reserve on an annual basis. As a result, the Federal Reserve conducted stress tests as part of the Comprehensive Capital Analysis and Review (CCAR) in 2011, 2012 and 2013.

EU-wide stress tests were initiated by the Committee of European Banking Supervisors (CEBS) in 2009 and 2010. The CEBS became the European Banking Authority (EBA) on January 1, 2011, which coordinated a new stress test the same year. As opposed to US stress tests by the Federal Reserve, European stress tests are conducted in a bottom-up fashion:

banks submit their stress test results to national supervisory authorities (NSAs) for review before NSAs submit to the EBA. For this reason, the EBA qualifies the EU-wide stress test exercise as a microprudential stress test. These stress tests are however the outcome of a global macroeconomic scenario defined by the European Central Bank (ECB) and share the objective of an overall assessment of systemic risk in the EU financial system.

The disclosure level of a stress test outcome is a strategic decision of the regulator that is well discussed in Goldstein and Sapra (2012), Petrella and Resti (2013) and Schuermann (2012). Some stress tests only disclose the stress scenario; other stress tests disclose an aggregate outcome of the stress scenario. On the opposite, the EBA 2011 stress test has an unprecedented level of transparency for the EU banking system. The amount of information at the bank level is very high and downloadable in an excel file from the EBA website.[6] To our knowledge, only three US and two EU-wide macroprudential stress tests publicly disclosed a bank-level outcome of the stress test exercise; the SCAP 2009, the CCAR 2012 and the CCAR 2013 in the US; the CEBS 2010 and the EBA 2011 in the European Union. These five macroprudential stress tests with bank level disclosure are the sample of stress tests of this paper. Their outcomes are summarized in Appendix A.

2.2. An alternative to stress tests: Vlab

Next to stress tests conducted by US and European regulators, a team of researchers at NYU Stern School of Business developed an alternative methodology to measure the systemic risk of financial institutions purely based on publicly available information (see Acharya et al. (2010a, 2012); Brownlees and Engle (2011)). An important breakthrough of this methodology is that systemic risk does not come from the failure of a firm as such but comes from the firm's failure when the whole financial system is undercapitalized. If a firm fails in isolation, other financial firms will step in and take over its activities. However, in

[6] http://www.eba.europa.eu/EU-wide-stress-testing/2011/2011-EU-wide-stress-test-results.aspx

a period of aggregate stress where the whole financial sector is undercapitalized, financial firms cannot find the resources to take over other firms' activities and failing firms impose negative externalities to the real economy.

In Acharya et al. (2012), the real systemic risk of a firm is defined as *the real social costs of a crisis per dollar of capital shortage* $\times$ *Probability of a crisis* $\times$ *Expected capital shortfall of the firm in a crisis*, where the last term is presented as a useful tool or a substitute for stress tests. Brownlees and Engle (2011) describe a method to derive the expected capital shortfall of a firm in a crisis (called SRISK) based on its size, its market leverage and its stock return under aggregate stress (called Long-Run Marginal Expected Shortfall or LRMES). The return of the firm in a crisis is estimated from a bivariate daily time series model, where volatilities are asymmetric GARCH processes and correlations follow a Dynamic Conditional Correlation (DCC) model. The six-month returns of the firm and the market index are simulated many times based on the estimated dynamic volatilities and correlations, and sampling from a joint distribution that allows for further dependence in the tails. LRMES is the average of the firm's returns across the simulation paths where the market return falls by 40% over a six-month time window.[7]

Defining MV as today's market capitalization of a firm, $LRMES * MV$ is the expected market cap loss that equity holders would face during the six-month crisis scenario described above. The capital shortfall of a firm i at time t ($SRISK_{it}$) is then derived assuming the book value of its debt (D_{it}) stays unchanged over the six-month scenario while its market cap falls by $LRMES_{it} * MV_{it}$

[7]The equity market return is the S&P500 for US banks, the MSCI ACWI World ETF index for European banks. Note that for European banks, the long run simulation is not yet implemented and LRMES is approximated by 1-exp(-18*MES) where MES is the expected daily return of the bank if the daily market return is less than -2%.

$$
\begin{aligned}
SRISK_{it} &= \mathrm{E}_t \left[k(Debt_{it+h} + MV_{it+h}) - MV_{it+h} | R_{mt+h} \leq -40\%\right] \\
&= kDebt_{it} – (1-k)(1 - LRMES_{it}) * MV_{it}
\end{aligned} \tag{1}
$$

where k is the prudential capital ratio, and h is the crisis scenario horizon (six months). The results of this methodology are available on the Volatility Laboratory website (Vlab), where systemic risk rankings are updated weekly both globally and in the United States (see http://Vlab.stern.nyu.edu/).

Vlab uses a prudential capital ratio k of 8% for US banks and a milder k of 5.5% for European banks to account for the difference in market leverage due to different accounting standards in the two regions: EU banks report under IFRS whereas US banks report under US GAAPs. Under US GAAPs, banks are allowed to report their derivatives on a net basis. The netting of derivatives is most of the time not allowed under IFRS norms leading to a substantial increase of the size of the balance sheet. Engle et al. (2012) indicate that the total assets of large US banks would be between 40% and 60% larger under IFRS norms.

As the stress is on the market value of equity, Vlab methodology can be viewed as a mark-to-market stress test. A market-based capital requirement would be difficult to implement in a regulatory context given the (excessive) volatility of market prices and its procyclicality implying higher capital requirements in a downturn. The higher requirements in a crisis have the potential of worsening the credit crisis where banks cannot raise equity and have to sell even more of their assets to restore their capital ratios. This observation makes SRISK only an adequate *ex ante* (before a crisis) measure of the capital shortfall of a firm.

Vlab does not have the information granularity of the supervisory data of regulatory stress tests but the use of publicly available market data allows for real-time forward looking measures. Moreover, the simplicity of the Vlab scenario compared to the complex multi-factor scenarios of stress tests makes Vlab outcomes robust to various economic environments.

Vlab is therefore viewed as a macroprudential benchmark that regulators may be interested to use in the assessment of their own stress tests outcomes. Differences in scenarios and data of Vlab and regulatory stress tests are further discussed in Appendix B.

In the next sections, we compare the outcome of stress tests to this market-implied benchmark and banks' real outcomes during the European sovereign debt crisis. We first compare the capital requirements implied by stress tests with SRISK in Section 3. Second, we compare the regulatory risk weights of stress tests and the market risk weight of Vlab with realized measures of risk in Section 4. Then, a third assessment compares the projected losses of stress tests against the market capitalization loss of Vlab and actual realized losses in Section 5. This third "test" allows us to disentangle the effect of different stress scenarios that directly impact projected losses from the impact of different risk measures.

3. Testing stressed capital ratios and shortfalls

Capital ratios are the most important output of stress tests. They determine which bank failed the test under the stress scenario and the resulting supervisory measures or recapitalization plans. In US Dodd-Frank-Act stress tests, banks have to pass regulatory thresholds on four ratios each quarter of the stress scenario: a 5% Tier 1 Common Capital Ratio (T1CR), 4% Tier 1 Capital Ratio (T1R), 8% Total Risk-Based Capital Ratio (Total CR) and a bank-specific 3 or 4%[8] Tier 1 Leverage Ratio (T1 LVGR).[9] The only ratio to be passed in the Euopean stress test of 2011 is a 5% Core Tier 1 Capital Ratio, which is considered equivalent to the US definition of the Tier 1 Common capital ratio.

The numerators of ratios are different qualities of capital based on Basel requirements:

[8]"3 percent only for a BHC with a composite supervisory rating of "1" or that is subject to the Federal Reserve Board's market-risk rule." (Board of Governors of the Federal Reserve (2013a))

[9]The disclosed ratios are actual ratios before the stress scenario (actual), stressed ratios at the end of the stress scenario (projected) assuming all capital actions, and minimum ratios over the 9 quarters of the stress scenario (min) assuming all capital actions or assuming no capital actions.

the Tier 1 common or core capital, the Tier 1 capital, and the Total Risk-Based Capital. Tier 1 common capital (US) and Core Tier 1 capital (EU) are the closest to common shareholders equity. The Total Risk-Based Capital is the addition of Tier 1, Tier 2, and possibly additional Tier 3 capital to cover market risk activities. As for the denominator, the Tier 1 leverage ratio is the only ratio defined in terms of total assets (average of total assets over the last quarters) while the denominators of all other ratios are the risk-weighted assets.

We show the aggregate output of common banks between stress tests and Vlab in Table 1 (all converted in Euros for European banks)[10]. All the banks in US and EU stress tests are not available in Vlab mainly because some banks in the stress tests samples are not publicly traded. Vlab reports the results of 18 of the 19 US banks (all except Ally Financial Inc.) and close to 60% of the banks in European stress tests.

3.1. Responding to outcomes: the goals of stress tests

Stress tests can have different goals depending on their economic and political context: the US stress test of 2009 is a good example of a crisis management tool as opposed to pre-emptive, forward-looking stress tests that followed. The SCAP 2009 led to a substantial recapitalization of the US financial system by forcing 10 bank holding companies to raise a 75 $ bn capital buffer. This objective of recapitalizing the US financial sector and that the additional capital buffer would be made available from the government was clear from its early announcements in February 2009.[11]

The European stress test disclosed in July 2011 also intended to serve as a confidence-building tool during the European sovereign debt crisis. However, the plans for failing firms were less clear. In March 2011, the EBA announced that it would be working with

[10] Data source for stock prices and exchange rates: Datastream. Data source for market caps: Vlab. To compare Vlab and EBA forecasting performance, Vlab results were downloaded on the closest date (June 30, 2011) prior the disclosure of the EBA results in July 2011.

[11] See joint statement by the Treasury, FDIC, OCC, OTS, and the Federal Reserve, February 23, 2009 (available at http://www.federalreserve.gov/newsevents/press/bcreg/20090223a.htm).

national authorities on remedial backstop measures for firms failing the stress scenario but never mentioned capital injections. Without appropriate recapitalization plans for the failing banks, a severe stress test could not be implemented to the banks and the goal of restoring confidence in the soundness of banks balance sheets was somehow missed. Evidence of this is the launch by the EBA in early December 2011 of a separate recapitalization plan of the European financial sector called the "Capital Exercise". The Capital Exercise is not a stress test (no stress scenario) but recommended the creation of an "exceptional and temporary capital buffer to address current market concerns over sovereign risk and other residual credit risk related to the current difficult market environment". The estimated capital buffer of 115 EUR bn (including 30 EUR bn for Greek banks)[12] was well above the 2.5 EUR bn estimate of the stress test disclosed five months earlier.

In the US, the function of annual Dodd-Frank Act stress tests is targeted at sanctioning capital distribution plans of poorly performing firms. The Federal Reserve uses its objection to banks' capital plans to force them to improve on some detected deficiencies during the stress test. A potential unintended consequence of applying a similar specific scenario and methodology annually on banks is the risk of banks specializing on a particular stress test. Banks adjust their portfolios to appear less risky to one specific stress scenario and methodology, but this does not necessarily make them more robust to the next crisis (which could be very different from the stress test scenario). Ideally, the stress test should reflect the outcomes of multiple stress scenarios. An additional backstop is also provided by comparing the stress test risk assessments to the Vlab outcomes where the Vlab scenario (a 40% drop in a broad market index) encompasses a wider range of specific scenarios.

The Vlab methodology provides a benchmark to these stress tests by applying the same

[12]Greek banks are treated separately in the EBA capital exercise where their capital buffers are defined in order not to conflict with pre-agreed arrangements under the EU/IMF program (European Banking Authority (2011c)).

scenario and the same requirement rule in different states of the world. The independence of Vlab from government policies gives an additional discipline to banks and makes it possible to identify stress tests differences in terms of severity and rankings as well as their perverse manifestations.

