A Work Project, presented as part of the requirements for the Award of a Master's degree in Finance from the Nova School of Business and Economics

Stress Test Methodologies Under Scrutiny: Examining the Interplay of Regulatory and Market Risk Metrics in Europe – A Comparative Analysis of 2011 and 2021

Geronimo Duckert – 58336

Work project carried out under the supervision of:

Tim Eisert

23/01/2025

Abstract

This thesis replicates and extends the analysis from "Testing Macroprudential Stress Tests: The Risk of Regulatory Risk Weights" by Acharya, Engle, and Pierret (2014). It investigates the relationship between capital shortfall metrics and the effectiveness of risk factors in predicting stress-induced risks, using EBA stress tests from 2011 and 2021 alongside market-based metrics like SRISK. The findings validate critiques of risk-weighted measures, highlight the alignment of leverage-based and market-based metrics, and advocate for combining market-based and regulatory approaches. Additionally, the analysis suggests the successful implementation of regulatory standards, such as Basel III, and their impact on regulatory stress tests.

Keywords

Systemic risk, Stress test, Regulation, Capital shortfall, Replication

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

Table of Contents

List of Figures

List of Tables

List of Abbreviations

BuffSov	Buffer for Sovereign Debt Exposures
CET1	Common Equity Tier 1
EBA	European Banking Authority
ECB	European Central Bank
GFC	Global Financial Crisis
ICAAP	Internal Capital Adequacy Assessment Process
IRB	Internal ratings-based
LRMES	Long-Run Marginal Expected Shortfall
PEPP	Pandemic Emergency Purchase Programme
RWA	Risk-weighted Assets
SCAP	Supervisory Capital Assessment Program
SSM	European Single Supervisory Mechanism
T1LVGR	Tier 1 leverage ratio
TREA	Total Risk Exposure Amount
VaR	Value-at-Risk

1 Introduction

1.1 Motivation

The Global Financial Crisis of 2008 (GFC) exposed significant weaknesses in the international financial system and revealed the vulnerabilities of the European Monetary Union. Banks operated with low capital ratios, and banking supervision was under the responsibility of national authorities. This resulted in systemic risks not being effectively mitigated by national regulations (Rutkowski, Schäfer, and Schnabel 2018, 10). The crisis depleted the capital reserves of several banks in the Eurozone and generated a massive volume of non-performing loans (Kalhoefer and Lang 2019, 192). The distress of the banking sector, high levels of credit growth, and bailout packages for certain countries were additional consequences of the financial crisis, which gradually evolved into a sovereign debt crisis in subsequent years (Lane 2012, 55). Beckert (2010, 9) further emphasizes that the GFC also represents a crisis of trust, which could have led to widespread liquidity problems or the collapse of interbank trading without government intervention.

In response to the GFC, stricter requirements were introduced for banks, and regulatory frameworks were reformed. Among these measures was Basel III, an internationally agreed-upon framework first published in December 2010. It established capital requirements, aimed at improving risk management, enhancing transparency in bank disclosures, and addressing the issues highlighted by the financial crisis (Basel Committee on Banking Supervision 2010, 1).

Furthermore, it became evident that ensuring financial system stability requires not only the supervision of individual banks through a microprudential perspective but also a holistic approach that incorporates a macroprudential view of the financial system (Hirtle, Schuermann, and Stiroh 2009, 1). Consequently, stress tests gained increasing importance for regulatory authorities in the wake of the financial crisis and became established as macroprudential standard tools aimed at supporting regulations and testing the resilience of the entire financial system (Acharya, Engle, and Pierret 2014, 36). Despite their predominantly macroprudential perspective, regulatory stress tests also evaluate the ability of individual banks to withstand adverse economic developments and shocks (Ferrari, Roy, and Vespro 2011, 105). Additionally, the results of stress tests are published by supervisory authorities to strengthen market confidence in the long term or restore it during crises (Acharya, Engle, and Pierret 2014, 37).

In this context, it is undeniable that stress tests play a crucial role in maintaining financial market stability. Crisan (2014, 255) even argues that their application forms the foundation for a healthy and robust banking system, particularly in the event of another crisis. However, there is a call for further research and the derivation of best practices that consider aspects such as the impact of stress events on other parts of the financial system or behavioral changes in banks under stress, as these interconnections are often underrepresented in conventional stress tests (Bookstaber et al. 2013, 11). Moreover, concerns persist about whether they comprehensively capture the issues and systemic risks in the financial sector (Steffen 2014, 29).

