DEUTSCHE BUNDESBANK
EUROSYSTEM

Discussion Paper

Deutsche Bundesbank
No 19/2021

System-wide and banks' internal stress tests: Regulatory requirements and literature review

Kamil Pliszka

Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-0

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

Internet http://www.bundesbank.de

ISBN 978–3–95729–827–0

Non-technical summary

Research question

The financial crisis of 2007-09 revealed the vulnerability of banks. In response, stress testing was established as a key supervisory element and as an essential risk management tool. Based on these developments, this paper first deals with system-wide stress tests and provides an overview of their evolution, compares programs in major jurisdictions, and reviews the academic literature. Second, this paper summarizes regulatory requirements and expectations for banks' internal stress tests and gives an overview of the relevant literature.

Contribution

This paper contributes to the literature, as it deals with both system-wide stress tests and banks' internal stress tests and discusses the differences between the two types. Moreover, it provides an update of earlier surveys by covering discussions from 2019-20 on the envisaged changes to system-wide stress testing in the EU and the US.

Results

System-wide stress tests have gained in importance. For instance, they feed into the calculation of capital requirements in the EU. The literature shows that the disclosure of stress test results reveals new information to the market. Furthermore, banks that participate in system-wide stress tests increase their capital ratios and shift lending to less risky borrowers. Banks' internal stress tests are deeply embedded in the Basel III framework and the banks that apply internal models are subject to more stringent stress testing requirements. For example, these banks have to ensure capital adequacy if the internal risk parameters are being stressed. Research on banks' internal stress tests shows that stress scenarios derived from expert judgment should be complemented by scenarios, which consider historical characteristics of the risk factors. As stress testing is exposed to considerable model and estimation risk, banks should carry out extensive robustness checks. In sum, both system-wide and banks' internal stress tests play a complementary role in ensuring resilience of individual banks and the financial system to adverse shocks.

Nichttechnische Zusammenfassung

Fragestellung

Als Reaktion auf die Finanzmarktkrise 2007 bis 2009 wurden Stresstests als ein zentrales Element der Aufsicht und als ein wichtiges Instrument im Risikomanagement etabliert. Ausgehend von diesen Entwicklungen befasst sich dieses Papier zunächst mit systemweiten Stresstests und gibt hierbei eine Übersicht über die Entwicklung, vergleicht Programme aus verschiedenen Jurisdiktionen und rezensiert die wissenschaftliche Literatur. Zum anderen fasst das Papier regulatorische Erwartungen zu bankinternen Stresstests zusammen und gibt einen Überblick über die einschlägige Literatur.

Beitrag

Der Beitrag des Papiers liegt darin, dass es sich sowohl mit systemweiten als auch mit bankinternen Stresstests beschäftigt und dabei die Unterschiede diskutiert. Darüber hinaus geht es auf Diskussionen aus 2019 und 2020 zu angedachten Änderungen bei systemweiten Stresstests in der EU und in den USA ein.

Ergebnisse

Systemweite Stresstests haben an Bedeutung gewonnen und fließen beispielsweise in die Bestimmung von Kapitalanforderungen in der EU ein. Die Literatur zeigt, dass die Offenlegung von Ergebnissen aus Stresstests dem Markt neue Informationen liefert. Darüber hinaus zeigt sich, dass Banken, die an systemweiten Stresstests teilnehmen, ihre Kapitalquoten erhöhen und die Kreditvergabe zu weniger riskanten Schuldnern verlagern. Bankinterne Stresstests sind tief in Basel III eingebettet, wobei Banken, die interne Modelle verwenden, strengeren Anforderungen unterliegen und beispielsweise Kapitaladäquanz auch mit "gestressten" internen Risikoparametern sicherstellen müssen. Die Forschung zeigt unter anderem, dass Szenarien neben Expertenschätzungen auch historische Eigenschaften von Risikofaktoren berücksichtigen sollten. Da Stresstests erheblichen Modell- und Schätzrisiken ausgesetzt sind, sollten Banken Robustheitsüberprüfungen durchführen. Zusammenfassend kann festgehalten werden, dass sich systemweite und bankinterne Stresstests bei der Sicherstellung der Widerstandfähigkeit von einzelnen Banken und des gesamten Finanzsystems gegenüber negativen Schocks ergänzen.

Bundesbank Discussion Paper No 19/2021

System-wide and banks' internal stress tests: Regulatory requirements and literature review*

Kamil Pliszka
Deutsche Bundesbank

Abstract

This paper deals with both system-wide and banks' internal stress tests. For system-wide stress tests it describes the evolution over time, compares the stress test design in major jurisdictions, and discusses academic research. System-wide stress tests have gained in importance and nowadays serve as a key regulatory tool. For instance, they feed into the calculation of capital requirements in the EU. The literature shows that the disclosure of stress test results reveals new information to the market. Furthermore, banks that participate in system-wide stress tests increase their capital ratios and shift lending to less risky borrowers. For banks' internal stress tests, this paper gives an overview of the regulatory requirements under Pillars 1 to 3 of Basel III and reviews the academic literature. Stress testing is deeply embedded in the Basel III framework. Banks that choose to apply internal models for calculating capital requirements are subject to more stringent stress testing requirements and, for example, have to ensure capital adequacy if the internal risk parameters are being stressed. The academic research on banks' internal stress tests shows that stress scenarios derived from expert judgment should be complemented by scenarios which are selected on the basis of algorithms that consider historical characteristics of the risk factors. Furthermore, banks' conventional credit risk models can be modified and used for stress testing. As stress testing is exposed to considerable model and estimation risk, banks should carry out extensive robustness checks. In sum, both system-wide and banks' internal stress tests play a complementary role in ensuring the resilience of individual banks and the financial system to adverse shocks.

Keywords: Literature survey, regulatory expectations, regulatory requirements, stress testing
JEL classification: G21; G32; G38

*Contact address: Kamil Pliszka, Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main, Germany, E-mail: kamil.pliszka@bundesbank.de.
I want to thank an anonymous referee and the participants of the Bundesbank Seminar. The views expressed in this paper are those of the author and do not necessarily coincide with the views of the Deutsche Bundesbank or the Eurosystem.

1 Introduction

The financial crisis of 2007-09 revealed that banks are vulnerable to macroeconomic shocks and liquidity dry-ups. In response, regulators and supervisory authorities have established stress testing as a key supervisory element and as an important risk management tool. In this regard, system-wide stress test programs[1] have been established as a tool to assess whether banks and the financial sector as a whole can withstand stress scenarios. Furthermore, the finalized Basel III framework provides more stringent requirements and expectations under Pillars 1 to 3 of Basel III for banks' internal stress tests.

Based on these developments, this paper first deals with system-wide stress tests and provides an overview of their evolution, compares programs in major jurisdictions and reviews the academic literature. Second, this paper summarizes regulatory requirements and expectations under Pillars 1 to 3 of Basel III for banks' internal stress tests and then gives an overview of the relevant literature. It contributes to the literature, as it deals with both system-wide stress tests and banks' internal stress tests and discusses the differences between them. Moreover, this paper provides an update of earlier surveys by covering discussions from 2019-20 on the envisaged changes to system-wide stress testing in the EU and the US.

System-wide and banks' internal stress tests both follow the same process steps (see Figure 1).

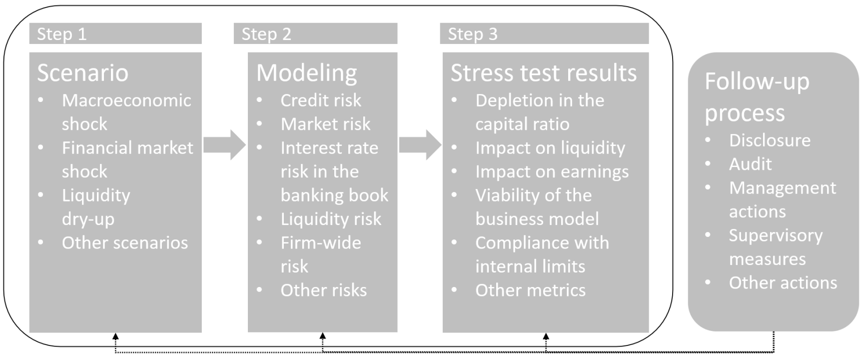

Figure 1: This figure sketches the process steps of a stress test. A scenario needs to be selected (step 1) and, in step 2, mapped onto bank-specific parameters. Then, these parameters are used to calculate the impact on the metric of interest (step 3). A follow-up process concludes the exercise.

In step 1, a scenario needs to be selected. Scenarios can consist of shocks to multiple risk factors or, in the case of sensitivity analyses, of a shock to a single risk factor (e.g., an interest rate increase by 200 bps for all maturities). Then, in step 2, a scenario is mapped onto bank-specific parameters. The modeling, i.e., the projection of a scenario onto risk parameters, can be conducted either by banks ("bottom-up stress test") or by a central bank/supervisory

[1] For the purpose of this paper, system-wide stress tests are micro-/macroprudential programs, which are carried out by central banks and supervisory authorities that aim at assessing the resilience of banks and the financial system as a whole to macroeconomic stress scenarios.

authority ("top-down stress test").[2] In step 3, the bank-specific parameters are used to calculate the impact on the metric of interest (e.g., depletion in the Common Equity Tier 1 (CET1) capital ratio or the Liquidity Coverage Ratio (LCR)). Although most stress tests aim at banks' solvency and thus assess regulatory capital ratios, thematic stress tests with an other main topic (e.g., liquidity, interest rate risk, climate risk, etc.) gain in importance. By contrast, so-called reverse stress tests start from a stress test result (e.g., a given loss or a violation of regulatory capital requirements) and then search for the most likely scenario leading to this outcome. One of the largest differences between system-wide and banks' internal stress tests is the subsequent follow-up process once the stress test results are available. In system-wide stress tests, supervisors can top-up capital requirements, follow-up on identified deficiencies, and disclose the results. By contrast, banks' internal stress test results feed into the Internal Capital Adequacy Assessment Process (ICAAP), are expected to be reported to the governing body, and be used effectively (see BCBS (2018, p. 5)).[3] Hence, a bank has some discretion on the conclusions it wants to draw from the results of internal stress tests.[4]

As system-wide stress tests and banks' internal stress tests follow the same process steps and thus have several commonalities, drawing a clear distinction between them proves to be difficult. This is even more true if one considers that models used for system-wide stress tests are often adjusted versions of internal risk models. Figure 2 visualizes both types of stress tests interacting with each other.

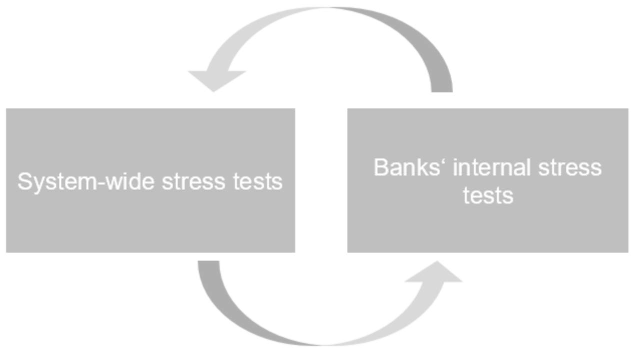

Figure 2: System-wide and banks' internal stress tests interact with each other.

Given the conceptual commonalities between both stress test types, supervisors explicitly emphasize that stress tests are designed to foster discussions about risk management practices, including at governing body level. This is consistent with the supervisory expectation that internal risk management and stress testing will benefit from the experience of system-wide stress tests (see BCBS (2018, pp. 4-5)).

2 System-wide stress tests can be bottom-up stress tests. In this case, banks project the impact of a scenario onto their risk parameters and then the central bank/supervisory authority has to assure the quality and appropriateness of the banks' numbers. This can be achieved by running supervisory challenger models, for example. By contrast, top-down stress tests first require central banks/supervisory authorities to calibrate a model, which is then used to project a scenario's impact onto banks' risk parameters.

3 Supervisors can, however, set capital surcharges for Pillar 2 based on results of internal stress tests.

4 However, in some jurisdiction guidance how stress test results should feed into the internal governance is given (see, e.g., ECB (2018)).

The remainder of this paper is structured as follows: Section 2 deals with system-wide stress tests. Specifically, it describes the evolution from the origins of stress testing until they became an essential supervisory tool, compares programs from major jurisdictions, and discusses academic research. In Section 3, regulatory requirements and academic research on banks' internal stress tests are discussed. Section 4 deals with the differences between both stress test types. Section 5 concludes.

2 System-wide stress tests

This section summarizes the evolution of stress tests in the past until system-wide stress tests became the essential tool they are today. Furthermore, the section provides an overview of stress testing programs in selected jurisdictions and discusses academic research.

2.1 Evolution

In 1999, the IMF and the World Bank introduced the Financial Sector Assessment Program (FSAP) in order to review a country's financial sector by using stress tests along with several other tools. Generally, only a single country[5] is reviewed, yet the idea of having a tool for reviewing the financial system as a whole by a central entity emerged. Furthermore, the Committee on the Global Financial System (CGFS) published a paper (see CGFS (2000)) in 2000 on so-called aggregate stress tests where banks had to translate a macroeconomic scenario into a risk measure and report it back to a central entity. This central entity (e.g., a central bank) would then compose a risk measure for the overall system vulnerability. Even though a composite risk measure is not the focus of more topical stress testing programs, the idea of computing the impact of macroeconomic scenarios on individual banks for assessing the vulnerability of the banking sector was born.