3.2. Stress tests capital shortfalls vs. SRISK: the European case

This paper does not comment on the optimal level of severity of a stress test, it is however worth noting the contrast in severity between Vlab and stress tests estimates in Table 1. Vlab is more severe both in terms of the stress level applied to banks (visible on losses and ratios) and the capital requirement rule k (visible on capital shortfalls). The discrepancy is even larger in Europe where the capital shortfall estimates of the last two stress tests (resp. 0.2 EUR bn in 2010 and 1.2 EUR bn in 2011) appear extremely low compared to the corresponding SRISK (796 EUR bn and 886 EUR bn, respectively), and reflect the earlier discussion about the different goals of European stress tests. Most European banks have a zero capital shortfall in the stress test (see Figure 1a) as the disclosed shortfall is

$$\text{Disclosed Capital Shortfall} = \max(0, [k' * RWA_S - Capital_S]), \tag{2}$$

where k' is the prudential capital ratio threshold used in the stress test (5% in the EBA 2011), RWA_S and $Capital_S$ are respectively the risk-weighted assets and the capital level of a bank at the end of the stress scenario.

The rank correlation between SRISK and the estimated capital shortfalls of stress tests are reported in Table 2 (panel A). Most European banks actually end up with a capital excess at the end of the stress scenario when we remove the zero bound and derive the 'absolute' capital shortfall ($k' * RWA_S - Capital_S$). Banks with the highest estimated capital shortfall in Vlab are considered to be the safest and the most well capitalized in European stress tests. Therefore, the rank correlation with SRISK is highly negative, significant and is almost the same in the last two European stress tests (-0.791 in 2010 and -0.790 in 2011).

We show this result in Figure 2a for the 2011 EBA stress test and obtain a similar pattern for the stress test of 2010. The heterogeneity in size in the sample of European banks however plays a major role in this result. We may not want to completely remove the impact of the size[13] from the analysis of capital shortfalls as size is a major factor contributing to the systemic importance of a bank. Size, by amplifying correlations, also shows how important discretionary rules on the final outcomes are.

To attenuate the size effect, we also look at correlations on the subsamples of (very) large banks (with Core Tier 1 capital over 19 $ bn) and small banks. The 15 large banks include HSBC, Barclays, BNP Paribas, Deutsche Bank, etc. and are comparable to the 19 US participating bank holding companies. The negative correlation of the stress test capital shortfalls with SRISK is indeed very sensitive to size; the correlation decreases for small banks (-0.53 in the EBA 2011 stress test) and is not significant in the group of large banks.

It is hard to believe that some of the banks in the EBA stress test were so well capitalized when the stress test was disclosed in July 2011. For example, Dexia appeared with a capital excess of 7.9 EUR bn in the stress test but was the first bank to be bailed out in the context of the European sovereign crisis in October 2011. The resolution plan of Dexia included the sale of its Belgian banking subsidiary for 4 EUR bn to the Belgian government, the sale of some of its assets and businesses, and guarantees of up to 90 EUR bn provided by Belgium, France and Luxembourg.[14] The resolution plan led to a net loss of 11.6 EUR bn for 2011 due to the discounted sales of assets.

Another piece of evidence of the weak 2011 stress test are the new capital shortfall estimates disclosed in the EBA Capital Exercise in December 2011. The recommended capital buffer (the "Overall Shortfall") is defined by

[13] This is done in the analysis of ratios in Section 3.3.

[14] *Dexia agrees to Belgian bail-out*, Financial Times, October 10, 2011.

$$\text{EBA Overall Shortfall} = \max(0, [0.09 * RWA - T1C]) + BuffSOV. \tag{3}$$

The overall shortfall is not the outcome of a stress test but is the result of three main drivers: the target 9% Tier 1 core capital ratio (instead of 5%), the application of Basel 2.5 to derive risk-weighted assets (increasing the capital requirement for market risk), and an additional capital buffer ($BuffSOV \geq 0$) for Eurozone sovereign debt exposures (one third of the buffer).[15] The rank correlation of SRISK with the EBA overall shortfall is positive (0.133) but not significant at 5%. The exercise corrected for the underestimated sovereign risk weights with the additional sovereign buffer but many SRISK-top banks like Crédit Agricole still end up with a capital shortfall of zero in the Capital Exercise (see Figure 1b). We may also argue that the estimates of the exercise arrived too late in December 2011 as many banks were already in deep financial trouble. Dexia, with 6.3 EUR bn shortfall in the exercise, was bailed out a second time for 5.5 EUR bn by French and Belgian governments in November 2012 and reported a net loss of 2.9 EUR bn for 2012.[16] Crédit Agricole, with no capital shortfall in the exercise, announced a net loss of 6.5 EUR bn for 2012.[17]

3.3. Stress tests vs. Vlab ratio

To facilitate the comparison with stress test ratios, we define the Vlab market leverage ratio under stress (M-LVGR$_s$) as the ratio of market cap to quasi-market assets under the Vlab stress scenario

$$\text{Vlab M-LVGR}_s = \frac{MV(1 - LRMES)}{MV(1 - LRMES) + D}. \tag{4}$$

The rank correlations in panel B of Table 2 reveal two important components driving the difference between stress test and Vlab rankings: capital actions and risk-weighted assets.

[15] European Banking Authority (2011c)

[16] *Fresh Franco-Belgian bailout for Dexia*, Financial Times, November 8, 2012. *Dexia at 'turning point' amid more losses*, Financial Times, February 21, 2013.

[17] *Second year in red for Crédit Agricole*, Financial Times, February 20, 2013.

The aggregate outcome of stress tests with and without the effect of capital actions is presented in Appendix A. The impact of capital actions on ratios is negative in the CCAR since capital actions are capital distribution plans (submitted as part of the CCAR). Conversely, capital actions are capital raising plans in the SCAP and the EBA and have a positive impact on stress tests outcomes.[18] For all stress tests, rank correlations with Vlab measures increase when capital actions are ignored. We show in the next section, that rank correlations also increase substantially when risk-weighted assets in stress tests measures are replaced by total assets.

3.4. EU-US differences

The ultimate impact of the stress scenario is a decline in capital ratios under both US and EU stress tests. However, this result comes from a diminution of capital (the numerator) in the US whereas increasing risk-weighted assets (the denominator) is the main driver of lower capital ratios in Europe (see Tables 5 and 6).

The difference in stress levels between US and EU stress tests (further discussed in Appendix B) is consistent with different stress scenario trajectories and assumptions on the balance sheet growth. European stress tests have a static balance sheet evolution assumption whereas the CCAR assumes that the size of the balance sheet can change according to economic conditions. US stress scenarios tend to revert to a 'normal state' of the world at the end of each scenario, unlike European scenarios which assume further deterioration of the economic situation the second year of the stress scenario. This is the reason why the Federal

[18]Capital actions in the CCAR 2012 include all proposed future capital distribution plans (issuance of capital instruments, dividends payments and share repurchases) throughout the stress scenario. In the EBA 2011, capital actions include issuance of common equity, government injections of capital and conversion of lower-quality capital instrument into Core Tier 1 capital. The EBA additionally considers the effect of mandatory restructuring plans and the final outcomes only consider mandatory measures announced before disclosure. In the SCAP, the capital actions include the proposed capital actions and the effects of the results of the first quarter of 2009. The correlation between SRISK and the SCAP capital buffer also increases from 0.507 to 0.562 when capital actions are not included.

Reserve considers minimum ratios over the scenario horizon to determine which banks failed the stress test, while European stress tests consider ratios at the end of the stress scenario.

We will see in the next section that risk-weighted assets (RWA) definitions in US and EU stress tests are also different. Overall, different scenarios, assumptions and definitions lead to a RWA fall of 6.1% at the end of the US supervisory stress scenario of 2012 and a RWA increase of 14% under the European 2011 adverse scenario. An important implication is that a leverage ratio cannot reflect the European stress scenario as the stress mostly appears in risk weights and total assets are assumed constant over the scenario. Stressed leverage ratios also decline less in the US than stressed risk-based ratios. The impact of the projected total assets and RWA changes on US ratios is not straightforward according to the Federal Reserve however.

4. Testing the efficacy of regulatory risk weights

A major difference between Vlab and stress test measures comes from the denominator of regulatory ratios. Regulatory ratios and shortfalls are expressed as a function of risk-weighted assets whereas Vlab uses quasi-market assets. While the forbearance of the stress tests outcomes could be addressed by a more severe scenario, stricter modeling assumptions or simply a larger threshold k as in the EBA Capital Exercise, we show that the misallocation of capital shortfall across banks is caused by risk weights and call for additional measures of risk in a macro stress test.

4.1. Basel I and Basel II

Risk-weighted assets (RWA) definitions are not the same in US and European stress tests; RWA are derived under Basel I in the US (before 2013) and under Basel II in the EU. This leads to important differences in risk measures and stress test models. Risk weights are fixed for different asset categories under Basel I whereas banks can use their own models to derive RWA under Basel II.

Under Basel I, RWA are defined such that assets are assigned to four different asset categories with different static risk weights (0%, 20%, 50%, 100%). These four categories could be roughly described as exposures to sovereigns (0%), banks (20%), mortgages (50%), and corporates (100%). In the SCAP, the rank correlation of SRISK was indeed the highest with the most risky asset category. But Basel I risk weights cannot reflect the risk evolution of different asset categories; in 2011, SRISK was more correlated with the zero-weight risk category (including sovereign debt exposures) than with the most risky 100%-weight category.

The problem of static risk weights is addressed in Basel II where the capital requirement for credit risk (around 80% of RWA) is defined in terms of exposures at default (EAD) and risk parameters. Risk parameters (probability of default and loss given default) are used to assign weights to each exposure. In the EBA 2011 stress test, the increase of RWA under the stress scenario comes from the credit risk component; the changes are located in risk-weights (stressed LGDs and PDs) since exposures are considered invariant under the static balance sheet assumption. This is a major difference with the US methodology that considers fixed risk weights, even if credit rating migrations are allowed.[19]

Concerns on the robustness of Basel II risk weights have been raised in Haldane (2012), given their degree of over-parametrization and the risk parameter estimates purely based on in-sample statistical fit over short historical samples. The use of banks' internal models to derive their risk parameters under the Internal-Rating-Based (IRB) approach of Basel II has also been criticized. First, Basel II was designed so that the use of banks' internal models would allow them to derive lower RWA in order to incentivize banks to update their risk management practices. Le Lesle and Avramova (2012) indicate that this resulted in lower RWA under Basel II, and therefore lower capital charges than under Basel I whereas the internal

[19]The RWA methodology was however updated in the CCAR 2013 where the stressed RWA also included BHCs' projections of a market risk component defined under the stricter Basel 2.5 market risk rule.

models did not necessarily imply lower risks. Second, concerns about the consistency of risk weights across firms are raised in Haldane (2011, 2012); Le Lesle and Avramova (2012); Mariathasan and Merrouche (2013); Basel Committee on Banking Supervision (2013a,b); European Banking Authority (2013). The Basel Committee confirmed these concerns showing in their "Regulatory Consistency Assessment Programme" (RCAP) that differences in risk weights (in the trading book) across firms reflect modeling choices and supervisory decisions rather than actual risk taking.[20] Furthermore, Mariathasan and Merrouche (2013) attribute the decline in risk weights when banks switch to the IRB approach to strategic risk modeling, and that effect to be particularly important for weakly capitalized banks. Third, the internal models used to derive risk weights are completely opaque. Haldane (2012) indicates that risk weights are black boxes that investors do not understand or trust. These concerns have important implications for the European stress tests outcomes knowing that 59 of the 90 participating banks in the 2011 stress test are IRB banks, i.e. they use their own models to derive risk weights under the stress scenario.