The work of Acharya, Engle, and Pierret (2014) provides an important starting point in this context, as it addresses a major weakness of many regulatory stress tests: their reliance on static risk assessments that fail to reflect the dynamic nature of risks. Their study emphasizes the need for methodologies that better capture these dynamics and introduces the V-Lab stress test as a benchmark to evaluate banks' capital shortfalls and vulnerabilities to adverse scenarios. Furthermore, their findings suggest that SRISK, the alternative market-based measure from V-Lab, offers greater predictive power in identifying systemic risks and forecasting potential imbalances in the banking sector. However, the study is not without limitations. Its analysis relies on 2011 data, which does not reflect the significant changes in the economic and regulatory environment since the GFC. Additionally, the study focuses solely on a specific stress test in Europe, potentially limiting its generalizability. Furthermore, it also only partially accounts for interdependence between banks and their behavior under stress. These limitations underscore the importance of replicating and expanding the study to validate the robustness of its findings and derive new insights in the context of current stress testing practices.

1.2 Research questions

The purpose of this thesis is therefore to respond to the call for further research by replicating and subsequently extending the analysis of the study “Testing macroprudential stress tests: The risk of regulatory risk weights” by Acharya, Engle, and Pierret (2014), based on European stress tests conducted by the European Banking Authority (EBA). Like the original study, our work focuses on two main analyses: first, a detailed examination of the relationship and comparison between SRISK and regulatory capital shortfall metrics; and second, an empirical analysis of regulatory and market-based risk measures using regression models to shed light on the relative importance of these factors in predicting risks during stress scenarios.

Replicating the original study’s results is a crucial first step to verify the robustness and validity of the findings and their relevance for financial system stability and systemic risk measurement with updated data. Particularly in finance, where reliable risk assessments are essential, replication builds trust in the applied methodology, ensuring robust models and assumptions, and confirming that the conclusions remain applicable. This represents a key prerequisite for extending the stress test analysis at a later stage and also motivates the first research question:

RQ 1: Do the original conclusions of Acharya, Engle, and Pierret’s (2014) analysis remain valid when replicated using 2011 EBA stress test data and updated V-Lab benchmark data?

Over time, and since the publication of the 2011 stress test by the EBA, which served as the basis for the study by Acharya, Engle, and Pierret (2014, 39), several changes have occurred in the financial system, particularly for banks. Economic conditions have shifted due to major events such as the GFC, the subsequent low-interest-rate phase, and the COVID-19 pandemic, while regulatory frameworks have also been reformed, imposing new requirements on participating banks. It is therefore crucial to examine whether the original results remain valid or whether new developments alter the findings and outcomes of stress tests when applied to more recent data. To address this, stress test data 2021 will be analyzed in detail, offering an updated perspective on how changes in the financial and regulatory landscape have influenced their outcomes. The second research question thus aims to confirm the insights based on the 2011 stress test or identify potential deviations and new conclusions:

RQ 2: Do the replicated findings and the original conclusions of Acharya, Engle, and Pierret's (2014) study remain valid when extended using 2021 EBA stress test data?

1.3 Research method

This thesis employs a research method that combines replication and extension of prior analyses, focusing on the methodologies developed by Acharya, Engle, and Pierret (2014) and applying them to recent data. Replication, a cornerstone of scientific research, verifies the robustness and credibility of findings. It serves as a critical step in strengthening confidence in the general applicability of the conclusions or in identifying constraints that require refinement (Nosek and Errington 2020, 7), while also providing the foundation for subsequent analyses.

Beyond replication, this thesis extends the analysis by incorporating stress test data from 2021, enabling an evaluation of the stability and consistency of the original approach across different time periods and evolving economic conditions. This extension also identifies emerging patterns and enhances the understanding of the broader applicability of the methods and conclusions. As King (1995, 451) highlights, extending replicated studies advances scientific knowledge by challenging, confirming, or refining existing insights while contextualizing them for new conditions.

By combining replication and extension, this thesis rigorously evaluates the original study's framework while offering new perspectives on stress testing dynamics and systemic risk measurement. This dual approach ensures that conclusions remain relevant within the current financial and regulatory environment and enriches the broader discourse on financial stability.

The remainder of this paper is structured as follows: Section 2 provides a comprehensive literature review on stress tests, focusing on the EBA and the V-Lab approach. Section 3 outlines the methodology underpinning the replication and extension analyses. Section 4 presents and discusses the results of the replication study, while Section 5 examines the extended analysis, situating new findings within the replication and original study. Finally, Section 6 summarizes key insights, answers the research questions, and offers an outlook for future research.