In the following years, system-wide stress tests were not carried out regularly. However, in response to the 2007-09 financial crisis, system-wide stress tests did experience a real boom. The Federal Reserve conducted the Supervisory Capital Assessment Program (SCAP) in 2009, and published the design of this stress test (see Federal Reserve (2009a)) and its results (see Federal Reserve (2009b)). The Committee on European Banking Supervision (CEBS) and, under its new name since 2011, the European Banking Authority (EBA) have conducted several stress test exercises for European banks. The first two iterations were aimed at the resilience of the European banking system and were conducted in 2010 (see CEBS (2010a))[6] and 2011 (see EBA (2011)). The third exercise was conducted as part of the comprehensive assessment in preparation for the ECB's takeover of banking supervision in the euro area (see EBA (2014)). These applications of stress testing reflected the increased appetite for stress

5 The FSAP has been expanded to other regions such as the euro area, where it was carried out in 2018 (see IMF (2018)).

6 The disclosure of stress test results played a major role following the financial crisis of 2007-09 in strengthening the public's trust in banks. Against this background, the CEBS has been criticized for the fact the results of the 2010 stress test have not been disclosed publicly at bank level.

testing as a supervisory tool in major jurisdictions.[7]

Nowadays, a solvency system-wide stress test for assessing banks' vulnerability to macroeconomic stress scenarios is carried out annually by the Federal Reserve (since 2011) and every two years by the ECB jointly with the EBA (since 2014) for banks, which are under the direct supervision of the ECB. In addition to the regular application of system-wide stress tests in major jurisdictions, system-wide stress tests have also gained importance. For instance, the EU's Pillar 2 Guidance (P2G), i.e., the level of capital a supervisor expects a bank to hold, incorporates stress test results (see EBA (2018, pp. 6-7)). Another example is the Federal Reserve, which uses results from its stress test exercises to assess banks' capital planning.

System-wide stress tests will continue to evolve. In September 2019, Randal Quarles, Vice Chair of Supervision at the Federal Reserve Board, sketched a proposal for the stress capital buffer (see Quarles (2019)), in which non-stress and stress capital requirements were to be combined. The proposal was not fully fleshed out, but options include replacing the static Basel III capital conservation buffer of 2.5% with bank-specific results from stress tests and the introduction of a positive baseline counter-cyclical capital buffer.

Ongoing discussions on the role and design of stress tests also take place in the EU. A proposal by the ECB (see Enria (2019)) and the EBA (see EBA (2020)) sets out an approach which takes account of both bank and supervisory stress test calculations. The supervisory leg would be the basis for supervisory measures and would be used to determine the P2G. By contrast, the bank leg would allow banks to explain any discrepancies between the supervisory results and their own calculations. A reason for this envisaged methodological change of the stress test in the EU might be the resource-intensive design of the current "constrained bottom-up approach," which allows banks to calculate the impact of a given macroeconomic scenario onto their risk parameters. Taking into consideration that banks are incentivized to minimize the capital impact of a scenario (see Quagliariello (2019)), the banks' numbers might be biased. However, the supervisory quality assurance process and challenger models for banks' calculations can mitigate these distortions.[8] On the one hand, a stress test with both bank and supervisory calculations might save supervisory resources. On the other hand, significant gaps between both calculations will be difficult to explain to the public and might bear reputation risk for both parties.

2.2 Overview of selected system-wide stress testing programs

This section provides an overview of the main characteristics of system-wide stress testing programs in various jurisdictions. Given the macroprudential and systemic nature of these exercises, only jurisdictions domiciling globally systematically important banks (G-SIBs)

[7] In elaborating on that change in supervision, Borio (2011) outlines how financial stability and macroprudential policy became more important after the 2007-09 financial crisis. Regulation and supervision changed from a micro-orientated perspective to a macro-oriented one. Accordingly, the focus changed from the stability of individual banks to the stability of the financial system as a whole.

[8] The design of the quality assurance process may also affect the results. For example, a supervisor might overlook or ignore specific risks which are relevant for some banks. Similarly, supervisors might be incentivized to attest banks vulnerabilities in order to justify the need for their work.

are selected.[9] The emphasis is on solvency stress tests, noting that several jurisdictions run thematic stress tests with other main topics.[10] The following five jurisdictions with a total of eight stress tests are included: (i) Euro area (Single Supervisory Mechanism (SSM) stress test of the ECB/EBA), (ii) United Kingdom (Bank of England (BoE) Annual cyclical scenario (ACS)), (iii) Switzerland (Swiss National Bank (SNB) building block analysis, Swiss Financial Market Supervisory Authority (FINMA) stress test), (iv) USA (Dodd-Frank Act stress test (DFAST), Comprehensive Capital Analysis and Review (CCAR)), (v) Japan (Bank of Japan (BoJ) stress test, Financial Services Agency of Japan (JFSA) stress test). The information has been obtained mainly from a survey among member authorities of the Basel Committee on Banking Supervision (BCBS), which has been summarized in FSI (2018a) and FSI (2018b).[11] In addition, the stress test from the Bank of England (see Bank of England (2019)) and further information have been incorporated. The stress tests within the FSAP are not included because the IMF and the World Bank mainly use data that are already available in the relevant jurisdiction. Hence, the structure of the FSAP depends on the data available in the national authorities.[12]

Main objective: Stress testing programs differ in terms of their main objectives. Authorities that are responsible for microprudential banking supervision focus on the resilience and capital adequacy of individual banks and enrich their Supervisory Review Process (SRP) with stress tests (SSM stress test, BoE ACS, FINMA stress test, DFAST, CCAR, JFSA stress test) (see FSI (2018b, p. 2)). In this regard, stress test results can feed into the calculation of capital requirements (SSM stress test) or affect a bank's capital planning (CCAR) (see FSI (2018b, p. 2)). The SSM stress test, the FINMA stress test and the JFSA stress test will also challenge banks' risk management (see FSI (2018b, p. 2)). By contrast, authorities that are mandated to monitor financial stability concentrate more on the resilience of the financial system as a whole (BoE ACS,[13] SNB building block analysis, BoJ stress test) (see FSI (2018b, p. 2)). However, this does not preclude macroprudential authorities from assessing the resilience of individual banks (e.g., BoJ stress test) (see FSI (2018b, p. 2)).

[9] The Financial Stability Board (FSB) publishes an annual list of around 30 banks that are deemed to be globally systematically important. Jurisdictions in which G-SIBs are headquartered are included in this paper as far as possible. However, Canada and China are not included, as information on system-wide stress testing frameworks in these jurisdictions is only irregularly available publicly. Of course, other jurisdictions carry out stress tests as well. For example, Ulloa (2020) gives an overview of system-wide stress test programs in Latin America.

[10] For example, the ECB ran an interest rate stress test in 2017 (see ECB (2017)) and a liquidity stress test in 2019 (see ECB (2019)).

[11] A further useful survey on system-wide stress tests among BCBS member authorities is BCBS (2017).

[12] Moretti, Stolz, & Swinburne (2008) describe the IMF's experience in stress testing and give insights into the applied models. Moreover, their paper describes how the methodologies evolved over time and outlines the high degree of dependence on available data in jurisdictions. Blaschke, Jones, Majnoni, & Martinez Peria (2001) describe the objective of stress testing in the FSAP and provide methodologies for the considered risk types. Čihák (2007) also gives an overview of the applied methodologies for the considered risk types, but describes their applications in the IMF's stress testing tool. Schmieder, Hesse, Neudorfer, Puhr, & Schmitz (2012) deal with the role of liquidity risk in the FSAP, as this became important after the 2007-09 financial crisis.

[13] The Bank of England is responsible for both microprudential (through the Prudential Regulatory Agency (PRA)) and macroprudential supervision.

Coverage of banks: The coverage of banks included in a system-wide stress test differs significantly among the programs. The ECB carries out its stress test for around 100 banks[14] (see FSI (2018b, p. 7)). The BoE ACS is conducted for the largest banks and, in 2019, covered seven banks.[15] The SNB includes around 90 banks in its exercise but requires additional data requests for the two G-SIBs, i.e., Credit Suisse and UBS (see FSI (2018b, p. 7)). The FINMA conducts a stress test specifically for the two G-SIBs (see FSI (2018b, p. 7)). Stress test programs in the US are carried out for big banks only. Until 2017, banks with total assets of USD 50bn or more were subject to stress testing. This threshold was increased to USD 100bn leading to a sample of 35 banks for the 2018 exercise, and to USD 250bn covering 18 banks in 2019 (see Federal Reserve (2019)). In Japan, the BoJ includes around 370 banks in its exercise and covers 80% to 90% of total credit outstanding (see FSI (2018b, p. 7)). The JFSA stress test is tailored to big banks and encompasses the three G-SIBs and the four domestic systematically important banks (D-SIBs) (see FSI (2018b, p. 7)). These seven banks account for 70% of the total assets of the Japanese banking sector (see FSI (2018b, p. 7)).

Number of scenarios: All stress test programs prescribe a baseline scenario and one to five stress scenarios. It may be noted that programs that are run for fewer banks tend to prescribe more scenarios. The ECB and BoE prescribe one stress scenario (see FSI (2018b, p. 10)). Two stress scenarios are prescribed in the stress tests of the FINMA, in both US stress tests (DFAST and CCAR) and in the exercise of the BoJ (see FSI (2018b, p. 10)). On top of the two stress scenarios of the DFAST, the CCAR also comprises scenarios with banks' own forecasts for the baseline scenario and the adverse scenario[16] (see FSI (2018b, p. 10)). The SNB provides an extended set of three to five stress scenarios (see FSI (2018b, p. 10)). The JFSA provides narratives for potential risks (e.g., economic downturn) and requires banks to run a non-specified number of stress scenarios (see FSI (2018b, p. 10)).

Data: Authorities initiate dedicated data collections in all system-wide stress test programs. In some exercises data points are pre-filled from regular supervisory reporting systems (SNB building block analysis, DFAST, and CCAR) (see FSI (2018b, p. 12)). Furthermore, the BoJ and the JFSA stress tests are augmented by publicly available data (see FSI (2018b, p. 12)). Stress test programs with a high number of participating banks apply proportionality in the sense that smaller banks have to provide less or less granular data (ECB stress test, SNB building block analysis, and JFSA stress test) (see FSI (2018b, p. 12)).

Concept and balance sheet assumption: Unlike top-down stress tests, bottom-up stress tests require banks to calculate the projections of a macroeconomic scenario onto their risk parameters and, eventually, the relevant key metrics. It turns out that microprudential authorities tend to carry out system-wide stress tests as bottom-up exercises, i.e., banks map the scenarios onto their risk parameters and the authorities quality-assure and challenge the banks' calculations.

[14] This is necessary, as stress test results feed into the calculation of capital requirements, i.e., P2G. For smaller euro area banks (less significant institutions (LSIs)), national authorities carry out LSI stress tests.

[15] The banks are Barclays, HSBC, Lloyds Banking Group, Nationwide, The Royal Bank of Scotland Group, Santander UK and Standard Chartered.

[16] This is required for calculating a bank's own capital planning numbers. Accordingly, the main difference between the CCAR and the DFAST is that the CCAR takes bank's capital planning into account (see European Parliament (2019)).

Accordingly, the SSM stress test (see FSI (2018b, p. 1)), the BoE ACS[17] (see Bank of England (2019, p. 25)), the FINMA stress test and the JFSA stress test are bottom-up stress tests (see FSI (2018b, p. 1)). By contrast, authorities that oversee financial stability tend to conduct top-down stress tests. This is the case for the SNB building block analysis and the BoJ stress test (see FSI (2018b, p. 1)). One exception are the stress tests in the US, which are designed as a top-down exercise (see FSI (2018b, p. 1)). Regarding balance sheet assumptions and the treatment of maturing positions, all system-wide stress tests, except the SSM stress test, permit banks to assume a dynamic balance sheet, i.e., banks can make assumptions about the characteristics of new business (see FSI (2018b, p. 18)). However, the FINMA stress test, the CCAR, and the DFAST do not permit a shrinkage of balance sheets to avoid mechanically improving the capital ratios (see FSI (2018b, p. 18)). By contrast, the SSM stress test requires banks to replace maturing positions by business with identical characteristics as the matured instrument (see FSI (2018b, p. 18)).

Key metrics: The impact on capital adequacy and solvency plays a key role in all stress test programs. Therefore, the depletion in the CET1 capital and the leverage ratio is calculated in all system-wide stress tests considered here (see FSI (2018b, p. 21)). Additionally, the DFAST and the CCAR assess the Tier 1 and total capital ratios (see FSI (2018b, p. 21)). The SNB building block analysis adds the LCR to the CET1 capital and the leverage ratio in order to measure the liquidity stress (see FSI (2018b, p. 21)).

Communication of results: The communication of results of system-wide stress tests differs significantly across the eight programs. The results of the SSM stress test, the BoE ACS, the DFSAT, and the CCAR are published at the bank level (see FSI (2018b, p. 22)). By contrast, the FINMA, the BoJ, and the JFSA do not disclose their results but discuss them with banks (see FSI (2018b, p. 22)). The SNB drafts a summary of the building block analysis results in its Financial Stability Review (see FSI (2018b, p. 22)).

[17] The BoE ACS is used by the PRA for supervisory purposes.