4.2. Risk-based capital vs. leverage

The assessment of bank leverage using a Tier 1 leverage ratio (T1 LVGR) defined as the ratio of Tier 1 capital to total assets is a recommendation of Basel III to supplement the risk-based regime (Basel Committee on Banking Supervision (2011)). Haldane (2012) shows that this ratio significantly predicts the failure of financial firms whereas the risk-based core Tier 1 capital ratio (T1CR) does not.

The Tier 1 leverage ratio is one of the four ratios examined in Dodd-Frank Act stress tests. In 2012, two banks (Citigroup and MetLife) failed the leverage ratio under the stress scenario. In 2013, Goldman Sachs had the lowest stressed leverage ratio followed by Morgan

[20] The RCAP of the banking book disclosed in July 2013 however indicates that three quarters of differences in banking book risk weights across banks are explained by differences in banks' exposures (Basel Committee on Banking Supervision (2013a)).

Stanley and JP Morgan, and two firms (Ally Financial and American Express) failed the leverage ratio under stress when considering the effect of their original submissions of planned capital actions. We build a Tier 1 leverage ratio for the European banks of the 2011 stress test and find that Deutsche Bank would have failed the stress test if the Basel III 3% leverage backstop existed. In Figure 3, the visual correlation between the market leverage ratio under the Vlab stress (M-LVGR$_s$) and the stressed Tier 1 leverage ratios appears to be strong in the last US and European stress tests (CCAR 2013 and EBA 2011). This result is confirmed in Table 2 (panel B): the rank correlation with the Vlab ratio increases from 0.581 to 0.877 when risk-weighted assets, the denominator of capital ratios, are replaced by total assets in the CCAR 2013. We obtain similar results one year earlier (CCAR 2012), and in the European stress test of 2011.

The contrast between risk-based and leverage-based stress tests outcomes is even more obvious when the stress test capital shortfall is written as a function of total assets. We show in Table 2 (panel A) and Figure 2b that the correlation between SRISK and the capital shortfall of the EBA stress test becomes highly positive (0.679) and significant when the EBA shortfall is written as a function of total assets instead of risk-weighted assets. The leverage-based capital shortfall is given by

$$\text{Capital Shortfall (TA)} = k * TA_S - Capital_S, \tag{5}$$

where k is the same prudential ratio used in Vlab (5.5% for European banks), and TA_S are the total assets of the bank at the end of the stress scenario. The result holds when we control for the size; the rank correlation between (5) and SRISK remains high and significant at 1% in the groups of small (0.634) and large banks (0.743). With this definition, the required capitalization of 53 EU banks would have increased from 1.2 EUR bn to 390 EUR bn.

If instead of using a leverage-based shortfall, the capital requirement rule of the stress test (k') in eq. (2) had been increased from 5% to 9% of RWA (as in the Capital Exercise), the total risk-based capital shortfall would have been 139 EUR bn for the same 53 banks. The correlation of the 9% risk-based shortfall with SRISK is positive (0.418) and significant at 1% but this result is not robust when controlling for size. More importantly, many banks like Dexia would still end up with an estimated capital excess with this definition while having a positive capital shortfall with the leverage-based definition of eq. (5).[21]

4.3. Stress tests vs. Vlab risk weight

To complement the view on the bank risk, Acharya et al. (2012) establish a connection between Vlab estimates and regulatory risk-weighted assets by defining the effective aggregate market risk weight to quasi-market assets corresponding to a SRISK of zero. They show that imposing a financial firm to be adequately capitalized in a crisis implies that its current market capitalization is above a fraction k of some 'market risk-weighted' assets

$$MV \geq \frac{k}{1-(1-k)LRMES}(MV + Debt) \tag{6}$$

Therefore, the Vlab aggregate risk weight of the firm is

$$\text{Vlab risk weight} = (1-(1-k)*LRMES)^{-1} \tag{7}$$

and is comparable to the aggregate regulatory risk weight defined by the ratio of RWA to total assets. The risk weight could be interpreted as an aggregate measure of risk per unit of asset; the smaller the risk weight, the less risky the asset holdings of a bank.

As the Vlab risk weight is conditional on a crisis, we compare it to the stressed aggregate risk weights of stress tests. Figure 4a compares the projected risk weight at the end of the EBA stress scenario with the Vlab risk weight. These measures of risk have nothing in common; the rank correlation is negative (-0.238) and not significant at 5%. Dexia and

[21] Dexia would have been forced to raise 9.5 EUR bn and Crédit Agricole 35.8 EUR bn in July 2011 based on the leverage-based capital shortfall of eq. (5). With the 9% risk-based capital shortfall definition, the estimated capital excess of Dexia would have been 2 EUR bn and the estimated capital shortfall of Crédit Agricole 3 EUR bn.

Crédit Agricole are among the riskiest banks according to the Vlab measure and among the safest with the stress test risk weight; both banks have values above the 75% quantile of the Vlab risk weight distribution while both appear below the 25% quantile of the EBA risk weight distribution.

In the US, the Vlab risk weight also appears uncorrelated with some approximation of the stressed risk weight of the 2009 stress test; the rank correlation is slightly negative (-0.011) and not significant at 5%.

4.4. Forecasting risk during the European sovereign debt crisis

This section empirically validates the risk measures of stress tests and Vlab as aggregate measures of bank risk during a crisis. To that end, we compare the performance of these measures to predict a realized measure of risk; the six-month realized volatility defined by

$$RV_{i,t,W} = \sqrt{\frac{1}{W}\sum_{t+1}^{t+1+W}(r_{it} - \bar{r}_{it,W})^2} \tag{8}$$

where $W = 130$ days (six months), $\bar{r}_{it,W}$ is the six-month forward average stock return of bank i at date t (the disclosure date of the stress test).

It is important to note that stress tests outcomes are not forecasts; they are estimates of the bank performance conditional on a specific adverse macro-economic scenario. However, if the goal of a macro stress test is to make banks more robust to aggregate stress conditions, we would expect that stress tests outcomes identify banks vulnerabilities when there is realized aggregate stress. In other words, comparing stress tests outcomes to realized outcomes in a crisis can help determining whether the stress test scenario was credible as well as other deficiencies of the stress test that would prevent it from detecting banks most obvious vulnerabilities.

We focus on the EBA stress test disclosed on July 15, 2011 as it is the only stress test with bank-level disclosure followed by a global economic downturn. The realized returns in the last six months of 2011 of US (S&P500), European (Eurostoxx50) and global (MSCI ACWI World) indices were respectively -4.89%, -20.67%, and -13.47%. This outcome was less severe than the Vlab scenario (40% decline in the World equity index) and is closer to the ECB scenario (15% decline in stock prices in the Euro area).

A striking result appears in Table 3 (panel A) where only the Vlab risk weight predicts the ranking of banks' realized risk. The rank correlation between the realized volatility and the stress test risk weight is negative and insignificant. Similarly, Das and Sy (2012) find that risk-weighted assets do not, in general, predict market measures of risk. The absence of correlation between the *stressed* risk weights and the realized risk of banks during a crisis shows furthermore that Basel II risk weights are also misleading in a macro stress test.

When comparing the risk measures against realized *book* measures we find that both the Vlab risk weight and the regulatory risk weights are negatively correlated to the future book performance of banks (measured by the net income divided by total assets, and the book equity return during the last two quarters of 2011). The Vlab risk weight does not seem to predict the ranking of realized book performance in the wrong direction, in contrast with the regulatory risk weights when predicting realized market risk.

In Table 4, we show the estimates of different risk factors regressed on the realized volatility measure defined in (8). The effect of individual risk factors is reported in columns 2 to 4 where the impact of accounting-based vs. market-based risk measurement is accounted for by including the book-to-market ratio in each regression. According to the goodness-of-fit measures, the stressed Tier 1 leverage ratio is the most important factor followed by the Vlab risk weight.

The EBA stressed risk weight parameter is negative and not significant alone (column 4) but becomes positive and significant at 10% when we control for the other risk factors (column 6). This result suggests that regulatory risk weights can bring additional information on risk once we account for other more important risk factors like the Vlab risk weight, and the leverage ratio. The improvement in terms of adjusted R^2 is however small (3.76%, from columns 5 to 6) when the EBA stressed risk weight is added to the regression.

A final validation test looks at the change in risk from the disclosure of the EBA 2011 to six months after the disclosure. Specifically, we assess how the change in EBA risk weights over the stress scenario predicts the realized change in risk defined by $RV_{i,t,W} - RV_{i,t-W,W}$. We show in Table 3 (panel B) that the change in Basel II risk weights, i.e. the stress on PDs and LGDs, is well correlated to the realized risk change of small banks. We conclude that even if the EBA stressed risk weight is individually an incorrect measure of aggregate bank risk, the stress model applied on risk weights is correct. Projected changes in risk weights indeed predict the ranking of banks' risk increases during the European sovereign debt crisis.

The Vlab risk weight also predicts the change in risk, suggesting that the risk measure does not only predict a risk level but also reflects investors expectations on banks' risk evolution.

4.5. Bank risk measure and portfolio incentives

Regulatory risk weights give important information on the risks attached to the different exposures of the bank. However, even if risk weights were perfectly measured, a strong argument against risk-weighted assets as a measure of the overall bank risk comes from the observation that risk is not an additive concept.

We show in Appendix C the weakness of Basel regulatory risk weights as an aggregate measure of bank risk where the bank is seen as a portfolio of assets. The main observation is that the risk of a portfolio is always less than or equal to the sum of the risks of its components. The use of risk-weighted assets ignores this portfolio feature of risk and consequently there is no incentive from the regulatory perspective to diversify.

Then, if some risk weights are underestimated or are not adjusted to reflect increased risk during a crisis, a bank will choose its optimal asset with the most underestimated risk weight and this will automatically lead to excessive leverage. Consequently, banks will take excessive leverage if their risk weights are not adequately adjusted to more severe economic conditions.

The forward-looking stress applied to Basel II risk weights should correct for the problem of static risk weights. We however show that the mechanical relationship between risk weights and leverage also holds empirically after the application of the stress scenario in the European stress test of 2011 (Figure 4b). The rank correlation between stressed risk weights and stressed Tier 1 leverage is 0.62 for 53 European banks and increases to 0.89 for the 15 largest banks.

The incentives created by risk weights also explain the portfolio decisions of many Eurozone banks during the European sovereign debt crisis. Acharya and Steffen (2013) document that the increase of exposures to risky sovereign debt is partly explained by regulatory arbitrage; banks with higher risk weights increased their exposures to risky sovereign debt to reduce the cost of raising fresh capital, as these exposures have a zero capital requirement (zero-risk weight). To a large extent, it also explains the misguidance of stress tests about European banks risks. For example, Dexia was holding a portfolio of risky sovereign bonds of almost a third of its balance sheet largely financed with short-term debt. Acharya and Steffen (2013) further show that this type of behaviour was largly pervasive among Eurozone banks. Therefore, the reliance on Basel risk weights appears not only to have misguided the recapitalization of the financial sector but also incentivized the build up of risky sovereign debt exposures.

5. Testing stressed losses

Stress test models translate the stress scenario into a bank outcome. The most direct impact of the stress scenario should be visible on banks' projected losses. We show that, despite different scenarios, data and models, Vlab losses and US stress test losses correlate well. In Europe too, the projected capital returns in the stress test correlate well with realized returns during the European sovereign debt crisis.