2 Literature review

This section provides a comprehensive overview of the academic literature on the development of stress tests, their significance for banks and supervisory authorities, and the various application perspectives. Subsection 2.1 outlines the history of stress tests, while Subsection 2.2 examines their importance, particularly in a regulatory context. Subsection 2.3 discusses microprudential and macroprudential perspectives of stress testing. Finally, Subsection 2.4 compares the EBA and V-Lab stress tests, providing the context for the methodology of Acharya, Engle, and Pierret (2014) as well as for the replication and extension of their study.

2.1 Stress test history

1980s and 1990s: The beginnings of scenario analysis and Value-at-Risk

Stress testing emerged in the 1980s as an internal tool within large financial institutions to assess potential losses under adverse economic conditions (Anderson et al. 2018, 8). Initially limited to basic scenario analysis focusing on hypothetical situations, it lacked formal standards or regulatory oversight (Dent, Westwood, and Segoviano 2016, 131). In the 1990s, the adoption of Value-at-Risk (VaR) models became widespread, enabling banks to quantify risks over specific time periods for internal reporting (Taskinsoy 2022, 5). Despite its limitations, VaR provided a measurable framework for risk assessment and laid the foundation for stress testing. Concurrently, Basel I was introduced in 1988, establishing minimum capital requirements for banks, though without directly mandating stress testing (Taskinsoy 2019, 13).

Early 2000s: The introduction of Basel II and operational risk inclusion

The Basel II reform, introduced in 2004, incorporated operational risk into regulatory frameworks and encouraged banks to enhance their internal risk models, laying the foundation for more formalized stress testing practices. Its Pillar 2 guidelines required banks to conduct stress tests that addressed risks beyond market risk, as captured by VaR, although the specifics were largely left to the discretion of individual institutions (Taskinsoy 2022, 6). However, Basel II faced criticism for its reliance on external credit ratings, which incentivized riskier behavior and undermined the development of internal risk models by over-relying on rating agencies (Lall 2009, 14–17). This reliance, combined with Basel II's procyclical tendencies, exacerbated defaults and financial instability, contributing to the 2008 GFC (Taskinsoy 2022, 18–19).

2008: Global Financial Crisis – A turning point in regulatory stress testing

The 2008 GFC marked a major shift in stress testing practices, revealing deep systemic vulnerabilities and demonstrating the inadequacy of traditional risk models in addressing interconnected risks within the financial sector (Rutkowski, Schäfer, and Schnabel 2018, 10). In 2009, the U.S. Federal Reserve implemented the Supervisory Capital Assessment Program (SCAP), the first large-scale regulatory stress test, to evaluate and publicly disclose the capital adequacy of major banks under severe economic scenarios (Hirtle, Schuermann, and Stiroh 2009, 3). Following this, the EBA conducted its first EU-wide stress test in 2010, emphasizing transparency and systemic stability in the European banking sector, and underscoring the critical role of such stress tests as regulatory tools (Acharya, Engle, and Pierret 2014, 37).

2010s: The expansion and international standardization of stress testing

In the aftermath of the GFC, Basel III was introduced to address weaknesses in existing regulatory frameworks, bolstering banks against shocks and enhancing their resilience. It set stricter capital and liquidity requirements, defining Common Equity Tier 1 (CET1), additional Tier 1, and Tier 2 capital while raising the minimum CET1 ratio to 4.5% by 2015 (Basel

Committee on Banking Supervision 2010, 12). A capital conservation buffer, starting at 0.625% in 2016 and reaching 2.5% by 2019, helped banks to absorb losses during downturns, while a countercyclical buffer further mitigated procyclicality (Taskinsoy 2013, 4). Overall, the Basel reforms marked a shift towards standardizing stress testing and improving systemic stability.

2020s: The incorporation of climate risks in financial regulation

In the 2020s, stress testing evolved to address emerging challenges such as climate change and operational vulnerabilities, including cybersecurity threats. Regulatory bodies, including the European Central Bank (ECB), the Bank of England, and the U.S. Federal Reserve, initiated climate stress tests to evaluate the impact of physical climate risks and transition risks associated with shifts towards green economies (Bolton et al. 2020, 7). These developments signal an expanded understanding of stress testing, encompassing not only traditional economic risks but also long-term systemic vulnerabilities (Covas 2020, 2).