Jurisdiction	Euro area	United Kingdom	Switzerland
Name	SSM stress test	BoE Annual cyclical scenario (ACS)	SNB building block analysis FINMA stress test
Main objective	Assess resilience to adverse conditions, promote transparency through disclosure, challenge banks' risk management, inform SREP	Assess capital adequacy as a whole and for individual institutions	SNB: Assess the resilience of the financial system, tool for communicating financial stability issues FINMA: Capital adequacy, inform SRP challenge banks' risk management
Coverage of banks:	Around 100 banks	Largest banks and building societies only (seven institutions in 2019)	SNB: Both G-SIBs and around 90 banks FINMA: Both G-SIBs
Number of scenarios	Baseline and one adverse scenario	Baseline and one adverse scenario	SNB: Baseline and three to five adverse scenarios FINMA: Baseline and two adverse scenarios
Data	Data collection at bank level, proportionality in data collection	Data collection at bank level	SNB: Regular reporting data is complemented by additional data requests, proportionality in data collection FINMA: Data collection at bank level
Concept and balance sheet assumption	Bottom-up, static balance sheet	Bottom-up, dynamic balance sheet	SNB: Top-down, dynamic balance sheet FINMA: Bottom-up, dynamic balance sheet (no shrinkage)
Key metrics	CET1 capital ratio and leverage ratio	CET1 capital ratio and leverage ratio	SNB: CET1 capital ratio, leverage ratio and LCR FINMA: CET1 capital ratio and leverage ratio
Communication of results	Public disclosure at bank level	Public disclosure at bank level	SNB: Summary in the Financial Stability Review FINMA: Shared with banks only

Table 1: Overview of selected regulatory system-wide stress testing frameworks. Own representation of FSI (2018b).

Jurisdiction	USA	Japan
Name	Dodd-Frank Act stress test (DFAST) Comprehensive Capital Analysis and Review (CCAR)	BoJ stress test JFSA stress test
Main objective	DFAST: Assess the ability of banks to absorb losses CCAR: Capital adequacy, capital planning	BoJ: Capital adequacy as a whole and for individual banks JFSA: Capital adequacy for individual banks, challenge banks' risk management
Coverage of banks:	DFAST/CCAR: Banks with total assets > USD 100bn (2019: Banks with total assets > USD 250bn)	BoJ: Around 370 banks JFSA: Four G-SIBs and three D-SIBs
Number of scenarios	DFAST: Baseline, adverse and severely adverse scenario CCAR: Baseline, adverse, severely adverse and banks forecasts for baseline and adverse scenario	BoJ: Baseline and two adverse scenarios JFSA: Not specified (various supervisory and bank-specific scenarios)
Data	DFAST/CCAR: Regular reporting data is complemented by additional data requests	BoJ: Supervisory and public (macroeconomic) data JFSA: Supervisory and public data
Concept and balance sheet assumption	DFAST/CCAR: Top-down, dynamic balance sheet (no shrinkage)	BoJ: Top-down, dynamic balance sheet JFSA: Bottom-up, dynamic (bank discretion)
Key metrics	DFAST: CET1 / Tier 1 /Total capital and leverage ratio CCAR: As DFAST, but minimum levels for key measures	BoJ: CET1 capital ratio JFSA: CET1 capital ratio, sometimes other measures
Communication of results	DFAST/CCAR: Public disclosure at bank level	BoJ/JFSA: Shared with banks only

Table 1 (continued): Overview of selected regulatory system-wide stress testing frameworks. Own representation of FSI (2018b).

2.3 Academic research on system-wide stress tests

Acknowledging the general structure of stress tests (see Figure 1), the majority of research papers focus on modeling (step 2 in Figure 1), stress test results (step 3), and the follow-up process after the completion of the exercise. Accordingly, scenario selection[18] (step 1) is not discussed as vigorously in academic circles. With this in mind, this paper summarizes (i) the literature on modeling (step 2) and (ii) the assessment of stress test results (step 3), and gives an overview of areas from the feedback process which are, to the best of the author's knowledge, the subject of the most lively debate. This covers the literature on (iii) the information value of stress tests, i.e., analyzes whether system-wide stress tests reveal new information to the market, and studies dealing with (iv) side effects of stress tests.

Modeling

First, one strand of literature focuses on the methodology of system-wide stress tests. Generally, such papers take a bird's view and deal with the general suitability of system-wide stress tests for providing meaningful results. Borio, Drehmann, & Tsatsaronis (2014) point out the limits of stress tests. They show that macroeconomic stress tests are not designed to identify vulnerabilities under normal market conditions. Thus, the authors suggest at least including non-linearities and feedback effects in the methodology. Although linear models might be adequate to capture the impact of small shocks, stress tests assume severe shocks, which are likely to lead to non-linear behavioral reactions and losses.[19] As a solution, the authors suggest regime-switching models where a stress reaction kicks in if a critical threshold is exceeded.[20] Furthermore, feedback effects within the financial sector as well as between the financial and real sectors of the economy should be part of stress tests. The authors mention the Bank of England's stress testing framework (see Aikman et al. (2009)) as a positive example, as it takes into account counterparty credit risk in the interbank market and considers feedback channels arising from market and funding liquidity. Cerutti & Schmieder (2014) argue that system-wide stress tests are based on banks' consolidated balance sheets and do not capture risks embedded in the geographical structures of banking groups. As a solution, the authors present a stress test model that takes ring-fencing and regulatory differences among countries into account.

Further studies suggest fully-fledged frameworks for system-wide stress tests. Busch, Koziol, & Mitrovic (2018) develop their own stress testing framework for German small and medium-sized banks. They emphasize that these banks play a significant role in providing credit to the real economy and that tailored stress tests for these banks are necessary. The framework consists of two elements – an income component and a credit risk component –

[18] In the euro area, for example, the European Systemic Risk Board (ESRB) provides a narrative of the most severe threats to the stability of the EU financial sector (see ESRB (2018)), which is then converted into macroeconomic and financial variables for the adverse scenarios. Although there has been some feedback in the media and press on the scenario design, academic research on this scenario decision is limited.

[19] For example, Drehmann, Patton, & Sorensen (2007) show that credit losses of a non-linear vector autoregressive model are different to credit losses using linear models, particularly for severe shocks.

[20] For example, Klein & Pliszka (2018) explain corporate bond spreads with a regime-switching models. The authors show that factors that are related to expectations about corporate earnings and, in turn, about default rates seem to have a much stronger effect on spread changes under stressed market conditions than in normal times. Put differently, the relationship between explanatory variables and the independent variable changes significantly under stress conditions and thus a linear model would not be adequate.

that are needed to calculate the stress capital ratios. The income component is estimated in a panel regression and credit risk is modeled in a CreditMetrics-style framework. The authors conclude that credit impairments (as modeled in a CreditMetrics-style framework) have a stronger effect on stress test results than the income component. Another example can be found in Buncic & Melecky (2013), who combine several approaches from the literature in order to obtain a flexible and useful stress testing framework. The authors aim to provide a model, which could also be used by policy makers and practitioners, and apply it to banks from an east European country. The main features are country-specific macroeconomic scenarios,[21] which are generated by a vector autoregressive model, and a credit risk model that considers bank-specific and asset-specific characteristics.

Overall, academic researchers propose using more granular methodologies (e.g., non-linearities, feedback effects, intra-group linkages) which come at the cost of higher data requirement and more complexity. Furthermore, there are proposals for fully-fledged system-wide stress testing frameworks – sometimes with a focus on a special group of banks, such as small and medium-sized banks.

Assessment of stress test results

Second, literature on stress test results assesses whether the measures of interest in system-wide stress tests are reasonable. Quagliariello (2019) points out that banks have higher incentives to show good stress test results than to identify actual risks. In line with this, Leitner & Yilmaz (2019) provide a study about incentives when banks are able to use their own models for regulatory purposes. This is also the case in bottom-up stress tests. Specifically, their paper uses a theoretical model to investigate the optimal level of monitoring for supervisors, as monitoring banks can provide information for supervisors but sets incentives for banks to generate less information. Interestingly, the authors conclude that, in the equilibrium state, monitoring sets incentives for banks to produce less information. Furthermore, Niepmann & Stebunovs (2018) find indications that banks optimize stress test results by taking advantage of the discretion in the bottom-up design of system-wide stress tests in the EU.[22] However, this study should be seen with the caveat that it draws on findings for banks' internal models based on public data and does not control for structural changes in the 2014 and 2016 EU-wide system-wide stress tests.

In a related context, Acharya, Engle, & Pierret (2014) compare the results of regulatory stress tests for the US and the euro area with their own stress test (V-lab stress test)[23] based on publicly available data. The authors find that the capital depletion in regulatory stress tests is uncorrelated with the V-lab stress test if the depletion is measured relative to risk-weighted assets. Thus, the authors propose relating capital depletion to total assets. Dissem & Lobez (2020) expand the idea of Acharya et al. (2014) and compare the 2014 EU-wide stress test results with various market-based measures for systemic risk, i.e., marginal expected shortfall (see Acharya, Pedersen, Philippon, & Richardson (2017)), SRISK and Delta CoVaR (see

21 This seems to be a valid aspect. The SSM stress test assumed the same shock sizes for all countries within the euro area until 2016. However, country-specific shocks have been introduced from 2018 onwards.

22 The Bank Policy Institute, however, questions the methodology in Niepmann & Stebunovs (2018) (see Covas (2018)).

23 The V-lab stress test is a variant of SRISK (see Brownlees & Engle (2017)).

Adrian & Brunnermeier (2017)).[24] The authors conclude that SRISK is most closely related to the results of the 2014 EU-wide stress test.

There are incentives for banks to minimize the impact of a scenario in stress tests. However, a comparison of results of system-wide stress tests with market-based measures for systemic risk (i.e., SRISK) shows some consistency between both proceedings – this points to some reasonableness of the results of the system-wide stress test tests.

Information value

A third strand in the literature deals with the information value and, more specifically, the market's reaction after the disclosure of stress test results. In an event study, Petrella & Resti (2013) discuss whether the disclosure of the 2011 EBA stress test results affected banks' stock prices. The authors conclude that the stress test results did indeed have a significant impact on stock prices and, thus, provided new information. Similarly, Lazzari, Vena, & Venegoni (2017) show in another event study that the disclosure of the results of the 2014 EBA stress test affected banks' stock prices. However, the paper concludes that the exercise was not able to identify troubled banks. Georgescu, Gross, Kapp, & Kok (2017) show in another event study that the disclosure of 2014 and 2016 EU system-wide stress test results revealed new information, and CDS spreads of stress-tested banks were affected by the disclosure accordingly.

Breckenfelder & Schwaab (2018) analyze how banks' stock prices reacted after the disclosure of the 2014 results of the comprehensive assessment[25] and analyze a potential spillover of bank risk to sovereigns. First, the authors conclude that banks' share prices in stress countries,[26] which were more strongly affected by the European sovereign debt crisis, declined, on average, by 12% after the disclosure. In contrast, share prices remained mostly unchanged for banks from non-stress countries.[27] Second, the authors observe a spillover of banks' risk to sovereigns (as measured by CDS spreads) in non-stress countries, as the co-movement between sovereign risk and bank risk increased for non-stress countries but remained unchanged for the stress countries after the disclosure. This points to the existence of risk-sharing across borders in the euro area in 2014.

It may be concluded that there is strong empirical evidence that the disclosure of stress test results provides new information for the market, i.e., that banks' stock prices and CDS spreads are affected by the publication of stress test results.

24 It should be noted that Delta CoVaR possesses a different interpretation than the other measures and might thus lack comparability. This is because bank-specific stress tests, the marginal expected shortfall, and SRISK measure the impact on individual banks conditional on the economy being in distress. Delta CoVaR, however, is defined the other way round. It measures the impact on the economy conditional on an individual bank being in distress.

25 The comprehensive assessment includes the 2014 EU-wide stress test.

26 Greece, Ireland, Italy, Portugal, and Spain are assumed to be stress countries.

27 Austria, Belgium, France, Germany, and the Netherlands are non-stress countries.

Side effects

Fourth, there are theoretical and empirical discussions about potentially unintended behavior as a side effect of system-wide stress testing. In a theoretical model, Gick & Pausch (2012) show that announcing the disclosure of the stress test methodology and the results can increase welfare. In this regard, the authors provide an optimum disclosure strategy and also conclude that disclosing all information about the state of the banking sector might lead to withdrawals of investors, which could, in turn, harm the economy. Similarly, the theoretical model in Goldstein & Sapra (2014) concludes that disclosure of stress test results at the aggregate level promotes financial stability and is thus beneficial. However, disclosure at the bank-level can harm risk-sharing among banks and, in turn, impair the interbank market. Moreover, disclosure of stress test results can have a negative impact on market discipline because traders have fewer incentives to collect information. In a related context, Goldstein & Leitner (2018) derive optimal disclosure strategies and discuss their link to risk-sharing opportunities. Disclosure of stress test results can have a detrimental effect on risk-sharing; this became particularly relevant in the 2007-09 financial crisis, as banks almost stopped doing business with each other.

Several empirical studies use a difference-in-differences (DID) approach to identify the change in behavior upon being included in a system-wide stress test ("treatment group").[28] Cornett, Minnick, Schorno, & Tehranian (2020) show that banks increase their capital ratios in the quarters in which a system-wide stress test is carried out. Moreover, banks that undergo a system-wide stress test lower dividends and spend more money on lobbying. Using US data, Doerr (2019) shows that banks that undergo a stress test cut small business loans for which borrowers use their homes as collateral. The reason for this is that most stress scenarios assume a large decline in residential home prices, and thus strongly reduce the value of the collateral. In turn, the lower credit supply leads to a relative decline in entrepreneurship in US counties with higher exposure to banks that are subject to the CCAR. Cortés, Demyanyk, Li, Loutskina, & Strahan (2020) also analyze the impact on lending for US stress tests. The authors show that banks affected most by stress tests shift credit from riskier to safer markets. However, in sum, banks do not reduce lending. This finding is broadly corroborated by Acharya, Berger, & Roman (2018), who show that stress-tested banks in the US reduce their credit supply, particularly to risky borrowers (small business and large corporate borrowers, commercial real estate, and credit card). Unlike Cortés et al. (2020), the authors also observe a reduction in aggregate credit supply. Moreover, the authors conclude that system-wide stress tests also lead to safer banks as measured by various capital ratios and the risk-weighted asset density (average risk weight per balance sheet unit).