5.1. Stress tests vs. Vlab losses

We compare the projected losses of stress tests with Vlab's market cap loss ($MV * LRMES$) in Table 1. From this table, we notice the important gap between the "Loss" and the "Net Loss" of stress tests (especially in Europe) due to the effect of projected revenues under the stress scenario. The net loss is the main driver of capital diminution under stress and is the accounting equivalent of the Vlab loss. Vlab appears again more severe; the order of amplitude of Vlab losses is similar to the amplitude of 'pure' losses of stress tests that do not include the stressed revenues. Moreover, the Vlab loss is a six-month loss whereas stress tests losses are projected over two years.

The rank correlations of the Vlab loss with the total losses of stress tests are very high and significant in all stress tests (see Table 2, panel C). Huang et al. (2012) do the same test for the SCAP and find that stress test losses are well correlated to several market-based measures of systemic risk;[22] our rank correlation estimate of 0.68 is close to their estimate of 0.67 for the Marginal Expected Shortfall (MES) of Acharya et al. (2010b).[23] We also report the correlations of the Vlab loss with the stress tests loan losses and trading losses since they are the most important sources of losses (85%) according to the CCAR 2012. The correlations of the Vlab loss with the loan and trading losses are also very high and significant, making Vlab's ranking and the ranking of losses under supervisory stress scenarios very consistent.

The correlations of the Vlab loss with the total *net* loss (including stressed revenues) are

[22]The Marginal Expected Shortfall (MES) of Acharya et al. (2010b), the CoVaR of Adrian and Brunnermeier (2010), and the marginal contribution to the Distress Insurance Premium (DIP) of Huang et al. (2009).

[23]Huang et al. (2012) use the marginal expected shortfall (MES) instead of the long-run counterpart LRMES, and the MES is multiplied by the Tier 1 capital instead of the market capitalization.

smaller for all stress tests and negative in Europe; banks with larger profits under European stress scenarios are predicted to have larger losses in Vlab. Some banks actually report positive profits under the stress scenario of stress tests where stressed revenues cover stressed losses.[24] The profits are then reported in the balance sheet so that the divergence with Vlab is also visible in capital changes. We show in Figure 5 that the projected profits under the EBA stress scenario lead to increasing capital levels for many banks with the largest Vlab losses. Controlling for the size effect, the correlation between the Vlab market cap return (LRMES) and the return on core Tier 1 capital over the EBA stress scenario is less important (-0.133) and not significant at 5%.

5.2. Predicting banks' real losses during the European sovereign debt crisis

Table 3 shows the performance of Vlab and the EBA 2011 stress test in predicting the actual ranking of banks' realized six-month losses and six-month returns after disclosure of the EBA stress test (i.e. the last two quarters of 2011). The six-month realized return is $-\sum_{t+1}^{t+131} \ln(p_{it}/p_{it-1})$, where p_{it} is the daily stock price of the bank, and the six-month realized loss is the product of the six-month realized return with the market cap of the bank.

For predicting realized losses (panel C), the Vlab market cap loss has the highest rank correlation (0.832) with the six-month realized loss. The correlation of the realized loss with the EBA projected total *net* loss is negative (except for large banks) since many banks with positive projected profits in the stress test actually endured the highest losses during the sovereign debt crisis.

For predicting realized returns (panel D), Vlab long-run marginal expected shortfall

[24] First, the stress scenario is not an absolute scenario as in Vlab but is defined as a deviation from a baseline scenario. If some banks are projected to make large profits in the baseline scenario, they will make lower but still positive profits under the adverse scenario. Second, the EBA explains that the stress scenario may lead to a higher net interest income where some banks assume that the impact of higher interest rates will be passed onto customers without a corresponding increase in the cost of funding for the bank. Then, the EBA considers a directional market risk stress test; depending on the direction of their exposures banks can realize trading gains on certain portfolios.

(LRMES) is a better predictor of the amplitude of realized returns according to the root mean square error (RMSE). However, the estimated core Tier 1 capital return over the EBA stress scenario better predicts the ranking of realized six-month returns than Vlab LRMES suggesting that the ranking of stressed returns in the EBA stress test was correct (after controlling for size), only the stress applied to banks was too mild compared to what happened during the following six months.

6. Concluding remarks

This paper provides an assessment of the outcomes of macroprudential stress tests conducted by US and European regulators. Concerns have been raised on the lack of robustness, severity and transparency of stress test measures. In this paper, we compare the risk assessments and outcomes of stress tests to a market-based macroprudential benchmark (Vlab) that addresses many of those concerns.

Our assessment of stress tests outcomes reveals a misallocation of capital shortfalls across banks due to the reliance on regulatory risk weights in determining required levels of capital in a stress test. The reliance on risk weights is even more important for European stress tests ratios where a mild stress on capital makes the impact of the stress scenario only apparent on the denominator (the risk-weighted assets). Therefore, the stressed risk weights appear as the main driver of the ranking of the lower capital ratios.

We show that the higher regulatory risk weights of stress tests have no link with the realized risk of banks during a crisis. For example, Dexia appeared to be one of the safest bank in the latest European stress test disclosed in July 2011, but ranked among the riskiest banks in Vlab. Three months later, Dexia was the first bank to be bailed out in the context of the European sovereign debt crisis.

Risk weights tend to be informative only when we control for a market measure of risk and the Tier 1 leverage ratio. Furthermore, Basel risk standards based on risk-weighted assets reduce the incentives for banks to diversify as they ignore the subadditivity feature of portfolio risk. Therefore, banks are encouraged to invest their entire portfolio in one asset category, and the underestimation of risk weights automatically leads to excess leverage.

Overall the results indicate that stress tests would be more effective if capital requirements were measured differently from the current static risk-weights approach. This paper recommends that regulatory stress tests complement their assessment of bank and system risks by using leverage-based and market-based measures of risk. The paper therefore welcomes the new Basel III Tier 1 leverage ratio, but the misguidance of the asset risk-return allocation is likely to be present in future stress tests as long as the reliance on static regulatory risk weights prevails under Basel III.

References

Acharya, V., Brownlees, C., Engle, R., Farazmand, F., Richardson, M., 2010a. Measuring systemic risk. In: Acharya, V., Cooley, T., Richardson, M., Walter, I. (Eds.), Regulating Wall Street: The Dodd-Frank Act and the New Architecture of Global Finance. John Wiley & Sons, Ch. 4., Ch. 4.

Acharya, V., Engle, R., Richardson, M., 2012. Capital shortfall: a new approach to rankings and regulating systemic risks. American Economic Review Papers and Proceedings 102:3, 59–64.

Acharya, V., Pedersen, L., Philippon, T., Richardson, M., 2009. Regulating systemic risk. In: Acharya, V., Richardson, M. (Eds.), Restoring Financial Stability: How to Repair a Failed System. John Wiley & Sons, Ch. 13.

Acharya, V., Pedersen, L., Philippon, T., Richardson, M., 2010b. Measuring systemic risk, Working Paper Stern School of Business.

Acharya, V., Steffen, S., 2013. The "greatest" carry trade ever?: Understanding eurozone bank risks, NYU Working Paper.

Adrian, T., Brunnermeier, M., 2010. CoVaR, Federal Reserve Bank of New York Staff Reports No. 348.

Alfaro, R. A., Drehmann, M., 2009. Macro stress tests and crises: What can we learn. BIS Quarterly Review December, 29–41.

Basel Committee on Banking Supervision, June 2011. Basel III: A global regulatory framework for more resilient banks and banking systems, Bank for International Settlements.

Basel Committee on Banking Supervision, July 2013a. Regulatory Consistency Assessment Programme (RCAP) - Analysis of risk-weighted assets for credit risk in the banking book, Bank for International Settlements.

Basel Committee on Banking Supervision, January 2013b. Regulatory Consistency Assessment programme (RCAP) - Analysis of risk-weighted assets for market risk, Bank for International Settlements.

Board of Governors of the Federal Reserve, 2009a. The Supervisory Capital Assessment Program: Design and Implementation, April 24, 2009.

Board of Governors of the Federal Reserve, 2009b. The Supervisory Capital Assessment Program: Overview of Results, May 7, 2009.

Board of Governors of the Federal Reserve, 2011. Comprehensive Capital Analysis and Review: Objectives and Overview, March 18, 2011.

Board of Governors of the Federal Reserve, 2012. Comprehensive Capital Analysis and Review 2012: Methodology and Results for Stress Scenario Projections, March 13, 2012.

Board of Governors of the Federal Reserve, 2013a. Comprehensive Capital Analysis and Review 2013: Assessment Framework and Results, March 14, 2013.

Board of Governors of the Federal Reserve, 2013b. Dodd-Franck Act Stress Test 2013: Supervisory Stress Test Methodology and Results, March 7, 2013.

Borio, C., Drehmann, M., 2009. Towards an operational framework for financial stability: "fuzzy" measurement and its consequences, Central Bank of Chile Working Paper No. 544.

Borio, C., Drehmann, M., Tsatsaronis, K., 2012. Stress-testing macro stress testing: does it live up to expectations?, BIS Working Paper No 369.

Breuer, T., Jandačka, M., Rheinberger, K., Summer, M., 2009. How to find plausible, severe, and useful stress scenarios. International Journal of Central Banking 5, 205–224.

Brownlees, C., Engle, R., 2011. Volatility, correlation and tails for systemic risk measurement, NYU Working Paper.

Das, S., Sy, A., 2012. How risky are banks' risk-weighted assets? Evidence from the financial crisis, IMF Working Paper WP/12/36.

Engle, R., Jondeau, E., Rockinger, M., 2012. Systemic risk in Europe, Swiss Finance Institute Research Paper No. 12-45.

European Banking Authority, 2010. Aggregate outcome of the 2010 EU wide stress test exercise coordinated by CEBS in cooperation with the ECB, July 23, 2010.

European Banking Authority, 2011a. 2011 EU-wide stress test: Aggregate report, July 15, 2011.

European Banking Authority, 2011b. 2011 EU-wide stress test: Methodological note, March 18, 2011.

European Banking Authority, 2011c. EBA recommendation on the creation and supervisory oversight of temporary capital buffers to restore market confidence, December 8, 2011.

European Banking Authority, 2013. Interim results update of the EBA review of the consistency of risk-weighted assets - low default portfolio analysis, August 5, 2013.

Flood, M., Korenko, G., 2013. Systematic scenario selection, Office of Financial Research Working Paper 0005.

Galati, G., Moessner, R., 2011. Macroprudential policy - a literature review, BIS Working Papers No 337.

Glasserman, P., Kang, C., Kang, W., 2012. Stress scenario selection by empirical likelihood, preprint. Available at SSRN: http://ssrn.com/abstract=2101465.

Glasserman, P., Kang, W., 2013. Design of risk weights, Columbia Business School Research Paper No. 13-39.

Goldstein, I., Sapra, H., 2012. Should banks' stress test results be disclosed? an analysis of the costs and benefits, working Paper, Booth School of Business, University of Chicago.

Greenlaw, D., Kashyap, A., Schoenholtz, K., Shin, H., 2012. Stressed out: Macroprudential principles for stress testing, Chicago Booth School of Business Working Paper No. 71.

Haldane, A., 2011. Capital discipline, Bank of England speech, January 9, 2011. URL `http://www.bankofengland.co.uk/publications/documents/speeches/2011/speech484.pdf`

Haldane, A., 2012. The dog and the frisbee, Bank of England speech, August 31, 2012. URL `http://www.bankofengland.co.uk/publications/documents/speeches/2012/speech596.pdf`

Hanson, S., Kashyap, A., Stein, J., 2011. A macroprudential approach to financial regulation. Journal of Economic Perspectives 25:1, 3–28.