The evolution of stress testing reflects the increasing complexity and interconnectedness of global finance. Originally an internal risk management tool, stress testing has transformed into a critical regulatory instrument, shaped by major financial crises and evolving economic conditions. Its development underscores the dynamic nature of financial systems and the ongoing need for adaptive regulatory frameworks to safeguard systemic stability.

2.2 Importance of stress tests

Stress tests are sophisticated risk assessment tools employed by financial institutions and regulators to evaluate the ability of banks or financial systems to withstand economic shocks (Vodenska et al. 2021, 1). By simulating hypothetical or historical adverse scenarios, they provide insights into how financial institutions might perform under extreme conditions, such as recessions, asset devaluations, or sudden interest rate hikes, enabling the early identification of vulnerabilities (Hirtle and Lehnert 2015, 342). Stress tests assess critical components of financial stability, including capital adequacy, expected losses under stress, and risk exposures, addressing both microprudential concerns at the institutional level and macroprudential risks across the broader financial system (Baudino et al. 2018, 7–9).

The importance of stress tests lies in their dual role as proactive and protective measures within the financial sector. By identifying weaknesses, they allow institutions and regulators to prepare for adverse conditions and build capital buffers or implement regulatory interventions (Hirtle and Lehnert 2015, 346). Additionally, stress tests serve as transparency tools; public disclosures by central regulatory authorities bolster market confidence by demonstrating that banks are equipped to manage potential crises. This transparency reassures investors and the public that financial institutions are taking proactive steps to prevent the failures witnessed during the 2008 GFC (Acharya, Engle, and Pierret 2014, 37; Schmieder, Puhr, and Hasan 2011, 4). Moreover, stress tests incentivize banks to strengthen their risk management practices by adopting conservative lending strategies and maintaining robust capital reserves. This contributes to the resilience of both individual institutions and the financial system as a whole (Borio, Drehmann, and Tsatsaronis 2014, 10).

While stress tests offer numerous benefits, they also have notable limitations. One major drawback is their reliance on predefined scenarios, which often cannot encompass the full range of potential economic shocks (ECB 2013, 15). When unforeseen events occur, such as the COVID-19 pandemic, stress tests may fail to predict the full extent of potential losses or vulnerabilities, as scenarios based on past crises may not account for novel factors (Hirtle, Schuermann, and Stiroh 2009, 14). Additionally, the complex financial models underpinning stress tests are constrained by assumptions and simplifications. These models, while useful for risk estimation, may overlook intricate interdependencies within the financial system, potentially leading to biased or overly optimistic outcomes (Baudino et al. 2018, 13).

Despite these limitations, stress tests remain a cornerstone of modern risk management. Their prominence in the post-2008 regulatory framework highlights the necessity for ongoing advancements in modeling techniques and scenario design to adapt to evolving economic landscapes. While stress tests are not a panacea, they are indispensable for identifying vulnerabilities, fostering prudent risk management, and safeguarding financial stability.

2.3 Microprudential and macroprudential perspectives of stress tests

Over the years, two main perspectives have dominated the field of stress testing: microprudential and macroprudential, which differ fundamentally in scope, objectives, and methodology, reflecting their complementary role in addressing financial risks at both institutional and systemic levels.

Microprudential stress testing

Microprudential stress testing focuses on individual financial institutions, with the primary objective of ensuring their solvency and resilience, particularly during periods of economic stress. This approach evaluates institution-specific risks, including credit risk, market risk, liquidity risk, and operational risk (Taskinsoy 2019, 4–7). By concentrating on the internal vulnerabilities of banks, microprudential stress tests assess whether institutions maintain adequate capital and liquidity buffers to withstand adverse economic scenarios.

The methodology employed in microprudential stress testing is typically bottom-up, relying on granular data unique to each institution's risk profile (Greenlaw et al. 2012, 5). While this allows for tailored assessments of vulnerabilities at the individual bank level, this approach also presents challenges. The lack of standardized modeling techniques and assumptions across banks can hinder comparability and complicate systemic risk aggregation (Taskinsoy 2022, 31).

The origins of microprudential stress testing can be traced to regulatory initiatives that predate the 2008 GFC. The framework was further formalized in Basel II and subsequently strengthened in Basel III, where it became an integral component of the Internal Capital Adequacy Assessment Process (ICAAP) (Taskinsoy 2019, 6). Under ICAAP, banks are required to conduct stress tests as part of their capital planning processes to ensure they maintain sufficient buffers to address risks such as credit defaults and interest rate fluctuations. This regulatory emphasis underscores the role of microprudential stress tests in safeguarding individual institutions against a range of financial risks.