Theoretical literature concludes that disclosure of stress test results can improve the welfare if an optimal disclosure strategy is chosen. Furthermore, there is evidence that the participation in system-wide stress testing exercises is positively related to increases in banks' capital ratios and lending to less risky borrowers.

[28] In this context, the approach comes with a caveat, as the behavior of stress-tested banks in the absence of the stress test cannot be observed and the best proxy available is a comparison of stress-tested banks to banks with broadly similar characteristics ("control group"). However, these banks will generally differ in more attributes than participation in the stress test, as being included in a stress test exercise is linked to specified criteria (e.g., balance sheet sums over USD 250bn in the DFAST and CCAR), and thus the assumptions of a DID approach cannot be fully met.

3 Banks' internal stress tests

This section outlines the requirements for banks' internal stress tests under the Basel III framework and gives an overview of the relevant literature.

3.1 Regulatory requirements and expectations in Basel III

As a consequence of the 2007-09 financial crisis, the BCBS developed principles for banks' internal stress testing (see BCBS (2009)).[29] These principles describe expectations for banks' internal risk management practices under the Basel framework and cover integration into the governance structure, methodologies, scenario selection, and the scope of portfolios to which stress tests should be applied. Before these principles were published, the Basel II framework contained only a limited number of references to stress tests, but no dedicated regulatory document on internal stress testing from the BCBS existed. In 2018, the BCBS revised its stress testing principles (see BCBS (2018)) and included, among several other items, principles for supervisors on system-wide stress tests. Furthermore, more stringent regulatory requirements and expectations for internal stress tests were introduced with the Basel III framework (see BCBS (2019)). Specifically, stress testing forms an important part of Pillars 1 to 3 and needs to be applied for all relevant risk types. Under Pillar 1, banks are required to carry out stress tests for credit risk, market risk, and liquidity risk.[30] Furthermore, stress testing is supposed to be part of the ICAAP and the SRP under Pillar 2.[31] Against this background, more explicit guidance is given for credit risk, market risk, interest rate risk in the banking book (IRRBB), risk data aggregation and risk reporting, and liquidity monitoring metrics in Pillar 2. Stress testing for operational risk is indirectly covered by principles for risk data aggregation and risk reporting, which require banks to have a flexible and accurate IT infrastructure (see BCBS (2019, pp. 1157-58)).[32] It is noteworthy that the level of detail differs strongly between the individual frameworks in Pillar 2. For example, very comprehensive requirements for stress testing are given for IRRBB, whereas the requirements for market risk are formulated less extensively. Under Pillar 3, information on stress tests needs to be disclosed for market risk, IRRBB, and liquidity risk. Furthermore, stress testing also plays a role in assessing the Basel core principles,

29 Based on the BCBS principles, national regulatory authorities published their expectations of stress testing in more detailed documents (e.g., FSA (2009) and CEBS (2010b)). The FSA was abolished in 2013 and its responsibilities were split between the PRA and FCA. The CEBS was succeeded by the EBA in 2011.

30 For the purpose of this paper, liquidity refers to funding liquidity and describes the ability of an entity to obtain funding. However, Brunnermeier & Pedersen (2009) show a mutually reinforcing relationship between market liquidity risk and funding liquidity risk.

31 Indeed, the Basel principles need to be implemented in individual jurisdictions. For example, the guidelines for banks that are under direct supervision of the ECB are more prescriptive than the Basel principles and have a strong focus on the ICAAP. For instance, Principle 7 of the ECB ICAAP Guide deals with stress testing (see ECB (2018)).

32 These are, indeed, the only requirements for operational risk stress tests. A reason for the limited number of requirements might be the risk-insensitivity of the regulatory standardized measurement approach (SMA) under Pillar 1. This approach is based on profit and loss numbers and thus economically meaningful relationships between stress scenarios and risk parameters of the SMA cannot be estimated. As the Pillar 1 model is not well-suited for stress testing, it is possible that banks' Pillar 2 models for operational risk might not be able to incorporate stress tests. Therefore, only a very limited number of stress testing requirements for operational risk exist. However, operational risk is included in system-wide stress testing in some jurisdictions (see BCBS (2017, p. 29)).

a benchmark for effective supervision.

Pillar 1

Credit risk: The standardized approach requires banks to consider stress tests for calculating haircuts for the collateral if historical data might understate the volatility due to illiquidity (see BCBS (2019, p. 217)). Banks using the internal ratings-based approach (IRBA banks) are required to use stress testing for capital adequacy. These banks should use stress testing for identifying unfavorable events for the banks' credit exposure. In doing so, they are required to consider at least the effect of a mild recession on the risk parameters PD, LGD, and EAD (see BCBS (2019, pp. 334-35)). Stress tests are also to be applied for equity exposures that are subject to credit risk. In these cases, stress tests should be used to gain information about tail events that are beyond the confidence level of internal models (see BCBS (2019, p. 366)). Banks applying the internal model method (IMM) for counterparty credit risk are subject to more stringent stress testing requirements. Specifically, the expected positive exposure (EPE)[33] can be estimated if they fulfill a number of requirements – stress testing is one of these (see BCBS (2019, p. 532)). Furthermore, stress testing needs to be applied to detect market factors that are positively correlated with counterparty credit risk, i.e., for measuring wrong-way risk. In cases where a bank serves as a clearing member, it is to use scenario analysis and stress testing to ensure that a sufficient amount of capital is held against its exposures to the central counterparty (CCP) (see BCBS (2019, p. 547)).

Market risk: Banks that choose to use the internal models approach (IMA) for calculating RWAs for general market risk are required, among a number of other conditions, to conduct stress tests (see BCBS (2019, pp. 624, 676-77)). The requirements for these stress tests are comprehensive and require banks to conduct qualitative and quantitative stress tests, use their own and supervisory scenarios, and to consider stress test results for internal policies and limits (see BCBS (2019, pp. 690-92)). To calculate specific market risk for portfolios that are eligible for correlation trading[34] under the IMA, banks have to apply a prescribed set of reference periods, market stress scenarios, default events, and follow the technical/methodological guidance as given in BCBS (2019, pp. 739-43). Besides these prescribed stress tests, banks have to apply their own internal stress tests for identifying meaningful stress scenarios and assessing the effects of these scenarios on the market value of the correlation trading portfolio (see BCBS (2019, pp. 743-44)).

Liquidity coverage ratio: Although the incoming and outgoing cash flows that accompany the LCR already imply a stress event, banks are expected to conduct their own stress tests, which

[33] This is the expected exposure of an instrument (i.e., OTC derivative) on a future target date conditional on positive market values.

[34] Correlation trading portfolios cover securitizations and n-th-to-default credit derivatives where the reference obligations are single-name products. However, correlation trading positions cannot be resecuritizations nor derivatives of securitizations that do not provide a pro-rata share in the proceeds of a securitization tranche. Hedges to the securitizations and n-th-to-default credit derivatives where a liquid two-way market for the instrument or its underlying exists can be included as well. However, positions referencing to an underlying that would be treated as a retail exposure, a residential mortgage exposure, or a commercial mortgage exposure under the standardized approach to credit risk as well as positions that reference a claim on a special purpose entity cannot be included (see BCBS (2019, pp. 619-20)).

are more severe than the one included in the LCR (see BCBS (2019, p. 836)).[35] These internal stress tests are to consider a bank's business activities and cover longer time horizons than the 30 days of the LCR. If a jurisdiction has an insufficient amount of Level 1 high-quality liquid assets (HQLA) (alternatively the sum of Level 1 and Level 2 HQLA) in their domestic currency to meet the banks' demand, banks in this jurisdiction can use alternative liquidity approaches.[36] In this case, jurisdictions have three options.[37] They can include liquidity facilities from the central bank (option 1), use the HQLA of a foreign currency (option 2), or use additional Level 2 assets with a larger haircut (option 3) (see BCBS (2019, p. 866)). Additional requirements for stress testing apply to options 2 and 3. If a bank uses option 2, it has to demonstrate that it can reasonably convert the foreign currency HQLA into domestic currency in times of stress (see BCBS (2019, pp. 870-71)). If a bank uses option 3 it must conduct stress tests to ensure that the HQLA remains sufficient to fulfill the LCR in a market-wide stress event (see BCBS (2019, p. 872)). Furthermore, it is to apply a haircut, which is above the minimum level of the supervisory-prescribed haircut under option 3 (see BCBS (2019, p. 872)).

Pillar 2

Supervisory Review Process: The BCBS has detected four key principles for the SRP under Pillar 2 (see BCBS (2019, p. 1051)). Principle 1 states that banks should assess their overall capital adequacy in relation to their risk profile and that they should have a strategy in place for maintaining their capital levels. In doing so, banks should use forward-looking stress tests to identify events that could adversely affect their internal capital adequacy (see BCBS (2019, p. 1051)). More sophisticated banks are required to assess internal capital adequacy for market risk based on both value-at-risk and stress testing (see BCBS (2019, pp. 1054-56)). Furthermore, the risk management processes should be reviewed periodically, also including stress testing programs (see BCBS (2019, p. 1057)). Principle 2 requires supervisors to review and evaluate banks' internal capital adequacy assessments and strategies. In accordance with Principle 2, supervisors should consider the results of banks' internal sensitivity analyses and stress tests for assessing banks' capital planning (see BCBS (2019, p. 1058)). Furthermore, supervisors should assess the degree to which a bank has considered unexpected events (including stress events) in setting its capital levels (see BCBS (2019, p. 1059)). Principle 3 and 4 are unrelated to banks' internal stress tests.[38]

Risk management principles / ICAAP: As mentioned above in Principle 1 of the SRP, banks are required to implement an internal capital adequacy assessment process. This so-called ICAAP should contain stress testing as a complementary tool to help validate other approaches (see

[35] In 2019, the ECB conducted a liquidity stress test (see ECB (2019)) in which a survival period was calculated, i.e., the number of days until the net outflow exceeded a bank's available liquidity. It turned out that 90% of euro area banks survived for a period of more than two months.

[36] As Level 1 HQLA predominantly consist of sovereign bonds and central bank reserves, fulfilling the LCR for banks in jurisdictions with a low government debt (e.g., Norway) can be more difficult. Therefore, Norway is exempted and Norwegian banks are allowed to use the alternative approaches (see European Commission (2015)).

[37] These options can also be combined.

[38] Principle 3 states that banks should generally hold capital above the minimum capital requirements ('management buffer') and Principle 4 requires supervisors to act before a bank's capital falls below the minimum capital requirements.

BCBS (2019, p. 1063)). Stress testing should be part of a bank's firm-wide risk management and should be used for identifying potential losses as well as liquidity needs under stress conditions (see BCBS (2019, p. 1064)). A bank's management information system (MIS)[39] should be flexible enough to assess the impact of forward-looking bank-wide (stress) scenarios that reflect the management's interpretation of evolving market conditions (see BCBS (2019, p. 1068)). A bank's stress testing framework should also encompass risk concentrations (see BCBS (2019, p. 1071)) and reputation risks (see BCBS (2019, p. 1073)). For liquidity risk, banks should regularly conduct stress testing to identify and quantify their vulnerabilities to liquidity in terms of cash flows, profitability, and solvency. These vulnerabilities are also required to be discussed with the management (see BCBS (2019, p. 1079)). Moreover, banks are to follow the stress testing principles (see BCBS (2019, p. 1077) and BCBS (2018)), which are designed to be applied proportionately to the size, complexity, and riskiness of a bank (see BCBS (2018, p. 5)). For liquidity risk, the BCBS published principles in 2008 (see BCBS (2008)). These require banks to run a variety of stress scenarios to detect vulnerabilities, ensure that internal risk limits are not exceeded, make sure that payment/settlement obligations can be met (see BCBS (2008, pp. 20, 24), BCBS (2019, p. 1172)), and hold a liquidity cushion (see BCBS (2008, pp. 29-30)). Stress test results are to inform strategies, policies, positions, and support the development of liquidity contingency plans (see BCBS (2008, pp. 24-25)). Supervisors are expected to assess banks' liquidity risk management (including stress testing) and to communicate more intensively with other supervisors and authorities in times of stress to increase efficiency and enhance the supervisory process (see BCBS (2008, pp. 33-35)).

Credit risk: IRBA banks have to ensure they have sufficient capital to meet the Pillar 1 capital requirements and, if applicable, the capital shortfall from the credit risk stress test (see BCBS (2019, p. 1124)). If a bank has an insufficient amount of capital, the supervisor is expected to take mitigating actions (e.g., require the bank to reduce its risk or to increase its capital) (see BCBS (2019, p. 1124)). Moreover, banks should run stress tests for credit risk concentrations, both for obligors and protection providers, to identify market conditions that can have an adverse impact on the bank (see BCBS (2019, p. 1126)). Supervisors are expected to assess banks' stress tests for concentration risk and take action if the risks are not adequately addressed by banks (see BCBS (2019, p. 1127)). Furthermore, banks must have a stress testing program for counterparty credit risk in place. The results of these stress tests should be reviewed by the management, reflected in risk policies/limits and inform mitigation strategies if vulnerabilities have been identified (see BCBS (2019, p. 1128)). Stress tests should also cover securitizations (see BCBS (2019, p. 1131)). Specifically, they should consider scenarios in which assets cannot be securitized or have to be marked-to-market, and assess the impact on liquidity, earnings, and solvency (see BCBS (2019, p. 1132)). These stress test results are, jointly with other risk management tools, to be incorporated into the risk management processes, the ICAAP, and feed into the calculation of the Pillar 2 capital requirements. If risk mitigation techniques for risks arising from securitizations are employed, stress tests are to be used to inform about the effectiveness and potential side effects of these techniques. Furthermore, stress tests are to be used to help understand the pool performance of securitizations (see BCBS (2019, p. 1132)).