Harris, T., Herz, R., Doron, N., 2012. Accounting's role in the reporting, creation, and avoidance of systemic risk in financial institutions. In: Fouque, J.-P., Langsam, J. (Eds.), The handbook of systemic risk. Cambridge University Press, available at SSRN: http://ssrn.com/abstract=1987749.

Hirtle, B., Schuermann, T., Stiroh, K., 2009. Macroprudential supervision of financial institutions: Lessons from the SCAP, Federal Reserve Bank of New York Staff Report No. 409.

Huang, X., Zhou, H., Haibin, Z., 2012. Systemic risk contributions. Journal of Financial Services Research 42, 55–83.

Huang, X., Zhou, H., Zhu, H., 2009. A framework for assessing the systemic risk of major financial institutions. Journal of Banking & Finance 33:11, 2036–2049.

Le Lesle, V., Avramova, S., 2012. Revisiting risk-weighted assets: Why do RWAs differ across countries and what can be done about it?, IMF Working Paper WP/12/90.

Mariathasan, M., Merrouche, O., 2013. The manipulation of Basel risk-weights, CEPR Discussion Paper No. DP9494.

Petrella, G., Resti, A., 2013. Supervisors as information producers: Do stress tests reduce bank opaqueness? Journal of Banking & Finance.

Schuermann, T., 2012. Stress testing banks, Wharton Financial Institutions Center Working Paper No. 12-08.

Table 1: **Vlab vs. stress tests: aggregate results.** This table presents the aggregate outcome of the samples of common banks between Vlab and regulatory stress tests. Vlab output is available on the website Vlab.stern.nyu.edu, under "NYU Stern Systemic Risk Rankings of U.S. Financials with Simulation" for US banks (where $k = 0.08$ in eq. (1)), and "NYU Stern Systemic Risk Rankings of World Financials without Simulation" for European banks (where $k = 0.055$ in eq. (1)). Vlab output is downloaded on the last date before the scenario start date of each stress test exercise. Vlab download date: 12/31/2008 (SCAP), 09/30/2011 (CCAR 2012), 09/28/2012 (CCAR 2013), 12/31/2009 (CEBS), 12/31/2010 (EBA), 09/30/2011 (EBA Capital Exercise). Vlab MV loss = $MV * LRMES$, SRISK is the market-based capital shortfall defined in eq. (1), Vlab M-LVGR$_s$ is the ratio of market cap to quasi-market assets under Vlab stress scenario (eq. (4)). Stress tests ratios (T1CR for EBA and CCAR, T1R for CEBS) are cross-sectional averages at the end of the stress scenario in EU stress tests, and cross-sectional averages of min ratios over the stress scenario in US stress tests (without the effect of BHCs planned capital actions). Stress tests losses are the sum of projected losses over the stress scenario and across banks. "Loss" (SCAP) = *Total Loss estimates* , "Loss" (CCAR) = *Loan Losses + Trading and Counterparty Losses + Realized Losses on Securities + Other Losses*, "Loss" (CEBS & EBA) = *Impairment losses + Trading losses.* "Net Loss" (SCAP) = max(0, *Total Loss estimates - Resources Other Than Capital to Absorb Losses in the More Adverse Scenario*), "Net Loss" (CCAR) = max(0, - *Projected Net Income before Taxes*), "Net Loss" (CEBS) = max(0, Loss - *pre-impairment income after the adverse scenario*), "Net Loss" (EBA) = max(0, - *Net profit after tax*). In parentheses: number of banks failing the systemic risk criterion. In brackets: cross-sectional average ratio change with the stress scenario.

US		Stress tests estimates				Vlab estimates		
	Sample	Shortfall	Ratio	Loss	Net loss	SRISK	M-LVGR$_s$	MV loss
SCAP 2009	18 US BHCs	63.1 \$ bn (9)		590 \$ bn	229 \$ bn	674 \$ bn (18)	2.39% [-4.82%]	438 \$ bn
CCAR 2012	18 US BHCs		7.55% (0) [-3.34%]	529 \$ bn	226 \$ bn	669 \$ bn (17)	3.54% [-4.54%]	447 \$ bn
CCAR 2013	17 US BHCs		8.37% (0) [-2.68%]	457 \$ bn	197 \$ bn	494 \$ bn (14)	5.48% [-5.36%]	525 \$ bn
EU								
	Sample	Shortfall	Ratio	Loss	Net loss	SRISK	M-LVGR$_s$	MV loss
CEBS 2010	50 EU banks	0.2 EUR bn	8.98% (1) [-1.38%]	425 EUR bn	39 EUR bn	796 EUR bn (48)	2.6% [-2.01%]	399 EUR bn
EBA 2011	53 EU banks	1.2 EUR bn	7.98% (4) [-1.02%]	381 EUR bn	70 EUR bn	886 EUR bn (51)	2.26% [-1.73%]	402 EUR bn
EBA Capital Exercise	44 EU banks (excluding Greek banks)	72 EUR bn (22)				1061 EUR bn (42)	1.56% [-1.66%]	336 EUR bn

Table 2: **Vlab vs. stress tests: rank correlations.** This table presents the rank correlations of stress tests and Vlab results. Panel A: rank correlations with Vlab's capital shortfall SRISK (eq. (1)). Panel B: rank correlations with Vlab M-LVGR$_s$, i.e. the ratio of market cap to quasi-market assets under the Vlab stress scenario (eq. (4)). Panel C: rank correlations with Vlab's MV loss ($LRMES * MV$). Stress tests "Total Net Loss" (i.e. net of revenues) and "Total Loss" are defined in Table 1, without a zero bound on net losses. Stress tests ratios are defined in Table 6. "min" stands for the minimum ratio over the 9 quarters of the CCAR scenario or the minimum ratio over the two years of the EBA 2011 stress scenario, other ratios are ratios at the end of the stress scenario. The symbol * indicates ratios based on stress tests results without the effect of capital actions and restructuring plans. Vlab download date: 12/31/2008 (SCAP), 09/30/2011 (CCAR 2012), 09/28/2012 (CCAR 2013), 12/31/2009 (CEBS), 12/31/2010 (EBA), 09/30/2011 (EBA Capital Exercise). The symbol * indicates statistical significance at the 5% level; ** at the 1%. Sample size: 18 (SCAP and CCAR 2012), 17 (CCAR 2013), 50 (CEBS), 53 (EBA), 44 (EBA Cap. Ex.).

Panel A: Rank correlations with Vlab SRISK

Stress tests capital shortfalls	SCAP 2009	CCAR 2012	CCAR 2013	CEBS 2010	EBA 2011	EBA Cap. Ex.
max(0, Shortfall (RWA))	0.507*			-0.153	-0.273*	0.133
Shortfall (RWA)				-0.791**	-0.790**	
Shortfall (TA)					**0.679****	

Panel B: Rank correlations with Vlab M-LVGR$_s$

Stress tests projected ratios	SCAP 2009	CCAR 2012	CCAR 2013	CEBS 2010	EBA 2011
T1R, scenario end		0.204		0.043	0.280*
T1CR, scenario end		0.242			0.282*
T1CR*, scenario end			0.453		0.546**
T1 LVGR, scenario end		0.576*			**0.570****
min T1CR		0.463	0.078		0.274*
min T1CR*		0.797**	0.581*		0.530**
min T1 LVGR		0.684**	0.561*		0.550**
min T1 LVGR*		**0.846****	**0.877****		

Panel C: Rank correlations with Vlab MV loss

Stress tests projected losses	SCAP 2009	CCAR 2012	CCAR 2013	CEBS 2010	EBA 2011
Total Net Loss	0.280	0.604**	0.507*	-0.296*	-0.476**
Total Loss	**0.682****	**0.851****	**0.842****	0.830**	**0.760****
Loan losses	0.580*	0.555*	0.662**	**0.837****	0.751**
Trading losses	0.477*	0.660**	0.589*	0.731**	0.694**

Table 3: **Forecasting during the European sovereign debt crisis.** This table presents the rank correlations and root mean square errors (RMSE) of the EBA and Vlab outcomes with the realized outcomes of banks after disclosure of the EBA stress test in July 2011 (p-values in parentheses). Panel A: rank correlations and RMSE with the six-month realized volatility $RV_{i,t,130}$ (eq. (8)). Panel B: rank correlations and RMSE with the change in realized volatility $RV_{i,t,130} - RV_{i,t-130,130}$. Panel C: rank correlations and RMSE with the 6-month realized loss ($-MV_{it} * \sum_{t+1}^{t+131} \ln(p_{it}/p_{it-1})$). Panel D: rank correlations and RMSE with the 6-month realized return ($-\sum_{t+1}^{t+131} \ln(p_{it}/p_{it-1})$). EBA T1C return = $(\text{T1C}_0 - \text{T1C}_S)/\text{T1C}_0$; EBA risk weight (scenario end) = RWA_S/TA_S; EBA risk weight change = RWA_S/TA_S - RWA_0/TA_0, where subscript S (resp. 0) denotes quantities after (resp. before) the EBA stress scenario. Vlab output was downloaded before the disclosure date of the EBA stress test: 06/30/2011. Sample size: 15 (large), 38 (small), 53 (all).

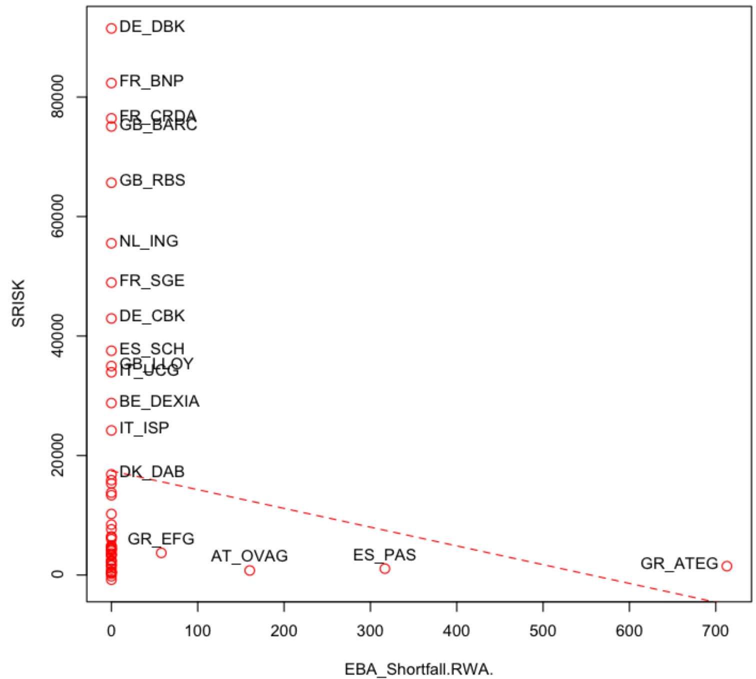
Panel A: 6-month realized volatility		**Rank correlations**			**RMSE**
	Estimated risk	Large	Small	All	All
Vlab	risk weight (eq. (7))	**0.554**	**0.561**	**0.535**	**3.395**
		(0.032)	(0.000)	(0.000)	
EBA	risk weight, scenario end	-0.111	-0.055	-0.140	4.539
		(0.694)	(0.742)	(0.318)	
Panel B: 6-month realized volatility change		**Rank correlations**			**RMSE**
	Estimated risk changes	Large	Small	All	All
Vlab	risk weight (eq. (7))	**0.521**	0.395	**0.434**	**1.305**
		(0.046)	(0.014)	(0.001)	
EBA	risk weight change	0.061	**0.397**	0.341	2.400
		(0.830)	(0.014)	(0.012)	
Panel C: 6-month realized EUR loss		**Rank correlations**			**RMSE**
	Estimated losses	Large	Small	All	All
Vlab	MV loss	0.293	**0.610**	**0.832**	5086
		(0.289)	(0.000)	(0.000)	
EBA	Total Net Loss	0.329	-0.100	-0.272	11202
		(0.232)	(0.549)	(0.048)	
EBA	Total Loss	**0.557**	0.527	0.803	**4945**
		(0.000)	(0.000)	(0.000)	
Panel D: 6-month realized return		**Rank correlations**			**RMSE**
	Estimated returns	Large	Small	All	All
Vlab	LRMES	0.350	0.314	0.299	**0.553**
		(0.201)	(0.055)	(0.029)	
EBA	T1C return	**0.546**	**0.339**	**0.354**	0.767
		(0.035)	(0.038)	(0.009)	

Table 4: **Realized volatility regressions.** Parameter estimates of cross-sectional regressions. Dependent variable: six-month realized volatility (eq. (8)) after disclosure of the EBA stress test in July 2011. EBA T1 LVGR (scenario end) = $T1_S/TA_S$; EBA risk weight (scenario end) = RWA_S/TA_S, where subscript S denotes quantities at the end of the EBA stress scenario. Vlab download date: 06/30/2011. White's heteroskedasticity-consistent standard errors in parentheses. The symbol * indicates statistical significance at the 5% level; ** at the 1%. Sample size: 53.