Despite their utility, microprudential stress tests are resource-intensive and inherently limited by their narrow scope. While they excel at identifying vulnerabilities within individual institutions, they often fail to account for systemic interdependencies, such as the potential for contagion effects arising from the failure of a single institution (Taskinsoy 2022, 5).

Macroprudential stress testing

In contrast to the bank-specific focus of microprudential tests, macroprudential stress tests assess the resilience of the entire financial system, aiming to identify and mitigate systemic risks such as interbank contagion, market-wide shocks, and macroeconomic instabilities. This approach gained prominence after the 2008 GFC, as regulators recognized the critical need to address vulnerabilities across interconnected financial markets (Hirtle and Lehnert 2015, 353). Macroprudential stress testing is typically conducted using top-down methodologies, where central banks or regulators apply standardized models and scenarios to multiple institutions simultaneously. This uniform approach facilitates the identification of systemic vulnerabilities and ensures consistency in assessing risks across the financial sector. By modeling interconnected risks, it allows regulators to anticipate how vulnerabilities in individual institutions might propagate through the system, thus preventing contagion during crises (Borio, Drehmann, and Tsatsaronis 2014, 3). Macroprudential tests are inherently forward-looking, incorporating macroeconomic models to anticipate emerging risks, which is crucial in a constantly evolving financial environment (Sorge 2004, 2). Unlike traditional analyses based solely on historical data, these stress tests focus on potential future challenges (Čihák 2007, 49). However, the standardization of macroprudential testing comes at the cost of granularity, as institution-specific risks may be overlooked (Taskinsoy 2022, 33).

While macroprudential stress testing gained prominence as a crisis-management tool, especially during the implementation of the U.S. Federal Reserve's SCAP in 2009, where it proved effective in restoring market confidence (Hirtle, Schuermann, and Stiroh 2009, 12), these tests also serve as valuable tools for policymakers, enabling them to design targeted interventions to strengthen financial stability.

Despite their effectiveness, the complexity of conducting system-wide tests makes them resource-intensive and costly. Additionally, reliance on assumptions about systemic linkages and economic dynamics can limit their accuracy, particularly when real-world conditions and interdependencies deviate from model projections.

2.4 EBA stress test vs. V-Lab stress test

This subsection compares the EBA stress testing framework with the V-Lab stress test methodology introduced by Acharya, Engle, and Pierret (2014). While both tests aim to assess financial stability under adverse conditions, their approaches differ markedly in methodology and assumptions. Therefore, this subsection explores their respective strengths, limitations, and implications for risk management.

2.4.1 The EBA stress testing framework

The EBA stress tests are designed to evaluate the capability of European banks to withstand adverse macroeconomic conditions. These tests are conducted within a standardized framework and under static balance sheet assumptions to ensure consistency and comparability across all participating institutions (EBA 2011a, 2; 2021a, 16). Thus, the EBA employs a bottom-up approach, where banks apply their internal risk models to simulate the effects of stress scenarios. These scenarios typically cover a horizon of two to three years, considering economic indicators such as GDP contraction, rising unemployment, and declining real estate values (EBA 2011a, 2). Results are reviewed by national supervisory authorities and subsequently by the EBA, ensuring adherence to the framework.

A key feature of the EBA framework is its reliance on risk-weighted assets (RWA) for measuring capital adequacy. RWAs, determined under Basel I and II, are calculated using banks' internal models tailored to their specific portfolios. However, despite the application of uniform scenarios, the calculations of risk parameters often lack full transparency, as they are subject to individual adjustments by banks (Acharya, Engle, and Pierret 2014, 42). This allows for institution-specific risk assessments, which introduces variability in results. For instance, in the 2011 stress test, 59 of 90 banks employed the internal ratings-based (IRB) approach (EBA 2011a, 24), enabling them to adjust risk weights and lower their capital requirements (Acharya, Engle, and Pierret 2014, 42). While the EBA framework provides detailed insights into individual bank vulnerabilities, it faces limitations. The static nature of the tests and their inability to capture systemic interlinkages reduce their effectiveness, especially during crises when risks propagate through interconnected markets. Moreover, the potential for manipulation of internal models undermines transparency and comparability, posing challenges for assessing systemic risks comprehensively.