Market risk: Banks using internal models for market risk have to ensure that the minimum

[39] This term refers to procedures that collect and produce data for supporting the decision-making process.

capital requirements under Pillar 1 also cover market risk stress tests (see BCBS (2019, pp. 1141, 1054-56)).[40] If there is a shortfall or if supervisors are not satisfied with the banks' calculations, they are expected to take mitigating actions (e.g., require a bank to reduce its risk or to increase its capital) (see BCBS (2019, p. 1141)).

Interest rate risk in the banking book: Banks should have an IRRBB stress testing framework in place that feeds into the ICAAP (see BCBS (2019, p. 1100)), the reporting of internal risk assessments to the management (see BCBS (2019, p. 1099)), the broader risk management, and governance processes (see BCBS (2019, p. 1090)). Within this framework, banks have to consider a broad range of interest rate scenarios[41] for which selection methods need to be endorsed by the governing body or its delegate (see BCBS (2019, p. 1085)). Banks need to be able to compute internal and supervisory prescribed risk measures for historical and hypothetical stress scenarios, and to incorporate supervisory-imposed constraints on banks' internal risk parameters (see BCBS (2019, pp. 1090, 1097)). The stress testing framework should be proportional to the extent of IRRBB (see BCBS (2019, p. 1091)). It should include clearly defined objectives, scenarios should be tailored to a bank's business and risks, have well-documented assumptions, and use sound methodologies. The stress testing framework should be used to assess the impact upon the bank's financial condition, have review processes, and be able to derive actions based on stress results. Experts from various divisions (e.g., trade, treasury, risk management) are to be involved in the stress test exercise. Moreover, the communication of risks from stress test results is to take place with internal and external stakeholders (see BCBS (2019, p. 1091)). Reverse stress tests are to be performed qualitatively and quantitatively (see BCBS (2019, pp. 1092-93)). The aim of these is to identify interest rate scenarios that can adversely affect the capital or the earnings and are intended to reveal vulnerabilities from hedging strategies and potential reactions of customers. Supervisors are expected to collect data from banks to assess banks' IRRBB. In doing so, they can collect and review information on the effects of various stress scenarios on earnings and economic value (see BCBS (2019, p. 1101-02)), and assess the effectiveness of the banks' IRRBB stress testing framework (see BCBS (2019, p. 1103)). If a bank's economic value is significantly sensitive to a shock or stress scenario, supervisors are required to assess the impact on capital resulting from banking book positions held at market value and from using market valuation for positions held at historical cost (see BCBS (2019, p. 1105)).

Liquidity monitoring metrics: In addition to LCR and NSFR, the BCBS developed other liquidity measures that are subject to stress testing and that banks have to monitor. First,

40 Illiquidity, concentration risk, one-way markets, non-linearities, jumps-to-default, and correlation risk should, where appropriate, be covered. Furthermore, other risks (e.g., recovery rate uncertainty) that may not be adequately captured in the value-at-risk shall be considered (see BCBS (2019, pp. 1141, 1055)).

41 Scenarios are to be applied for both earnings and economic value (see BCBS (2019, p. 1090)), and cover a wide range of parallel and non-parallel interest rate changes, effects of concentrations in liquidating positions, interactions of IRRBB with other risks, changes in spreads of maturing and new business, execution of interest rate options (i.e., interest rate level and volatility changes), changes in administered interest rates, and basis risk (see BCBS (2019, pp. 1091-92)). Basis risk in the context of IRRBB means the risk of incongruities between different currencies, reference rates, and maturities. In addition, scenarios should be forward-looking and consider changes of internal (e.g., business model) and external factors (e.g., competition) (see BCBS (2019, p. 1092)). Banks are to consider meaningful shock and stress conditions, including a potential breakdown of key assumptions under stressful market conditions, and take them into account when setting internal policies and limits (see BCBS (2019, pp. 1087, 1090)).

funding concentrations are to be assessed, as they might be more likely to be withdrawn in times of stress (see BCBS (2019, pp. 1173-74)). Second, currency-specific LCRs[42] (see BCBS (2019, p. 1177)) are to be monitored to track potential currency mismatches of incoming and outgoing cash flows, which could have adverse effects in times of stress. Third, banks are to assess their intraday liquidity under (i) own financial stress scenarios, (ii) counterparty stress scenarios, (iii) customer's bank's stress scenarios, and (iv) market-wide credit/liquidity stress scenarios (see BCBS (2019, pp. 1196-98)). The impact of these stress scenarios on banks' normal liquidity profile is to be evaluated and discussed with supervisors (see BCBS (2019, pp. 1198-1200)). Furthermore, positions from different currencies, although managed on a cross-currency basis, can be aggregated for reporting purposes if the bank is able to convert them reasonably to the reporting currency in times of stress (see BCBS (2019, p. 1184)).

Risk data aggregation and risk reporting: Banks should have internal risk reporting which allows them to perform stress testing flexibly and effectively, and provide a forward-looking risk assessment (see BCBS (2019, pp. 1151, 1162)). Specifically, risk data reports should be generated in normal and stress periods, be accurate, include stress test results, be provided in a timely manner, and be able to incorporate ad hoc requests (see BCBS (2019, pp. 1153, 1159, 1161, 1162, 1164)). Moreover, frequency should be increased during stress periods (see BCBS (2019, p. 1164)). It is noteworthy that these prescriptions concern operational risk that may arise due to an unreliable IT infrastructure. Supervisors are expected to assess a bank's capabilities to aggregate data and produce reports during times of stress (see BCBS (2019, p. 1165)).

Pillar 3 and Basel core principles

Disclosure: Banks have to disclose qualitative information about how stress tests relate to risk management objectives and policies (see BCBS (2019, p. 1253)). Besides this general integration of stress tests, banks have to provide disclosures on stress tests for market risk, IRRBB, and liquidity risk. For market risk, banks need to disclose qualitative information on stress tests applied to parameters (see BCBS (2019, p. 1449)). For IRRBB, banks have to describe interest rate shock/stress scenarios that are used for calculating changes in the economic value of equity and earnings. Moreover, they have to disclose information about the overall IRRBB risk management and mitigation strategies (see BCBS (2019, p. 1465), which are expected to encompass stress tests as well (see BCBS (2019, p. 1089)). Regarding liquidity risk, banks have to explain qualitatively how stress testing is used in their liquidity risk management (see BCBS (2019, p. 1494)).

Basel core principles: The Basel core principles constitute a minimum standard for regulation and supervisory practice in jurisdictions.[43] In this function, they are assessed as part of the FSAP to test the effectiveness of a jurisdiction's supervisory framework. Supervisors are required to use results of banks' internal stress tests (see BCBS (2019, p. 1564)), ensure that

42 This seems to be an emerging risk, as the IMF warned in IMF (2019, pp. 61-79) that mismatches of assets and liabilities in USD for non-US banks are increasing and might be a source of financial instability.

43 Demirgüç-Kunt, Detragiache, & Tressel (2008) show that banks in countries with a good implementation of the Basel core principles receive more favorable ratings. Interestingly, Demirgüç-Kunt & Detragiache (2011) do not find a significant relationship between the implementation of the Basel core principles and (idiosyncratic) banks' risks (as measured by the Z-score).

banks have adequate internal information systems under normal and stress periods (see BCBS (2019, p. 1579)), and have to ensure that banks prepare contingency plans for normal and stress periods (see BCBS (2019, p. 1580)). Moreover, supervisors should require banks to have a stress testing framework as part of their risk management in place (see BCBS (2019, p. 1581)). Specifically, supervisors have to ensure that banks carry out internal stress tests for credit risk (see BCBS (2019, p. 1587)), risk concentrations (see BCBS (2019, p. 1591)), country and transfer risk (see BCBS (2019, p. 1594)), market risk (see BCBS (2019, p. 1596)), IRRBB (see BCBS (2019, p. 1598)), liquidity risk (see BCBS (2019, pp. 1599, 1601, 1602)), and operational risk (see BCBS (2019, p. 1603)).[44] Furthermore, supervisors can require banks to run stress tests to set higher capital levels or to determine contingency arrangements in anticipation of adverse events (see BCBS (2019, p. 1583)).

3.2 Academic research on banks' internal stress tests

Academic papers on banks' internal stress tests cover the process steps of a stress test as shown in Figure 1. Accordingly, studies focus on approaches for selecting stress scenarios (step 1 in Figure 1), provide models for individual risk types or holistic stress testing frameworks (step 2), and assess the results of stress tests (step 3). Indirectly, a follow-up process is also considered, as most papers provide recommendations on stress testing practices and thus might help improving banks' risk management.

This section contains a survey of literature in five research areas, including among other things, the three above-mentioned process steps. Specifically, an overview of the literature is given on the (i) selection of stress scenarios, (ii) stress test models for credit risk, (iii) model and estimation risk in stress tests, (iv) reverse stress tests, and (v) liquidity stress tests. Area (i) provides insights into what criteria and what aspects should be considered when choosing stress scenarios, thus focusing on step 1. Areas (ii) to (v) are more technical and focus on modeling (step 2). Moreover, several reviewed studies can also be allocated to area (v), as they demonstrate an application to stylized portfolios and show stress test results. However, it should be emphasized that area (v) is very strongly related to the assessment of stress test results (step 3). The reason for this is that comprehensive simulation frameworks, which affect the amount and timing of banks' cash flows, are developed. These frameworks allow us to evaluate the impact of internal or external (liquidity) limits/policies on banks' liquidity and solvency risk in times of stress. Accordingly, this makes such frameworks suitable candidates for evaluating the effectiveness of these limits/policies and places an emphasis on the stress test results.

Selection of stress scenarios

First, literature deals with the selection of appropriate stress scenarios. This is particularly relevant for internal stress tests, as banks are expected to apply sufficiently severe, relevant, and varied scenarios (see BCBS (2018, p. 6)). Abdymomunov & Gerlach (2014) emphasize the central role of interest rate risk[45] in a low-interest rate environment. They propose a

[44] The requirements for operational risk mainly refer to the IT infrastructure, which should work appropriately in normal and stress periods.

[45] In this context, Foos, Lütkebohmert, Pliszka, & Markovych (2017) show that the interest rate sensitivity (as measured by stock prices) increased for euro area banks in the years after the financial crisis.

tool based on the Nelson-Siegel interest rate model for generating scenarios with a wider variety of slopes and shapes than the regulatory 200 bps parallel scenarios of Basel II for the standardized outlier test. The authors identify a valid aspect, as the BCBS expanded the parallel shocks to six scenarios and covered a wider variety of interest rate changes (see BCBS (2019, pp. 1106-08)). Still, Cerrone, Cocozza, Curcio, & Gianfrancesco (2017) propose using Monte Carlo methods for generating interest rate shock scenarios, which cover a wider variety of interest rate movements – even compared with the six new regulatory scenarios. The authors argue that a static set of scenarios might not be sufficient and that identifying interest rate scenarios, which are significant for business and consistent with market conditions, is crucial in order to properly assess and manage interest rate risk in the banking book.[46]

Flood & Korenko (2015) provide a methodology based on eigenvectors of the variance-covariance matrix for market risk factors to generate equally likely scenarios based on the Mahalanobis distance – a plausibility measure based on the (multivariate) distance of a stress scenario from the expected (baseline) scenario. The authors argue that using a range of scenarios with the same likelihood reduces the risk of overlooking relevant scenarios.[47] Packham & Woebbeking (2019) provide a factor model for stress testing correlations of risk factors. By using the Mahalanobis distance, this approach can identify the most severe scenario conditional on a given plausibility, i.e., the correlation structure between the risk factors that maximizes the portfolio loss.[48] In addition to correlation stress, the framework integrates volatility shocks; both elements usually increase in times of stress. The authors conclude that considering correlations in stress testing is particularly relevant for large portfolios, e.g., portfolios of central counterparties, as diversification benefits from tranquil periods might overestimate the diversification benefits during times of stress.

The findings in the academic literature emphasize the key role of selecting meaningful stress scenarios. In particular, academic research calls for using algorithms that consider historical characteristics like the correlation structure of the risk factors. These scenarios can sensibly complement scenarios that are selected by expert judgment in banks.

Stress test models for credit risk

Second, papers deal with the incorporation of stress scenarios in credit risk models. In this context, most academic papers are based on CreditPortfolioView (see Wilson (1997a,b)), a credit portfolio model that links the systematic variation of default rates to a set of macroeconomic variables. Macroeconomic variables are assumed to follow autoregressive processes, which feed into a logit equation for the default rate. This link allows us to compute stress default rates for given (macroeconomic) stress scenarios. For example, Boss (2002), Virolainen (2004), Sorge & Virolainen (2006), J. Wong, Choi, & Fong (2008) and Jokivuolle, Virolainen,

46 In this regard, it should be noted that BCBS (2019, pp. 1091-92) requires banks to develop their own interest rate stress scenarios beyond the six prescribed scenarios of the standardized outlier test.