		1	2	3	4	5	6
	Constant	4.39**	-0.12	6.34**	5.34**	1.70	0.12
		(0.27)	(1.82)	(0.83)	(0.88)	(1.89)	(1.90)
	Book-to-market	0.03**	0.03**	0.03**	0.03**	0.03**	0.04**
		(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.004)
Vlab	risk weight (eq. (7))		2.50*			2.62**	2.99**
			(0.96)			(0.79)	(0.78)
EBA	T1 LVGR, scenario end			-39.99*		-41.39*	-62.44*
				(16.82)		(19.02)	(26.39)
EBA	risk weight, scenario end				-1.75		3.56
					(1.52)		(2.08)
	F-test	11.48**	10.2**	11.88**	6.43**	12.72**	11.25**
	Adj. R^2 (%)	16.78	26.14	29.50	17.28	40.34	44.10

Figure 1: **EBA capital shortfalls vs. SRISK**

(a) Disclosed capital shortfall in the EBA 2011 stress test (eq. (2)) vs. SRISK (EUR millions). Vlab download date: 12/31/2010.



(b) EBA Capital Exercise "Overall Shortfall" (eq. (3)) vs. SRISK (EUR millions). Vlab download date: 09/30/2011.

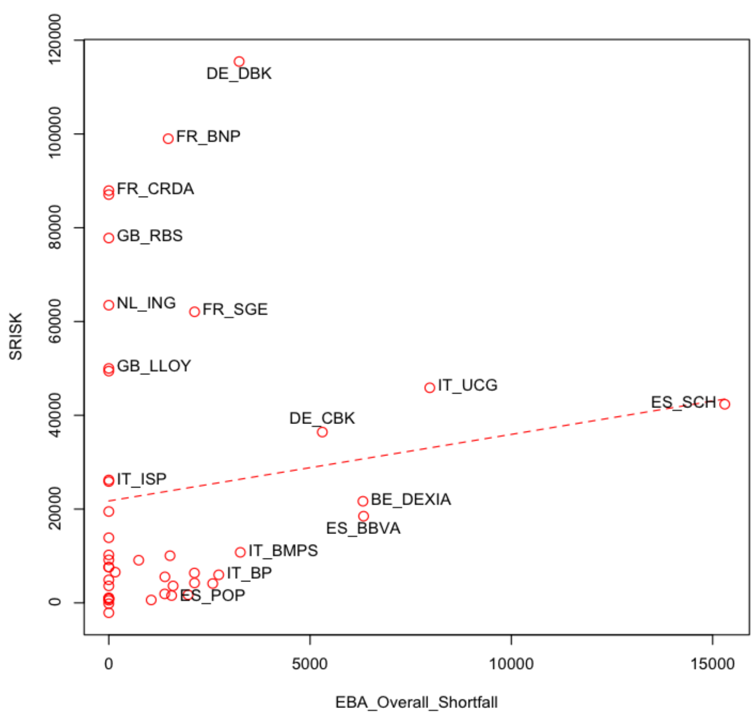
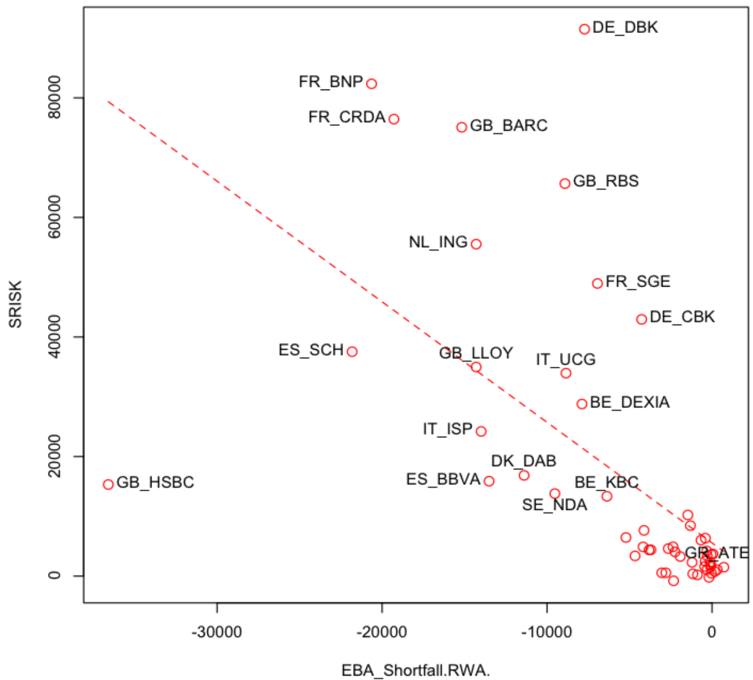


Figure 2: **EBA risk-based and leverage-based capital shortfalls vs. SRISK**

(a) EBA 2011 stress test 'absolute' risk-based capital shortfall/excess vs. SRISK (EUR millions). Vlab download date: 12/31/2010.



(b) EBA 2011 stress test leverage-based shortfall (eq. (5)) vs. SRISK (EUR millions). Vlab download date: 12/31/2010.

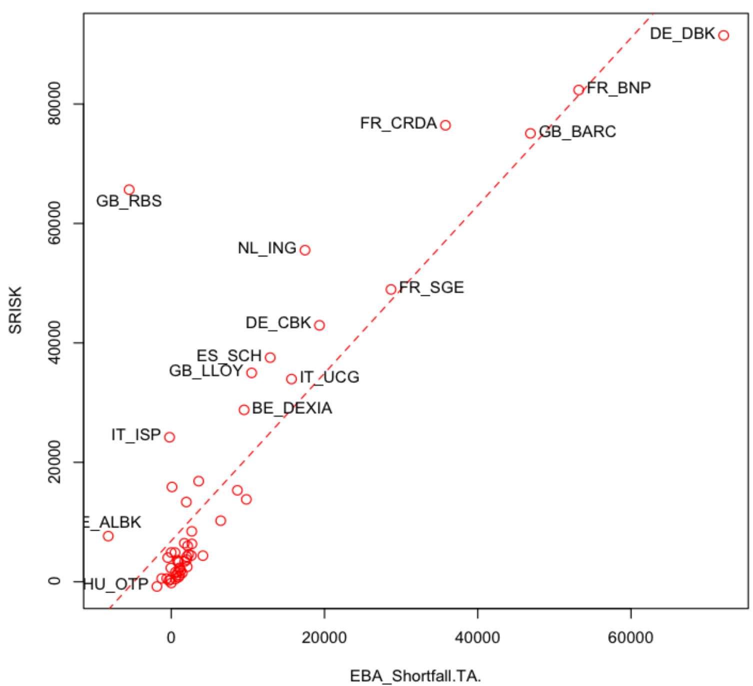
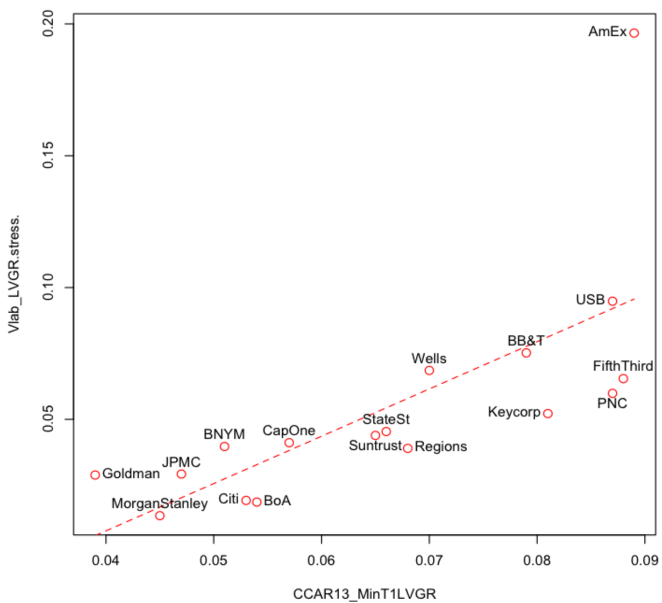


Figure 3: **Stress tests Tier 1 leverage ratios vs. Vlab market leverage ratio.** The Tier 1 Leverage ratio (T1 LVGR) is the ratio of Tier 1 Capital to total assets. The Vlab market leverage ratio (M-LVGR$_s$) is the ratio of market cap to quasi-market assets under the Vlab stress scenario (eq. (4)). "Min" stands for the minimum ratio across the 9 quarters of the CCAR 2013 scenario. CCAR 2013 ratios do not consider the effect of planned capital actions and are disclosed in the Dodd-Frank Act stress test (DFAST 2013). EBA 2011 ratios are the projected ratios at the end of the stress scenario.

(a) CCAR 2013 min T1 Leverage ratio (wihout the effect of capital actions) vs. Vlab market leverage ratio. Vlab download date: 09/28/2012.



(b) EBA 2011 stressed T1 Leverage ratio vs. Vlab market leverage ratio. Vlab download date: 12/31/2010.