2.4.2 The V-Lab market-based framework

In contrast, the V-Lab stress test, as benchmark methodology, adopts a dynamic, market-based approach to assessing systemic risk. The central metric is the estimated capital shortfall SRISK, which can be considered as the amount of capital a bank would need to raise during an economic crisis to restore a target capital ratio (Acharya, Engle, and Pierret 2014, 37). This metric is calculated exclusively using publicly available market data, making it independent of banks' internal models. SRISK is a function of the bank´s size, leverage, and its Long-Run Marginal Expected Shortfall (LRMES), which estimates the expected equity loss during a stress event (Brownlees and Engle 2017, 49). This event is defined as a 40% market decline over a six-month horizon, with the market decline and banks' equity losses simulated based on dynamic volatilities and correlations (Brownlees and Engle 2017, 12–13). The prudential capital ratio used in the SRISK calculation for the European context is set at 5.5% (Acharya, Engle, and Pierret 2014, 40). SRISK values are updated and publicly published by the NYU Stern Volatility and Risk Institute on its official website[1].

The V-Lab framework's strengths lie in its dynamic and forward-looking approach, enabling it to identify systemic vulnerabilities effectively. By avoiding reliance on RWAs or internal bank models, it eliminates opportunities for banks to manipulate outcomes. However, the methodology has its shortcomings. It lacks the granularity to analyze idiosyncratic risks at the institution level, which are often better captured by supervisory frameworks, and its accuracy relies on the availability of public market data, which may vary across regions and banks.

3 Methodology and data

This thesis adopts the methodology of Acharya, Engle, and Pierret (2014), applied consistently in both the replication and the extended analysis. The study employs a systematic approach to analyze the outcomes of macroprudential stress tests, focusing on systemic risks and capital requirements in the banking sector. The methodology is structured into three key steps: data preparation, analysis of capital shortfall metrics, and a regression analysis of risk factors.

Step 1: Data preparation

The analysis begins by constructing a sample of banks that are present in both the EBA datasets and the V-Lab database. Starting with the EBA stress test datasets from the relevant years, a matching process ensures consistency between the data sources. Banks are matched using identifiers such as names, tickers, and ISINs. Banks without publicly available market data are excluded, ensuring the final sample includes only institutions suitable for comparative analysis.

Step 2: Analysis of capital shortfall metrics

The primary objective of this step is to evaluate the relationship between market-based SRISK, and various regulatory capital shortfall metrics derived from EBA data. These metrics include

[1] Source: https://vlab.stern.nyu.edu

the disclosed capital shortfall, absolute capital shortfall, leverage-based capital shortfall, and overall shortfall. Each metric is calculated using specific information codes, date codes, and scenario codes provided by the EBA (see Appendix A). The data is harmonized across the stress test years to ensure comparability. Where data gaps or differences in metric definitions exist, adjustments and proxies are used to align the methodology with the original study. SRISK values, originally reported in US dollars, are converted to euros using ECB exchange rates for consistency.[2] Scatterplots and correlation analyses are then created to visualize and assess the strength and direction of relationships between SRISK and the regulatory shortfall metrics.

Step 3: Regression analysis of risk factors

The final step involves a regression analysis to examine the predictive power of various risk factors on realized market risk, measured by a six-month realized volatility, aligned with the EBA stress test publication dates. The analysis employs an Ordinary Least Squares (OLS) regression model across six specifications: (a) realized volatility and book-to-market ratio as baseline regression; (b) individual regressions of realized volatility on each of the risk factors: V-Lab risk weight, Tier 1 leverage ratio, and EBA risk weight; (c) a combined regression incorporating all variables except the EBA risk weight; (d) a full regression including all variables. Notably, all regressions control for the book-to-market ratio to account for the influence of accounting- versus market-based risk measurement. In some cases, banks lacking key data, such as book-to-market values, are excluded from the respective regression samples. By maintaining consistency with the original methodology, this study ensures the comparability of results in the replication phase. The extension to stress test data from 2021 evaluates the stability and evolution of the findings under different economic and regulatory conditions. This combined approach provides a comprehensive assessment of the original conclusions while exploring their applicability across multiple stress test periods.

---

[2] Source: https://www.ecb.europa.eu/stats/policy_and_exchange_rates/euro_reference_exchange_rates/html/eurofxref-graph-usd.en.html

4 Replication results

This section presents the results of our replication study in comparison with those of the original study, aiming to identify differences and assess the validity and robustness of the findings by Acharya, Engle, and Pierret (2014). The results and discussions are organized as follows: Subsection 4.1 introduces the bank sample, Subsection 4.2 explores the relationship between regulatory and market-based capital shortfall metrics, and Subsection 4.3 presents the regression analysis of risk factors.