47 The authors classify this approach between regular und reverse stress testing. Unlike in reverse stress tests, where a portfolio loss level is set, the paper specifies a likelihood for a scenario.

48 Indeed, this characteristic also makes this approach suitable for reverse stress testing. However, given that a similar methodology was presented in earlier studies (e.g., Breuer, Jandačka, Rheinberger, & Summer (2008)), and considering the clear focus on the application of the methodology for correlation stress scenarios, this paper has been classified as a study focusing on scenario selection.

& Vähämaa (2008) discuss the application of stress testing in CreditPortfolioView.

As an extension, Gaglianone & Schechtman (2012) present a variant of CreditPortfolioView where the processes for the macroeconomic variables are estimated with quantile regressions. The authors compare this extended model with the traditional CreditPortfolioView concerning Brazilian data covering the period 1995-2009. They conclude that the variant with quantile regressions yields more conservative estimates, i.e., tends to forecast higher default rates under the same macroeconomic scenario. The approach in Kanas & Molyneux (2018) is another extension based on CreditPortfolioView. The authors expand the original model by including non-linear effects of macroeconomic variables using semi-parametric quantile regression for aggregate US data from 1984 to 2013. The authors conclude that the macroeconomic variables are significant in their specification and that these significances are lost if they switch to a linear quantile regression model. Therefore, linear models might not provide reliable estimates. Jokivuolle & Virén (2013) apply a CreditPortfolioView-style framework to default rates and Loss Given Defaults (LGDs) using Finnish data ranging from 1989 to 2008. They argue that LGDs are counter-cyclical, i.e., positively correlated with default rates, and keeping LGDs constant would underestimate losses in times of stress. Besides CreditPortfolioView-style models, several other approaches are discussed. One example is the Bank of England's holistic stress testing framework (see Aikman et al. (2009) and Burrows, Learmonth, & McKeown (2012)) in which the credit risk element is modeled with a Bayesian vector autoregressive model (BVAR). Another example can be found in Düllmann & Kick (2014).[49] Using a vector autoregressive model, the authors shock sector-specific production indices by assuming a shock in the spread of borrowing and lending interest rates. Put differently, the increase in the spread leads to skyrocketing funding costs, which, in turn, reduces production. Then, a CreditMetrics-style model is used to estimate the impact on default probabilities for bank credit portfolios based on the falls in sector-specific production. Koziol, Schell, & Eckhardt (2015) discuss the impact of the dependence structure of risk factors on banks' capital ratios in credit risk stress test models. In using a Merton-style model, it turns out that heavy-tailed copulas, i.e., Clayton copulas and *t*-copulas, lead to a higher capital depletion for less severe scenarios, whereas a Gaussian copula leads to a higher capital depletion for extreme stress scenarios.[50] Therefore, the authors conclude that no copula can be assumed as the best selection and that a variety of dependence structures need to be investigated.

The literature shows that conventional credit risk models are often adjusted to incorporate stress testing. However, given the discretion for bank's modeling practices under Pillar 2, the heterogeneity of models used in the banking practice will likely be greater than in the academic literature.

Model and estimation risk in stress tests

Third, the literature discusses model and estimation risk in stress tests. Grundke, Pliszka, & Tuchscherer (2020) show that mapping macroeconomic risk factors in a CreditPortfolioView-style model is exposed to significant model and estimation risk. The paper estimates that

[49] Although this study uses data of 24 German banks, it can also be applied to individual banks.

[50] In a related context, Grundke & Pliszka (2018) show that heavy-tailed copulas (Clayton and *t*-copula) are more appropriate for fitting a stress test model to historical observations.

the initial default probabilities can increase by +20% to +80% – depending on the selected stress test model. Furthermore, the study shows that model and estimation risk are higher when mapping a stress scenario onto default probabilities, compared to non-stress scenarios. Vilsmeier & Siemsen (2018) also discuss model and estimation risk in credit risk stress tests and provide mitigation techniques for model uncertainty. Similar to Grundke et al. (2020), the authors find a significant dispersion between various stress test models for credit risk. Based on an autoregressive time series model, which explains default probabilities by its own lagged values and macroeconomic variables, the authors generate a distribution of forecasts for stress default probabilities by drawing numerous models with different specifications (e.g., number of included variables, different lag lengths, etc.). This leads to predictions of default probability increases between -90% and +7,000%. In the next step, they apply several plausibility criteria to filter out default probabilities generated by models with economically and statistically undesired characteristics (multicollinearity, autocorrelation, sign restrictions). Then, they estimate a Merton-style benchmark model to exclude those default probabilities from the autoregressive time series model that differ too strongly from the benchmark model. In sum, the filters in the two stages reduce the model uncertainty and reduce the predictions of default probability increases to a range between +60% and +150%. Kupiec (2018) compares the forecasting ability of bank performance in various stress test models. He hypothesizes that stress tests based on stepwise regression, which are also used by the Federal Reserve, might lack forecast ability. These models are selected based on the model fit within the estimation period and thus have a high in-sample fit, are parsimonious, and often assume that the selected variables point to the economically expected direction. The downside is, however, that these models can have issues with out-of-sample forecasts because they usually include bank characteristics (e.g., balance sheet information) as explanatory variables. Since bank characteristics are usually assumed to be constant over the stress scenario, the stress test model might lack sensitivity to macroeconomic factors.[51] He concludes that stepwise regression approaches prove to have high forecasting errors for a bank's performance and should therefore be supplemented by methods from information theory and machine-learning techniques. Furthermore, he calls for the development of reliable methods for validating the accuracy of stress test models.

Lin & Surti (2015) discuss the sensitivity of capital requirements for OTC derivatives that are traded with central counterparties. The authors conclude that capital requirements are highly sensitive to whether model parameters are calibrated on a point-in-time or on a stress period basis. The latter leads to higher capital requirements and is thus more likely to cover losses during periods of stress. Hence, the paper emphasizes the importance of regulators mandating the use of stress period inputs in risk models of central counterparties. Cetina, Paddrik, & Rajan (2018) assess the sensitivity of stress tests for counterparty credit risk. Similar to the approach in the CCAR, the study assumes the default of the largest counterparty of a bank and finds that second-round effects of defaults through the bank's other counterparties may be larger than the initial impact on the default of the largest counterparty. This calls for further internal stress tests for counterparty credit risk which go beyond simply assuming that the biggest counterparty defaults.

In a related context, some studies argue that calculating stress test losses for individual risk

[51] Borio et al. (2014) points out that stress tests should incorporate feedback effects between the real economy and banks.

types may understate the effects of stress scenarios – this issue is, however, not exclusively relevant to stress tests. For example, Pillar 1 of Basel III requires banks to calculate capital requirements for each risk type separately and then to sum them up. Drehmann, Sorensen, & Stringa (2010) propose a framework for jointly measuring credit risk and interest rate risk in a stress testing framework that depends on macroeconomic factors. The authors show that ignoring the interaction between credit risk and interest rate risk is quantitatively significant for stress tests and can underestimate the fall in net profits by over 50% in the first year but overestimate the fall by almost 100% in the third year. Without directly referring to stress tests, Breuer, Jandačka, Rheinberger, & Summer (2010) argue that for foreign currency loans a separate calculation of economic capital for market risk and credit risk may significantly underestimate the true risk.

These findings suggest that projecting the impact of a stress scenario onto risk parameters is exposed to significant model and estimation risk. This calls for extensive robustness checks.

Reverse stress tests

Fourth, several papers discuss the application of reverse stress tests, which have been introduced in the BCBS stress testing principles (see BCBS (2009)). Reverse stress tests look for the scenarios that correspond to a given stress test result. This process is usually accompanied by numerical and computational problems, as an inversion problem and an optimization problem need to be solved.[52] Against this background, research activities focus on technical approaches for identifying reverse stress scenarios.

Breuer et al. (2008) and Breuer, Jandačka, Mencia, & Summer (2012) measure the plausibility of a stress scenario by its distance from the expected scenario with the Mahalanobis distance.[53] Then, for a given distance, the authors search for the most adverse scenario, i.e., the scenario which maximizes the expected loss. McNeil & Smith (2012) introduce the "most likely ruin event" (MLRE) based on the concept of depth of a probability distribution as a way to solve the inversion problem of reverse stress tests. Glasserman, Kang, & Kang (2015) provide a methodology for market risk that allows identifying the most likely scenario exceeding a given portfolio loss. First, the authors estimate confidence regions for (on a given loss) conditional means of the risk factors by empirical likelihood. Second, they use scaling factors for risk factors, which are derived from the heaviness of the tails of the market risk factors, to identify the most likely reverse stress scenario. This procedure is demonstrated on simple market portfolios and requires a sufficiently high number of observations that exceed the given loss. Kopeliovich, Novosyolov, Satchkov, & Schachter (2015) draw scenarios based on principal components of risk factors to derive candidate scenarios leading to the same portfolio loss. Then, they select only those scenarios that meet (relative) plausibility criteria and that are composed only of the most important principal components. Although this approach allows the authors to identify the most likely reverse stress scenario, they argue that taking a single scenario might disregard important information and, in turn, negatively affect banks'

52 These technical difficulties might be a reason why reverse stress tests were neglected in the euro area in the first few years (see EBA (2013, p. 4)). However, this calls for further research on reverse stress tests in order to provide more information and to make them more appealing for real-world applications.

53 Čihák (2007) calls the idea of fixing a probability level and searching for the scenario that yields the most negative impact on a measure of interest the "worst case approach".

decision-making. Skoglund & Chen (2009) employ the Kullback information criterion that gives information on the relative importance of risk factors and thus helps in selecting relevant risk factors for reverse stress tests. The authors demonstrate the application of the methodology on a stylized equity derivatives portfolio.

Grundke (2011) proposes to use a CreditMetrics-based bottom-up model for reverse stress tests. This model considers interest rate risk, default risk, rating-specific credit spreads, and takes into account correlations between these risk factors. In performing the reverse stress test, the paper searches for the most likely scenario that exhausts the bank's economic value of equity. This is done by a grid search, which, however, requires the evaluation of the whole probability space of the risk factors and thus might lack computational tractability if more risk factors are used. The model is expanded in Grundke (2012) by, among other things, a time-varying bank rating, contagion effects between obligors, and correlations between the LGD and other risk factors. Grundke & Pliszka (2018) empirically implement the model from Grundke (2011), add some more realistic assumptions (e.g., hedging via interest rate swaps) and suggest methodologies (e.g., principal components) for making reverse stress tests computationally more tractable. Although the authors conclude that the model yields realistic results, they summarize that reverse stress tests require several robustness checks due to considerable model and estimation risk.

Given these findings in the academic literature, it can be concluded that reverse stress tests can add value to banks' risk management. However, reverse stress tests require long time series with stress data, are exposed to model and estimation risk, and likely require dimension reduction techniques to be computationally tractable.

Liquidity stress tests

Fifth, stress tests deal with funding liquidity.[54] BCBS (2013) presents a survey on liquidity stress testing and covers, inter alia, best practices of large banks. The survey states that banks use maturity mismatches of incoming and outgoing cash flows as their key liquidity risk measure and assume a stress period of six to twelve months. Stress scenarios are rating downgrades of the bank itself, which may be complemented by deposit withdrawals, credit losses, wholesale funding run-offs, operational risk events or other incidents. In addition, banks started combining liquidity stress events with other risk types (e.g., credit risk). Results of liquidity stress tests serve many purposes. The results support risk management, are used for calibrating risk limits/liquidity buffers, and they inform the Funds Transfer Pricing (FTP).[55] In

[54] The research area on the impact of a drop in market liquidity (i.e., the stress event) is covered in a different research area. For example, Soula (2017) assesses which banks' share prices react most strongly to a decline in market liquidity (as measured by the spread between bank and government interest rates) for euro area banks in the period 2005-12. Unsurprisingly, the author concludes that larger and better capitalized banks are less vulnerable to declines in market liquidity. In a related context, Dombret, Foos, Pliszka, & Schulz (2019) analyze the impact of market liquidity declines on the volume and credit spreads of loans in the euro area. The paper finds a strong positive relationship of market liquidity to loan volumes, and a strong negative relationship to credit spreads. Furthermore, this paper concludes that these relationships are stronger during stress periods and that a decline in market liquidity is more significant than an increase.

[55] The FTP is a mechanism which allocates earnings and costs of instruments (e.g., deposits) between different business units within a bank. It is generally used by big banks and makes possible product pricing and measures the profitability of individual business units.

referring to ECB (2008), stress tests are performed at various aggregation levels noting that the data availability hampers the quality of liquidity stress tests at higher aggregation levels, i.e., group level.