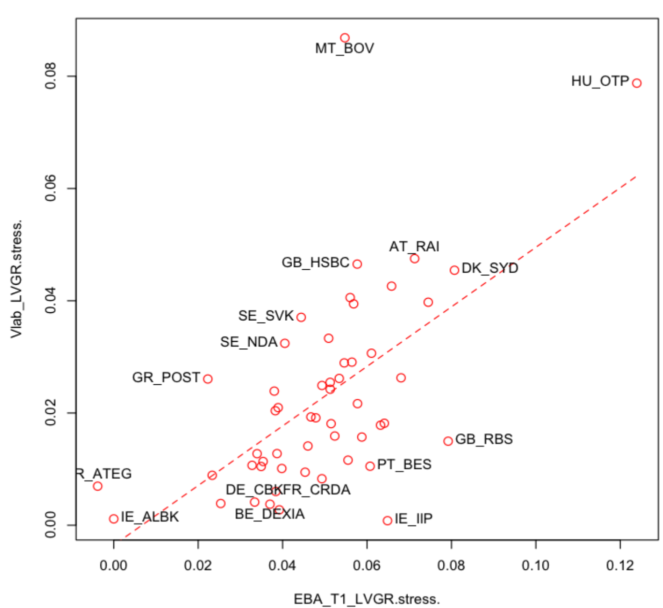
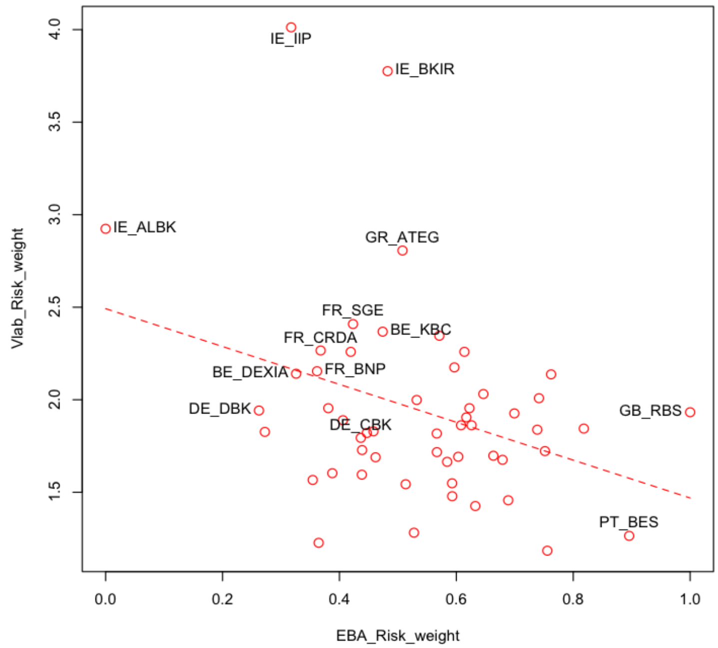


Figure 4: **Stress test risk weight vs. Vlab risk weight and T1 leverage ratio.** Projected regulatory risk weight at the end of the EBA 2011 stress scenario (RWA_S/TA_S, horizontal axis) against Vlab risk weight (a), and the projected Tier 1 Leverage ratio at the end of the EBA 2011 stress scenario (b). Vlab download date: 12/31/2010.

(a) Projected regulatory risk weight vs. Vlab risk weight



(b) Projected regulatory risk weight vs. projected Tier 1 Leverage ratio at the end of the EBA 2011 stress scenario ($T1_S/TA_S$)

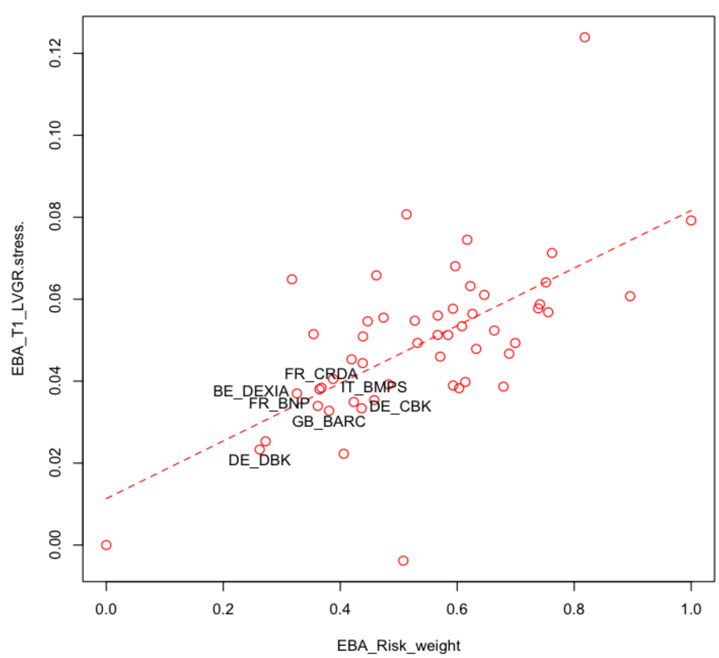
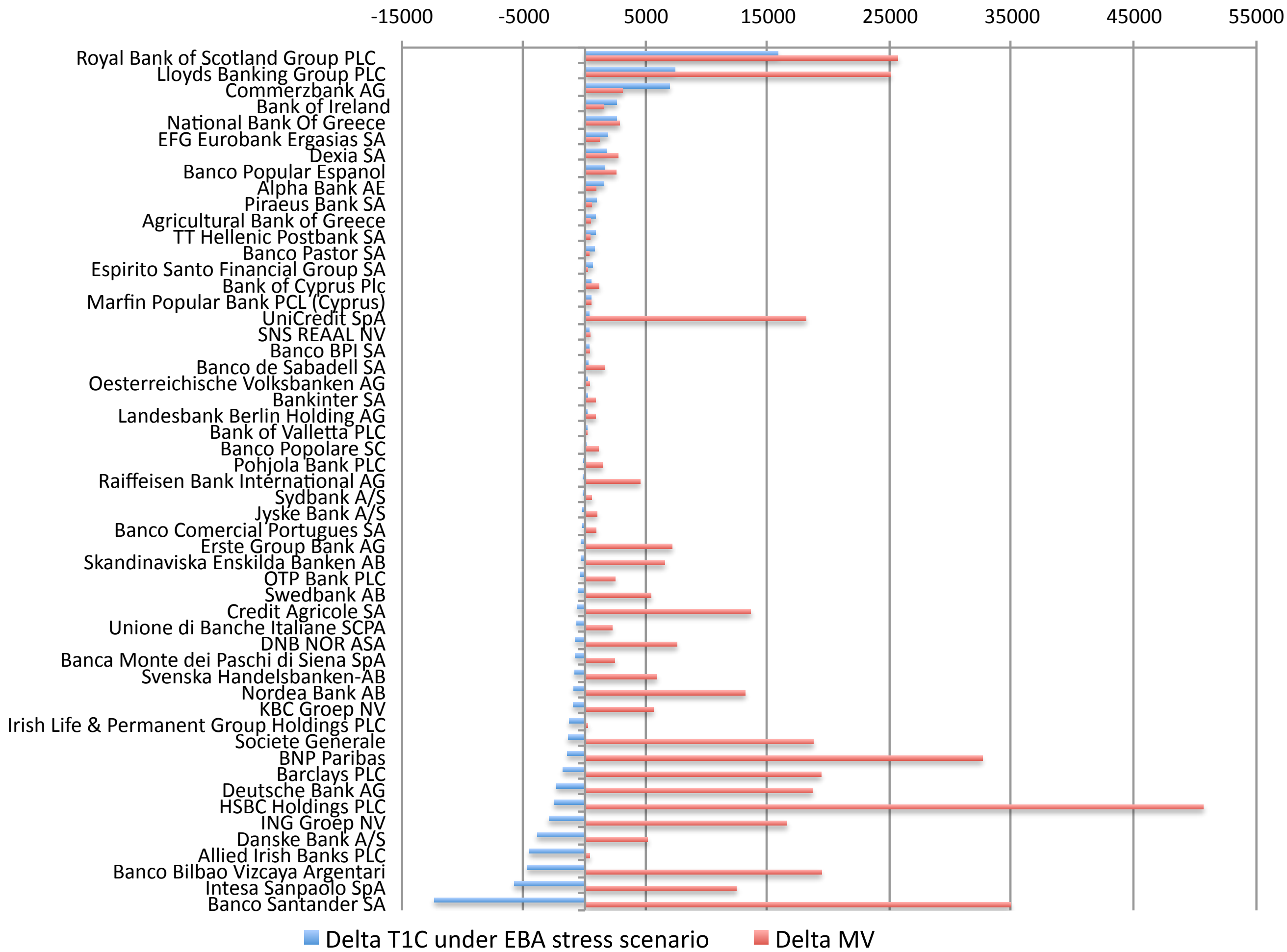


Figure 5: **Stress test change in capital vs. Vlab market cap loss (EUR millions).** Change in Core Tier 1 Capital (Delta T1C) under the EBA 2011 stress scenario (blue) against Vlab's market capitalization loss (Delta MV) (red). Negative changes represent a capital increase. Banks are ranked according to their changes in Core T1 Capital under the EBA stress scenario. Vlab download date: 12/31/2010.



Appendix

A. US and EU stress tests results

Table 5: **US stress tests results.** This table presents the aggregate outcome of US stress tests for which a bank-level outcome is publicly available. "T1" is the Tier 1 capital, "T1C" is the Tier 1 Common capital, and "RWA" are the risk-weighted assets. "T1CR" is the Tier Common Capital Ratio (T1CR = T1C/RWA); "T1R" is the Tier 1 Capital Ratio (T1R = T1/RWA); "Total CR" is the Total Risk-based Capital Ratio (Total CR = Total Capital/RWA); "T1 LVGR" is the Tier 1 Leverage Ratio (T1 LVGR = T1/Total Assets). In parentheses: number of banks failing the regulatory criterion. "min" stands for the cross-sectional average (unweighted) of banks minimum ratios over the 9 quarters of the CCAR scenario. The different minimum ratios may not happen on the same quarter. The column "After scenario*" presents the aggregate results of stress tests without the effect of BHCs planned capital actions (results disclosed in the Dodd-Frank Act stress test 2013 for the CCAR 2013).

	Disclosure	Sample	Scenario horizon	Measure and threshold	Before scenario	After scenario	After scenario*
SCAP 2009	05/07/2009	19 US banks	2009 - 2010	T1	837 $ bn		
		(19 BHCs)	(2 years)	T1C	413 $ bn		
				RWA	7815 $ bn		
				T1CR $\geq$ 4%	6.7% (2)		
				T1R $\geq$ 6%	11.5%		
				T1C shortfall $\leq$ 0		74.6 $ bn (10)	185 $ bn (10)
CCAR 2012	03/13/2012	19 US banks	Q4 2011 -	T1	907 $ bn	540 $ bn	
		(19 BHCs)	Q4 2013	T1C	741 $ bn	438 $ bn	
			(9 quarters)	RWA	7356 $ bn	6904 $ bn	
				T1CR $\geq$ 5%	10.1%	6.6% min (3)	7.3% min (1)
				T1R $\geq$ 4%	12.3%	8.0% min (0)	8.7% min (0)
				Total CR $\geq$ 8%	15.5%	10.8% min (2)	11.5% min (2)
				T1 LVGR $\geq$ 3-4%	7.4%	5.2% min (2)	5.7% min (1)
CCAR 2013	03/14/2013	18 US banks	Q4 2012 -	T1C	792 $ bn		
		(18 BHCs)	Q4 2014	T1CR $\geq$ 5%	11.3%	6.9% min (1)	8.0% min (1)
			(9 quarters)	T1R $\geq$ 4%	13.1%	8.3% min (0)	9.8% min (0)
				Total CR $\geq$ 8%	15.6%	10.7% min (0)	12.1% min (0)
				T1 LVGR $\geq$ 3-4%	8.8%	5.7% min (2)	6.8% min (0)

Table 6: **EU stress tests results.** This table presents the aggregate outcome of EU-wide stress tests for which a bank-level outcome is publicly available. "T1" is the Tier 1 capital, "T1C" is the Tier 1 Core capital, and "RWA" are the risk-weighted assets. "T1CR" is the Core Tier 1 Capital Ratio (T1CR = T1C/RWA); "T1R" is the Tier 1 Capital Ratio (T1R = T1/RWA). In parentheses: number of banks failing the regulatory criterion. Ratios are cross-sectional average ratios at a specific date (scenario start or scenario end). The column "After scenario*" presents the aggregate results of stress tests without the effect of capital actions and restructuring plans.