4.1 Bank sample in the replication

The 2011 EBA stress test considered 90 banks as the initial sample (EBA 2011a, 2). Banks without publicly available market data were excluded, reducing the selection to 54 institutions. The verification of tickers and ISIN codes confirmed this selection, with two exceptions.

First, Nordea Bank, identified as Swedish by the EBA (EBA 2011c, 1), is legally a Finnish corporation, although it operates branches in Sweden (Nordea Bank Abp 2023, 47). To maintain data consistency and due to uncertainty about the legal entity referenced by the EBA, Nordea was excluded. Second, Irish Life & Permanent was renamed Permanent TSB Group Holdings plc in 2012 after its profitable insurance division was sold and the remaining part of the bank was nationalized (permanent tsb Group Holdings p.l.c 2024). Since the EBA data referred to the pre-restructuring entity, while the V-Lab data corresponded to the renamed entity, consistent mapping was not feasible, leading to its exclusion. This process resulted in 52 banks.

A subsequent review to identify any overlooked connections prompted one final adjustment. Effibank, commonly known by its trade name Liberbank (Moreno de Alborán de Cominges 2016, 11), was included, as no significant events were identified that could have impacted the bank's operations. This resulted in a final sample of 53 banks, forming the replication basis.

In addition, special consideration was given to the 2011 EBA Capital Exercise, which included a capital buffer for sovereign debt exposures (BuffSov) not accounted for in the 2011 EBA stress test. Therefore, data from the capital exercise, which covered 61 banks and used the same bank codes and names as the stress test (EBA 2012, 6), was applied instead. Mapping these banks to the sample yielded 40 banks, which were used to replicate the overall shortfall metric. Appendix B lists all banks considered in the analysis.

4.2 Capital shortfall metrics in the replication

Before delving into the replication findings of the initial comparison between SRISK and the disclosed capital shortfall, it is essential to outline the underlying formula:

$$\text{Disclosed capital shortfall} = \max(0,[k' * \text{RWA}_S - \text{Capital}_S]) \quad . \qquad (1)$$

Here, k' represents the prudential capital ratio of 5% for the 2011 stress test (EBA 2011b, 2). RWA_S refers to risk-weighted assets, and $Capital_S$ denotes the capital level of the bank, both measured at the end of the regulatory stress scenario.

The comparison highlights two key findings. First, the regulatory estimator is zero for most banks, indicating that the majority of banks would meet regulatory capital requirements under an adverse scenario (see Figure 1).

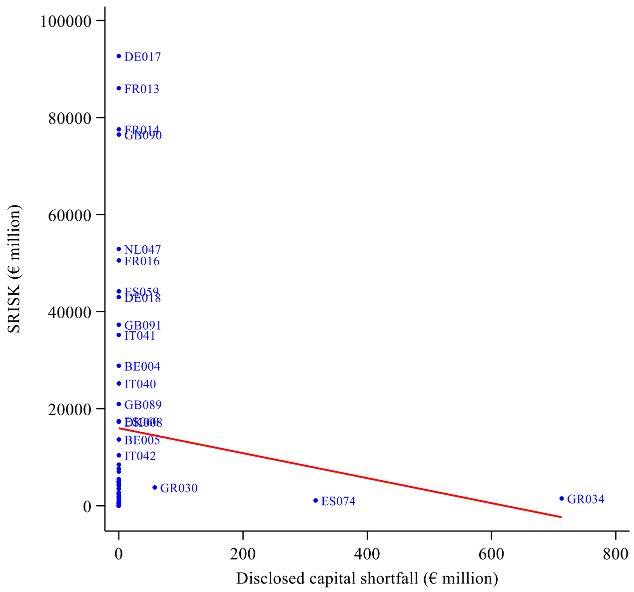


Figure 1. Disclosed capital shortfall based on 2011 EBA stress test data disclosed in July 2011 (Eq. (1)) vs. SRISK based on V-Lab data with reference date: 31.12.2010.
Source: Figure by G. Duckert

In contrast, SRISK values reveal that not all banks hold sufficient capital to withstand a systemic crisis. Second, the metrics exhibit a weak negative correlation, suggesting that banks with higher SRISK values often do not show a capital shortfall based on regulatory calculations. This finding highlights the differing approaches and objectives of the two metrics: SRISK, as a market-based risk measure, evaluates a bank's ability to absorb losses during systemic stress, whereas the disclosed shortfall is based solely on regulatory minimum capital requirements. When examining the absolute capital shortfall, a slightly modified metric is applied:

$$\text{Absolute capital shortfall} = k' * RWA_S - Capital_S \quad . \tag{2}$$

The results reveal a relatively strong negative correlation between SRISK and the regulatory shortfall (see Figure 2).