Several papers propose comprehensive simulation frameworks for liquidity stress testing, sometimes also evaluating the effects of complying with LCR and the NSFR. Although these papers include multiple banks, they apply to individual banks and thus suitable for banks' internal liquidity stress testing. However, it is important to emphasize that carrying out a comprehensive simulation-based liquidity stress test requires many assumptions. In particular, assumptions on a bank's reaction to stress scenarios (e.g., reinvestment of maturing business) are crucial and have a strong impact on the results. van den End (2012) provides a simulation approach that allows assessing the one-month impact of market shocks and funding liquidity shocks on banks' balance sheets. Using data for 85 Dutch banks from 2004 to 2010, this paper shows that negative effects (as measured by market disturbances and idiosyncratic reputation effects of banks' reactions to the shocks) of liquidity stress scenarios can be mitigated if banks fulfill the Basel III liquidity measures. A liquidity buffer as given by the LCR appears to work particularly well. Interestingly, the paper concludes that the impact of liquidity shocks on lending is limited.[56] Grundke & Kühn (2020) design a model that takes credit risk, interest rate risk, and liquidity risk into account, and simulate evolutions of stylized balance sheets over a three-year risk horizon. It turns out that the LCR and the NSFR have no clear impact on balance sheet growth and the equity return but reduce the (equity- and liquidity-driven) default risk and close gaps for maturities under one year (at the cost of negative cash flows for maturities above one year). Moreover, complying with the liquidity measures proves to be particularly difficult in times of stress.[57] Another example of a liquidity stress test is the simulation study of E. Wong & Hui (2009), who employ publicly available data of Hong Kong banks for the period 2006-08. In analyzing a forecast horizon of one year, the authors specify processes for market, credit, and liquidity risk that transmit the effects of stress scenarios onto banks' balance sheet positions. Then, liquidity measures are calculated from which liquidity shortfalls and default dates can be derived. The results call for holding liquid assets and underline the importance of considering interactions between liquidity and credit risk. Aikman et al. (2009) show how funding liquidity is modeled as part of the Bank of England's stress testing framework. The framework is applied to UK banks, employs approximately 650 balance sheet positions, and covers a three-year risk horizon. The first step projects the impact of a scenario on a bank's rating and, in turn, on the funding costs. Depending on bank-specific indicators (e.g., Tier 1 capital ratio), access to funding markets can be impaired for some banks. Therefore, a bank can fail, which can lead to adverse feedback effects for other banks (e.g., contagion effects, mark-to-market losses). In conclusion, the authors emphasize the importance of an amplifying effect of liquidity risk on other risks.

In sum, similarities and differences between banks' practices in BCBS (2013) and academic liquidity stress testing frameworks can be observed. Specifically, there are common approaches,

[56] This result contradicts the findings of de Haan & van den End (2013) who conclude that banks reduce lending after a funding liquidity shock. In a related context, Dombret et al. (2019) show a strong relationship between market liquidity and lending.

[57] This finding might be one reason why the LCR framework accounts for the possibility of falling below the minimum level of 100% in times of system-wide stress (see BCBS (2019, pp. 837, 839)).

i.e., gaps or ratios of incoming and outgoing cash flows for one or more points in time are used to measure liquidity risk. However, discrepancies exist for the length of the risk horizon. Banks use six to twelve months, whereas academic papers cover a range between one month and three years. Moreover, banks mostly carry out pure liquidity stress tests, whereas academic research embeds liquidity risk in an integrated framework with market and credit risk.

4 Differences between system-wide and banks' internal stress tests

Conceptually system-wide and bank's internal stress tests follow the same process steps (see Figure 1), i.e., a scenario needs to be selected and mapped onto bank-specific parameters. Then, these parameters are used to calculate the impact on the metric of interest; a follow-up process concludes the exercise. However, given different purposes and addressees, there are substantial differences in the specific execution of the two stress test types; Table 2 gives an overview of their main distinguishing features.

	System-wide stress tests	Banks' internal stress tests
Standardized set-up	Yes	No
Predetermined set of stress scenarios	Yes	No
Reverse stress tests	No	Yes
Focus on total capital impact	Yes	No
Predefined follow-up actions	Yes	No

Table 2: Main differences between system-wide stress tests and banks' internal stress tests.

Standardized set-up: The execution of system-wide stress tests follows a standardized set-up. The responsible central bank/supervisory authority defines macroeconomic scenarios, sets a timeline, provides data templates, and drafts instructions. Then, banks have to fill out a data template and send it back to the central bank/supervisory authority. Following a quality assurance process, the central bank/supervisory authority then uses the stress test results for its supervisory mandate. In contrast, most of the regulatory requirements and expectations for banks' internal stress tests are principle-based. This means that stress testing is to be part of banks' risk management and used for identifying potential losses and liquidity needs under stress conditions (see BCBS (2019, p. 1064)). However, banks can decide on the specific implementation and application of these principles.

Predetermined set of scenarios: One distinguishing feature of the two types of stress tests is that scenarios are prescribed in system-wide stress tests, whereas banks have to choose appropriate stress scenarios in their internal stress tests. Against this background, system-wide stress tests are composed of common risk factors that affect all banks. In internal stress tests, however, banks have to select scenarios that are specifically tailored to their business model and address their individual vulnerabilities (see BCBS (2018, p. 6)).[58]

[58] For example, a default of a central counterparty would not (directly) affect most banks but could have a large impact to individual banks.

Reverse stress tests: Quantitative reverse stress tests have so far not formed part of a system-wide stress test. This might be due to numerical and computational issues (see Section 3.2), which set high practical requirements for the quality assurance process. Another reason is that the results of reverse stress tests are difficult to compare across banks, as the outcomes are scenarios composed of different risk factors with different probabilities of occurrence. This strongly complicates the use of reverse stress tests in a system-wide stress test and hampers horizontal comparisons. However, reverse stress tests are well-suited for internal purposes and can inform the banks' risk management under Pillar 2. For instance, they can help to identify material risk factors in the risk inventory or provide insights for designing meaningful bank-specific (regular) stress scenarios.

Focus on total capital impact: System-wide stress tests assess the resilience of banks to adverse market conditions. In doing so, they compare a variety of banks by taking highly aggregated key risk metrics (e.g., the depletion in the CET1 capital ratio). By contrast, internal stress tests are often applied to smaller portfolios, and less aggregated risk metrics are assessed. This allows banks to analyze more granular questions, calibrate internal risk limits, and assess metrics that are unrelated to solvency and liquidity measures (e.g., ability to provide payment services). Indeed, this does not preclude banks' internal stress tests from being used to assess firm-wide solvency or liquidity (see Section 3.1).

Predefined follow-up actions: Since the primary addressees of stress tests are the supervisor and the banks' own management, the follow-up process and actions differ strongly between both two types of stress tests. Supervisors can take supervisory measures and raise capital requirements, follow-up on identified deficiencies, and disclose the stress test results. In contrast, banks' internal stress test results are only expected to be reported to the governing body and used effectively (see BCBS (2018, p. 5)). Accordingly, it is a bank's own decision and responsibility which actions it may derive from the results of a stress test. For example, it can change the riskiness of its business, adjust internal risk limits, voluntarily disclose results, revise models or change the funding structure.

5 Conclusions

System-wide stress tests have gained in importance after the 2007-09 financial crisis and are continuously evolving. Nowadays, comprehensive stress test exercises are performed regularly in major jurisdictions in order to assess the resilience of banks to adverse conditions. However, supervisors are considering the ways in which stress tests can be integrated best into the regulatory framework. In this regard, US authorities are discussing options for substituting the capital conservation buffer with stress test results, and the EU intends to switch to an approach in which both banks' and supervisory calculations are taken into account.

Overall, system-wide stress testing frameworks in different jurisdictions have a lot in common. Their main purpose is to assess the capital adequacy of individual banks and the financial system as a whole. In this function, stress test results are used to determine capital requirements (P2G) in the EU and affect capital planning in the US, for example.

Moreover, stress test results represent a key element for challenging banks' risk management and to identify banks' vulnerabilities. One key distinguishing factor in the implementation of system-wide stress tests in jurisdictions is the disclosure of results. Stress test results are disclosed at the individual bank level in the EU, the UK, and the US, whereas Switzerland and Japan do not disclose stress test results for individual banks.

Research on the modeling practices in system-wide stress tests detects some areas for improvement (e.g., consideration of non-linearities and feedback effects), which come at the cost of greater complexity in the models. In addition, incentives for banks to minimize the impact of a scenario in a (bottom-up) stress test are discussed. The circumstances concerning the data do not allow external researchers to assess the appropriateness of banks' models. Yet, approaches for measuring banks' vulnerability to systemic risk based on market data (e.g., SRISK) show results that are to some extent consistent with results obtained from system-wide stress tests. In all cases, the publication of the results of system-wide stress tests provides new information and affects stress-tested banks' stock prices and CDS spreads. Stress tests can, however, affect banks' behavior. Specifically, stress-tested banks tend to increase their capital and shift their portfolios from riskier to less risky borrowers.

As a consequence of the 2007-09 financial crisis, the BCBS developed principles for banks' internal stress testing and deeply embedded stress testing in the Basel III framework. In this context, banks have to perform stress tests under Pillars 1, 2 and 3 for all major risk types. Furthermore, banks that opt for internal models for credit and market risk are subject to stricter requirements on internal stress tests, i.e., they have to ensure that potential capital shortfalls from credit risk and market risk stress tests can be covered with Pillar 1 capital.

Academic research provides valuable insights for banks' internal stress tests. Specifically, algorithms that consider historical characteristics such as the correlation structure of the risk factors should be used to select meaningful stress scenarios. Another research finding is that conventional models can be modified in such a way that they can be used for stress testing. Furthermore, stress testing is generally exposed to considerable model and estimation risk, which calls for extensive robustness checks. Reverse stress tests are accompanied by numerical and computational problems for which studies provide solutions. Approaches for stress testing liquidity provide comprehensive frameworks which allow the assessment of risks arising from illiquidity. In sum, these papers show that complying with regulatory or internal liquidity risk limits lowers the probability of defaults.

Although internal risk management and stress testing should benefit from the experience of system-wide stress tests (see BCBS (2018, pp. 4-5)), there are substantial differences between the two types of stress tests. System-wide stress tests follow a standardized procedure, which allows supervisors to carry out horizontal comparisons and assess the resilience of individual banks to adverse macroeconomic conditions. By contrast, the principle-based regulation of internal stress tests gives banks the discretion to carry out stress test that are specifically tailored to a bank's business and risks. This means that both stress test types play a complementary role in ensuring the resilience of individual banks and of the financial system to adverse shocks.

References

Abdymomunov, A., & Gerlach, J. (2014). Stress testing interest rate risk exposure. *Journal of Banking & Finance*, *49*, 287-301.

Acharya, V., Berger, A. N., & Roman, R. A. (2018, April). Lending implications of U.S. bank stress tests: Costs or benefits? *Journal of Financial Intermediation*, *34*, 58-90.

Acharya, V., Engle, R., & Pierret, D. (2014, July). Testing macroprudential stress tests: The risk of regulatory risk weights. *Journal of Monetary Economics*, *65*, 36-53.

Acharya, V., Pedersen, L. H., Philippon, T., & Richardson, M. (2017, January). Measuring Systemic Risk. *Review of Financial Studies*, *30*(1), 2-47.

Adrian, T., & Brunnermeier, M. K. (2017, July). CoVaR. *American Economic Review*, *106*(7), 1705-1741.

Aikman, D., Alessandri, P., Eklund, B., Gai, P., Kapadia, S., Martin, E., … Willison, M. (2009). Funding liquidity risk in a quantitative model of systemic stability. *Bank of England Working Paper No. 372*.

Bank of England. (2019). Stress testing the UK banking system: 2019 guidance for participating banks and building societies.

BCBS. (2008). Principles for Sound Liquidity Risk Management and Supervision. *Bank for International Settlements*.

BCBS. (2009). Principles for sound stress testing practices and supervision. *Bank for International Settlements*.

BCBS. (2013). Liquidity stress testing: a survey of theory, empirics and current industry and supervisory practices. *Bank for International Settlements Working Paper No. 24*.

BCBS. (2017). Supervisory and bank stress testing: range of practices. *Bank for International Settlements*.

BCBS. (2018). Stress testing principles. *Bank for International Settlements*.

BCBS. (2019). The Basel Framework. *Bank for International Settlements*.

Blaschke, W., Jones, M. T., Majnoni, G., & Martinez Peria, S. (2001). Stress Testing of Financial Systems: An Overview of Issues, Methodologies, and FSAP Experiences. *IMF Working Paper WP/01/88*.

Borio, C. (2011). Rediscovering the macroeconomic roots of financial stability policy: journey, challenges and a way forward. *BIS Working Papers No 354*.

Borio, C., Drehmann, M., & Tsatsaronis, K. (2014, June). Stress-testing macro stress testing: Does it live up to expectations? *Journal of Financial Stability*, *12*, 3-15.

Boss, M. (2002). Ein makroökonomisches Kreditrisikomodell zur Durchführung von Krisentests für das österreichische Kreditportfolio. *Finanzmarktstabilitätsbericht 04/2002*(12), 68-88.

Breckenfelder, J., & Schwaab, B. (2018, December). Bank to sovereign risk spillovers across borders: Evidence from the ECB's Comprehensive Assessment. *Journal of Empirical Finance*, *49*, 247-262.

Breuer, T., Jandačka, M., Mencia, J., & Summer, M. (2012, February). A systematic approach to multi-period stress testing of portfolio credit risk. *Journal of Banking & Finance*, *36*(2), 332-340.

Breuer, T., Jandačka, M., Rheinberger, K., & Summer, M. (2008). Hedge the stress. using stress tests to design hedges for foreign currency loans. In D. Rösch & H. Scheule (Eds.), *Stress-testing for Financial Institutions - applications, regulations and techniques* (p. 111-126). Risk Books.

Breuer, T., Jandačka, M., Rheinberger, K., & Summer, M. (2010, April). Does adding up of economic capital for market- and credit risk amount to conservative risk assessment? *Journal of Banking & Finance*, *34*(4), 703-712.

Brownlees, C. T., & Engle, R. F. (2017, January). SRISK: A Conditional Capital Shortfall Measure of Systemic Risk. *Review of Financial Studies*, *30*(1), 48-79.