	Disclosure	Sample	Scenario horizon	Measure and threshold	Before scenario	After scenario	After scenario*
CEBS 2010	07/23/2010	91 banks,	2010 - 2011	T1	1162 EUR bn	1118 EUR bn	
		65% of	(2 years)	RWA	11.29 EUR tn	12.15 EUR tn	
		EU-27 assets		T1R $\geq$ 6%	10.3%	9.2% (7)	
				T1 shortfall ≤ 0		3.5 EUR bn	
EBA 2011	07/15/2011	90 banks,	2011 - 2012	T1	1218 EUR bn	1199 EUR bn	
		65% of	(2 years)	T1C	1006 EUR bn	1001 EUR bn	954 EUR bn
		EU-27 assets		RWA	11.37 EUR tn	13 EUR tn	13.12 EUR tn
				T1CR $\geq$ 5%	8.9% (3)	7.7% (8)	7% (20)
				T1C shortfall ≤ 0		2.5 EUR bn	25 EUR bn
EBA Capital	12/08/2011	65 banks		T1	1190 EUR bn		
Exercise		(excluding		T1C	987 EUR bn		
		Greek banks)		RWA	10.55 EUR tn		
				T1CR $\geq$ 9%	10.2% (27)		
				Sovereign buffer	39 EUR bn		
				Overall shortfall ≤ 0	85 EUR bn (31)		

B. Vlab vs. regulatory stress tests design

B.1. Scenarios

The design of robust and coherent scenarios has been an important topic of the stress testing literature (Breuer et al. (2009); Flood and Korenko (2013); Glasserman et al. (2012) and others). Stress test results are indeed all conditional on the scenario definition. The scenarios of the Federal Reserve, the EBA and Vlab are different on several dimensions: they consider different variables, horizons, stress levels and trajectories. Vlab scenario is the simplest one; it is a one-factor scenario featuring a 40% drop in equity prices over six months. Other variables are considered endogenous to the market factor.

The stress scenario of stress tests usually focuses on an adverse macro-economic scenario defined as a deviation from a baseline scenario. The EBA stress scenario developed by the European Central Bank features a macro-economic stress scenario defined as a deviation of the European Commission forecasts. Conversely, Vlab is an absolute market stress scenario and prevents the procyclical effects of deviation-based scenarios where stress scenarios, and consequent stress tests outcomes, are more severe during bad times. Also, the Vlab scenario horizon of six months is shorter than stress tests scenarios that typically last two years and better justifies the assumption that banks do not react to the stress scenario.

Unlike Vlab, stress tests scenarios are multi-factor scenarios and the principal challenge of the scenario design is coherence. Stresses have to be consistent across the multiple variables so that the joint outcome of the scenario is economically realistic. The challenge of coherence also grows as the number of variables increases. The SCAP 2009 considered only three factors: real GDP growth, the unemployment rate and house prices. The CCAR 2012 stress scenario defines trajectories for 25 macro-economic and financial variables, and additionally accounts for a global market shock on the six BHCs with the largest trading activities. The number of factors in the ECB scenario exceeds 70 variables and the ECB also considers a market stress scenario conditional on the macro-economic scenario.

The question of causality between macro-economic variables and financial variables is raised by Alfaro and Drehmann (2009), where severe macro-economic conditions are shown to be subsequent to the beginning of financial crises. This argument questions the whole relevance of macro-economic scenarios and the models used to map macro scenarios on the financial performance of banks. Macro-economic scenarios defined by exogenous shocks on macro-economic variables are rooted in the microprudential perspective as it ignores the endogenous nature of systemic risk where feedback effects from the financial sector amplify the macro-economic crisis.

Borio et al. (2012) indicate that the selection of a stress scenario is easier when the stress tests are crisis management tools since they are designed to address apparent risks. Indeed, in the SCAP 2009, the decline of local house prices was a significant factor in larger loan loss rates. In the EBA 2011, country-specific shocks to sovereign credit spreads translated into valuation haircuts of banks' sovereign exposures.

Other concerns come from the coherence and plausibility of stress scenarios, which could dilute their severity. Relating to Vlab's 40% equity market index decline over six months, the EBA stress scenario considers a fall of between 10% to 20% in equity prices over two years. The CCAR 2012 assumes a 50% drop in the Dow Jones total stock market index in the middle of the scenario (late 2012) but reverts to a higher level at the end of the scenario. We argue that this focus on plausibility, while can facilitate the buy-in of all stakeholders involved in a macro stress test (Borio et al. (2012)), severely undermines the reliability of stress test results.

B.2. Data

B.2.1. Public vs. supervisory

US and EU-wide stress tests are implemented using extended bank supervisory data. US BHCs submit their data confidentially to the Federal Reserve using standardized FR Y-14A forms. These forms contain detailed information on capital composition, loan and security portfolios, trading and counterparty exposures, and historical P&L. The reports additionally collect BHCs' own projections of losses and revenues as well as their estimates of exposure sensitivities to a set of risk factors specified by the Federal Reserve. In Europe, banks implement stress tests themselves and use their own data. Banks are encouraged by the EBA to use all the time series available on credit risk parameters and P&L figures for the application of the macro scenario.

Relating to EU and US stress tests, Vlab could be qualified as a non-invasive stress test. Vlab results are obtained from a reduced dataset of publicly available data including historical market prices, market capitalization, and leverage.

B.2.2. Market vs. accounting

Stress scenarios are generally applied to accounting data in supervisory stress tests whereas Vlab applies to the market value of equity. In that respect, Vlab may be considered a mark-to-market stress test. Market prices are believed to reflect expectations of all market participants on bank performance and are available in real time. Harris et al. (2012)

further argue that balance sheet data, since they are single-point estimates at a specific reporting date, cannot reflect the dynamic risks embedded in assets and liabilities. For these reasons, market data are considered to be more informative than accounting numbers.

Another problem concerning accounting data comes from their lack of uniformity. Accounting standards are different; US banks report under US GAAP and European banks under IFRS. This leads to the large differences in the leverage ratios mentioned earlier. But even under the same standards there still exist large cross-border differences as pointed out by the EBA. Moreover, accounting rules are subject to different interpretations at the bank level.

For these reasons, the reliability of accounting data may be questionable during a crisis. We have shown in Section 4.4 that banks with high book-to-market ratios experienced higher realized volatility during the European sovereign debt crisis. The positive and significant correlation between the book-to-market ratio and risk indicates that investors do not trust banks estimates of capital levels during crises. In the US too, the book-to-market ratio increased after the financial crisis; market capitalizations were far above regulatory capital levels in 2008 but the market cap is generally below the regulatory capital in 2011.

B.3. Vlab as a macroprudential benchmark

According to Borio et al. (2012), any stress test has four elements: the scenario, the risk exposures, the model, and the outcome. The scenario specifies the shocks that will be applied to bank data using a specific model, and the resulting measures are the final outcome of the stress test. The simplicity of Vlab's data and scenario may represent its main weakness. It does not have the information granularity at the asset category or exposure level that stress tests possess. Vlab does not, however, do worse than stress tests when it comes to forecasting real bank outcomes during a crisis, as shown in Section 5.2. Instead, the use of publicly available market data makes Vlab estimates richer and more transparent; it provides real-time forward-looking measures of risk and addresses concerns on the reliability and consistency of accounting-based measures across firms. The simplicity of the Vlab scenario makes its outcomes robust to various economic environments and supports the argument in Haldane (2011, 2012) that complex uncertain environments call for simpler rules. Vlab severity makes SRISK the binding constraint for banks whereas stress tests estimates are not binding (Hanson et al. (2011)). Vlab SRISK may therefore be a useful complementary tool for regulators to assess their own stress tests outcomes.

C. Portfolio choice under regulatory risk weights

We demonstrate in this section the weakness of Basel regulatory risk weights as an aggregate measure of bank risk where the bank is seen as a portfolio of assets. The bank chooses its allocation of resources to maximize its return subject to a tolerable level of risk. Regulators implement several standards of prudent risk but these may sometimes be misguided. Here we consider the allocation of a fixed investment budget to asset categories subject to the regulatory requirement implemented in a stylized version of Basel standards.

Let TA be the total assets to be allocated between cash, C (equivalent to the capital requirement for credit risk in Basel II), and other risky assets. Let there be N risky assets with conditional expected returns given by the $(N \times 1)$ vector m, and conditional covariance matrix given by the $(N \times N)$ matrix H. According to Basel rules, each of these assets has a risk weight w_j between zero and one that we assemble in a $(N \times 1)$ vector w. The solution is a $(N \times 1)$ vector of dollars to be invested in each asset, q. The vector q will also determine the optimal exposures at default under Basel II and the optimal RWA, $w'q$. The risk budget requires that $C \geq kw'q$, where k is the prudential capital ratio and $C = TA - \iota' q$, where ι is a $(N \times 1)$ vector of ones.

To maximize asset returns subject to these constraints the firm must solve

$$\begin{aligned} &\max_q \; q'm \\ &\text{s.t. } TA - \iota' q \geq kw'q, \quad q \geq 0 \end{aligned} \tag{9}$$

The Lagrangian of this maximization problem is

$$L(q, \lambda, \mu) = m'q - \lambda \left(TA - \iota' q - kw'q \right) - \mu' q \tag{10}$$

where the scalar λ and the $(N \times 1)$ vector μ are Lagrange multipliers. The first order condition of equation (10) with respect to q is given by

$$m' + \lambda \left(\iota' + kw' \right) - \mu' = 0 \tag{11}$$

Multiplying equation (11) by q and recognizing that either q or μ will be zero for each asset (from the first order condition of (10) w.r.t. to μ), then

$$
\begin{aligned}
m'q + \lambda\left(\iota' q + kw'q\right) &= 0 \qquad (12)\\
m'q &= -\lambda TA\\
\lambda &= \frac{-m'q}{TA}
\end{aligned}
$$

Replacing λ in (11), we obtain

$$
m' - \left(\frac{m'q}{TA}\right)\left(\iota' + kw'\right) - \mu' = 0 \qquad (13)
$$

Hence all non-zero allocations q_j, must satisfy

$$
\begin{aligned}
m_j - \left(\frac{m'q}{TA}\right)\left(1 + kw_j\right) &= 0 \qquad (14)\\
\frac{m_j}{1 + kw_j} &= \frac{m'q}{TA}
\end{aligned}
$$

Supposing that each asset has a different value of $m_j(1+kw_j)^{-1}$, then the maximum will occur if the entire portfolio of the bank $\iota' q$ is invested in the asset with the greatest value of this ratio. The amount invested in this asset will be

$$
q_j = \frac{TA}{1 + kw_j} \qquad (15)
$$

If there are multiple assets with the same value of this ratio, the performance will be the same for any feasible allocation to these assets.

The main observation is that the risk of a portfolio is always less than or equal to the sum of the risks of its components. The use of risk-weighted assets ignores this portfolio feature of risk and consequently there is no incentive from the regulatory perspective to diversify. The only case where this measure is appropriate is when all assets are perfectly correlated.

For firms with risk aversion, risk weights act as an additional cost on assets.[25] Glasserman and Kang (2013) show that risk weights that are optimal from both banking and regulatory perspectives have nothing to do with risk but are instead proportional to the asset returns m. These optimal risk weights do not distort the portfolio a bank would chose without the

[25]It can be shown that the additional cost will be greater if the threshold k is large and for a bank with low risk aversion and low capital.

risk-based capital constraint and satisfy the regulator's objective to limit the bank's portfolio riskiness.[26]

Then, if some risk weights are underestimated or are not adjusted to reflect increased risk during a crisis, a bank will choose its optimal asset with the most underestimated risk weight and this will automatically lead to excessive leverage. If w_j is the risk weight of the optimal asset and since $q_j = \iota' q = TA - C$, the leverage ratio C/TA from (15) is $1 - (1 + kw_j)^{-1}$. Consequently, banks will take excessive leverage if their risk weights are not adequately adjusted to more severe economic conditions.

[26] Also note that the portfolio distortion problem does not exist for banks that are only leverage-constrained since the additional charges are the same for all assets.