Brunnermeier, M. K., & Pedersen, L. H. (2009, June). Market Liquidity and Funding Liquidity. *The Review of Financial Studies*, *22*(6), 2201-2238.

Buncic, D., & Melecky, M. (2013, September). Macroprudential stress testing of credit risk: A practical approach for policy makers. *Journal of Financial Stability*, *9*(3), 347-370.

Burrows, O., Learmonth, D., & McKeown, J. (2012). RAMSI: a top-down stress-testing model. *Bank of England Financial Stability Paper No. 17*.

Busch, R., Koziol, P., & Mitrovic, M. (2018, May). Many a little makes a mickle: Stress testing small and medium-sized German banks. *The Quarterly Review of Economics and Finance*, *68*, 237-253.

CEBS. (2010a). Aggregate outcome of the 2010 EU wide stress test exercise coordinated by CEBS in cooperation with the ECB.

CEBS. (2010b). Guidelines on Stress Testing (GL32).

Cerrone, R., Cocozza, R., Curcio, D., & Gianfrancesco, I. (2017, June). Does prudential regulation contribute to effective measurement and management of interest rate risk? Evidence from Italian banks. *Journal of Financial Stability*, *30*, 126-138.

Cerutti, E., & Schmieder, C. (2014, April). Ring fencing and consolidated banks' stress tests. *Journal of Financial Stability*, *11*, 1-12.

Cetina, J., Paddrik, M., & Rajan, S. (2018, April). Stressed to the core: Counterparty concentrations and systemic losses in CDS markets. *Journal of Financial Stability*, *35*, 38-52.

CGFS. (2000). Stress Testing by Large Financial Institutions: Current Practice and Aggregation Issues. *Bank for International Settlements*.

Čihák, M. (2007). Introduction to Applied Stress Testing. *IMF Working Paper WP/07/59*.

Cornett, M. M., Minnick, K., Schorno, P. J., & Tehranian, H. (2020, January). An examination of bank behavior around Federal Reserve stress tests. *Journal of Financial Intermediation*, *41*, Article 100789.

Cortés, K. R., Demyanyk, Y., Li, L., Loutskina, E., & Strahan, P. E. (2020, April). Stress tests and small business lending. *Journal of Financial Economics*, *136*(1), 260-279.

Covas, F. (2018). What a Recent Fed Paper Gets Wrong on European Stress Tests. *Bank Policy Institute*.

de Haan, L., & van den End, J. W. (2013, October). Banks' responses to funding liquidity shocks: Lending adjustment, liquidity hoarding and fire sales. *Journal of International Financial Markets, Institutions & Money*, *26*, 152-174.

Demirgüç-Kunt, A., & Detragiache, E. (2011, December). Basel Core Principles and bank soundness: Does compliance matter? *Journal of Financial Stability*, *7*(4), 179-190.

Demirgüç-Kunt, A., Detragiache, E., & Tressel, T. (2008, October). Banking on the principles: Compliance with Basel Core Principles and bank soundness. *Journal of Financial Intermediation*, *17,*(4), 511-542.

Dissem, S., & Lobez, F. (2020). Correlation between the 2014 EU-wide stress tests and the market-based measures of systemic risk. *Research in International Business and Finance*, *51*(C), Article 100939.

Doerr, S. (2019). Unintended Side Effects: Stress Tests, Entrepreneurship, and Innovation. *BIS Working Papers No 823*.

Dombret, A. R., Foos, D., Pliszka, K., & Schulz, A. (2019). What are the real effects of financial market liquidity? Evidence on bank lending from the euro area. *Journal of International Financial Markets, Institutions & Money*, *62*, 152-183.

Drehmann, M., Patton, A., & Sorensen, S. (2007). Non-linearities and stress testing. In *Risk measurement and systemic risk, Proceedings of the Fourth Joint Central Bank Research Conference*. ECB.

Drehmann, M., Sorensen, S., & Stringa, M. (2010, April). The integrated impact of credit and interest rate risk on banks: A dynamic framework and stress testing application. *Journal of Banking & Finance*, *34*(4), 713-729.

Düllmann, K., & Kick, T. (2014). Stress testing German banks against a global credit crunch. *Financial Markets and Portfolio Management*, *28*, 337-361.

EBA. (2011). 2011 EU-wide stress test aggregate report.

EBA. (2013). Report on the peer review of the EBA Stress Testing Guidelines (GL 32).

EBA. (2014). Results of 2014 EU-wide stress test.

EBA. (2018). Guidelines on common procedures and methodologies for the supervisory review and evaluation process (SREP) and supervisory stress testing.

EBA. (2020). Discussion paper on the future changes to the EU-wide stress test. *EBA Discussion Paper EBA/DP/2020/01*.

ECB. (2008, November). EU banks' liquidity stress testing and contingency funding plans.

ECB. (2017). Sensitivity Analysis of IRRBB – Stress test 2017.

ECB. (2018). ECB Guide to the internal capital adequacy assessment process (ICAAP).

ECB. (2019). Sensitivity Analysis of Liquidity Risk – Stress Test 2019.

Enria, A. (2019). The future of stress testing - realism, relevance and resources. *Speech delivered at the European Systemic Risk Board (ESRB) Annual Conference*. (Frankfurt am Main, 26 September 2019)

ESRB. (2018). Adverse macro-financial scenario for the 2018 EU-wide banking sector stress test.

European Commission. (2015). Implementing regulation (EU) 2015/2344. *Official Journal of the European Union*.

European Parliament. (2019). Towards a fundamental re-design of Banks' Stress Tests in the EU?

Federal Reserve. (2009a). The Supervisory Capital Assessment Program: Design and Implementation.

Federal Reserve. (2009b). The Supervisory Capital Assessment Program: Overview of Results.

Federal Reserve. (2019). Comprehensive Capital Analysis and Review 2019: Assessment Framework and Results.

Flood, M. D., & Korenko, G. G. (2015). Systematic scenario selection: stress testing and the nature of uncertainty. *Quantitative Finance*, *15*(1), 43-59.

Foos, D., Lütkebohmert, E., Pliszka, K., & Markovych, M. (2017). Euro area banks' interest rate risk exposure to level, slope and curvature swings in the yield curve. *Deutsche Bundesbank Working Paper No 24/2017*.

FSA. (2009). Stress and Scenario Testing, Feedback on CP08/24 and Final Rules. Policy Statement 09/20.

FSI. (2018a). FSI Insights on policy implementation No 12: Stress-testing banks – a comparative analysis. *Bank for International Settlements*.

FSI. (2018b). FSI Insights on policy implementation No 12: Stress-testing banks – a comparative analysis. Online Appendix. *Bank for International Settlements*.

Gaglianone, W. P., & Schechtman, R. (2012, September). Macro stress testing of credit risk focused on the tails. *Journal of Financial Stability*, *8*(3), 174-192.

Georgescu, O.-M., Gross, M., Kapp, D., & Kok, C. (2017). Do stress tests matter? evidence from the 2014 and 2016 stress tests. *ECB Working Paper Series No 2054 / May 2017*.

Gick, W., & Pausch, T. (2012). Persuasion by stress testing - optimal disclosure of supervisory information in the banking sector. *Deutsche Bundesbank Discussion Paper 32/2012*.

Glasserman, P., Kang, C., & Kang, W. (2015). Stress scenario selection by empirical likelihood. *Quantitative Finance*, *15*(1), 25-41.

Goldstein, I., & Leitner, Y. (2018, September). Stress tests and information disclosure. *Journal of Economic Theory*, *177*, 34-69.

Goldstein, I., & Sapra, H. (2014). Should Banks' Stress Test Results be Disclosed? An Analysis of the Costs and Benefits. *Foundations and Trends in Finance*, *8*(1), 1-54.

Grundke, P. (2011, March). Reverse stress tests with bottom-up approaches. *The Journal of Risk Model Validation*, *5*(1), 71-90.

Grundke, P. (2012, June). Further recipes for quantitative reverse stress testing. *The Journal of Risk Model Validation*, *6*(2), 81-102.

Grundke, P., & Kühn, A. (2020, February). The impact of the Basel III liquidity ratios on banks: Evidence from a simulation study. *The Quarterly Review of Economics and Finance*, *75*(C), 167-190.

Grundke, P., & Pliszka, K. (2018). A macroeconomic reverse stress test. *Review of Quantitative Finance and Accounting*, *50*(4), 1093-1130.

Grundke, P., Pliszka, K., & Tuchscherer, M. (2020, July). Model and estimation risk in credit risk stress tests. *Review of Quantitative Finance and Accounting*, *55*(1), 163-199.

IMF. (2018). Euro Area Policies: Financial System Stability Assessment. *IMF Staff Country Report No. 18/226*.

IMF. (2019). Global Financial Stability Report: Lower for Longer.

Jokivuolle, E., & Virén, M. (2013, April). Cyclical default and recovery in stress testing loan losses. *Journal of Financial Stability*, *9*(1), 139-149.

Jokivuolle, E., Virolainen, K., & Vähämaa, O. (2008). Marco-model-based stress testing of Basel II capital requirements. *Bank of Finland Research Discussion Papers 17/2008*.

Kanas, A., & Molyneux, P. (2018, May). Macro stress testing the U.S. banking system. *Journal of International Financial Markets, Institutions & Money*, *54*, 204-227.

Klein, A. C., & Pliszka, K. (2018). The time-varying impact of systematic risk factors oncorporate bond spreads. *Deutsche Bundesbank Discussion Paper 14/2018*.

Kopeliovich, Y., Novosyolov, A., Satchkov, D., & Schachter, B. (2015). Robust Risk Estimation and Hedging: A Reverse Stress Testing Approach. *The Journal of Derivatives*, *22*(4), 10-25.

Koziol, P., Schell, C., & Eckhardt, M. (2015). Credit risk stress testing and copulas - is the Gaussian copula better than its reputation? *Deutsche Bundesbank Discussion Paper No 46/2015*.

Kupiec, P. H. (2018, October). On the accuracy of alternative approaches for calibrating bank stress test models. *Journal of Financial Stability*, *38*, 132-146.

Lazzari, V., Vena, L., & Venegoni, A. (2017, October). Stress tests and asset quality reviews of banks: A policy announcement tool. *Journal of Financial Stability*, *32*, 86-98.

Leitner, Y., & Yilmaz, B. (2019, February). Regulating a model. *Journal of Financial Economics*, *131*(2), 251-268.

Lin, L., & Surti, J. (2015, March). Capital requirements for over-the-counter derivatives central counterparties. *Journal of Banking & Finance*, *52*, 140-155.

McNeil, A. J., & Smith, A. D. (2012, May). Multivariate stress scenarios and solvency. *Insurance: Mathematics and Economics*, *50*(3), 299-308.

Moretti, M., Stolz, S., & Swinburne, M. (2008). Stress Testing at the IMF. *IMF Working Paper WP/08/206*.

Niepmann, F., & Stebunovs, V. (2018). Modeling Your Stress Away. *FRB International Finance Discussion Papers No. 1232*.

Packham, N., & Woebbeking, C. F. (2019, April). A factor-model approach for correlation scenarios and correlation stress testing. *Journal of Banking & Finance*, *101*, 92-103.

Petrella, G., & Resti, A. (2013, December). Supervisors as information producers: Do stress tests reduce bank opaqueness? *Journal of Banking & Finance*, *37*(12), 5406-5420.

Quagliariello, M. (2019). Are stress tests beauty contests? (and what we can do about it). *EBA Staff Paper Series Nr. 4*.

Quarles, R. (2019). Refining the Stress Capital Buffer. *Speech delivered at the Program on International Financial Systems Conference*. (Frankfurt am Main, 5 September 2019)

Schmieder, C., Hesse, H., Neudorfer, B., Puhr, C., & Schmitz, S. W. (2012). Next Generation System-Wide Liquidity Stress Testing. *IMF Working Paper WP/12/3*.

Skoglund, J., & Chen, W. (2009). Risk contributions, information and reverse stress testing. *The Journal of Risk Model Validation*, *3*(2), 61-77.

Sorge, M., & Virolainen, K. (2006, June). A comparative analysis of macro stress-testing methodologies with application to Finland. *Journal of Financial Stability*, *2*(2), 113-151.

Soula, J.-L. (2017, November). Measuring heterogeneity in bank liquidity risk: Who are the winners and losers? *The Quarterly Review of Economics and Finance*, *66*, 302-313.

Ulloa, B. (2020). Stress testing in Latin America: A comparison of approaches and methodologies. *BIS Papers No 108*.

van den End, J. W. (2012). Liquidity stress-tester: do Basel III and unconventional monetary policy work? *Applied Financial Economics*, *22*(15), 1233-1257.

Vilsmeier, J., & Siemsen, T. (2018). On a quest for robustness: about model risk, randomness and discretion in credit risk stress tests. *Deutsche Bundesbank Discussion Paper 31/2018*.

Virolainen, K. (2004). Macro stress testing with a macroeconomic credit risk model for Finland. *Bank of Finland Discussion Papers No. 18/2004*.

Wilson, T. C. (1997a). Portfolio credit risk: part I. *Risk*, *10*, 111-117.

Wilson, T. C. (1997b). Portfolio credit risk: part II. *Risk*, *10*, 56-61.

Wong, E., & Hui, C.-H. (2009). A Liquidity Risk Stress-Testing Framework with Interaction between Market and Credit Risks. *Hong Kong Monetary Authority Working Paper 06/2009*.

Wong, J., Choi, K.-F., & Fong, T. (2008). A framework for stress-testing banks' credit risk. *The Journal of Risk Model Validation*, *2*(1), 3-23.