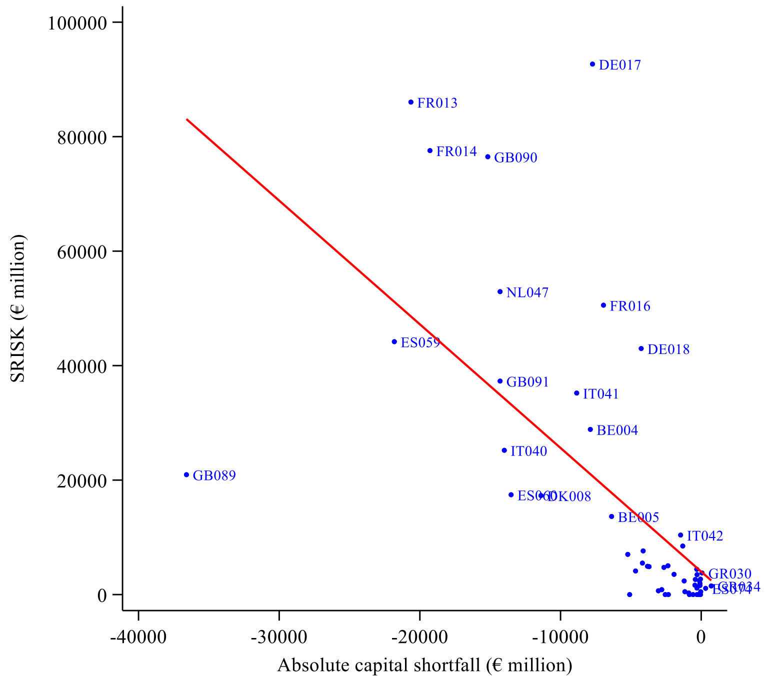


Figure 2. Absolute capital shortfall based on 2011 EBA stress test data disclosed in July 2011 (Eq. (2)) vs. SRISK based on V-Lab data with reference date: 31.12.2010.
Source: Figure by G. Duckert

This indicates that banks with higher SRISK tend to have greater regulatory capital surpluses. Most banks exhibit moderate to significant absolute capital surpluses, suggesting they exceed regulatory requirements and typically hold more capital than required under stress scenarios.

This corroborates the preceding analysis, underscoring that banks considered well-capitalized under the 2011 stress test appear to be among the least capitalized when assessed using SRISK. This pattern may stem from several factors. For instance, the EBA allowed banks to raise capital or implement restructuring plans during the first four months of 2011, which were considered in the stress exercise (EBA 2011a, 8). Additionally, the application of Basel II, particularly the IRB approach, contributed to a reduction in RWAs compared to Basel I, facilitating compliance with capital requirements (Basel Committee on Banking Supervision 2004, 13). The IRB approach enabled banks to apply their own methods for estimation, potentially resulting in overly optimistic assessments or limited transparency in the derivation of these RWAs (Le Leslé and Avramova 2012, 7–8). These factors underscore the potential misalignment between regulatory capital metrics and market-based risk measures like SRISK, which can identify additional risk factors, especially systemic risk, overlooked in regulatory frameworks. Acharya, Engle, and Pierret (2014, 50) further attributed differences to the milder stress scenario used by the EBA compared to V-Lab, possibly influenced by political considerations.

Both results align with the findings of Acharya, Engle, and Pierret (2014), who observed similar distributions, including many banks without regulatory shortfalls, and negative correlations between the metrics (see Appendix C for disclosed and Appendix D for absolute capital shortfall). Their study also reported statistically significant rank correlations: -0.273 for SRISK and the disclosed shortfall at the 5% level and -0.790 for SRISK and the absolute shortfall at the 1% level (Acharya, Engle, and Pierret 2014, 47). Our replication confirms these findings in terms of magnitude, with rank correlations of -0.136 for the disclosed shortfall and -0.747 for the absolute shortfall, the latter being statistically significant at the 0.1% level (see Table 1).

Table 1: Rank correlations between capital shortfall metrics based on 2011 EBA data disclosed in July and December 2011 vs. SRISK based on V-Lab data with reference date: 31.12.2010, except 30.09.2011 for the EBA overall shortfall.

Disclosed capital shortfall	Absolute capital shortfall	Leverage-based capital shortfall	EBA overall shortfall
-0.1356	-0.7467***	0.6922***	0.2382

* p<0.05, ** p<0.01, *** p<0.001

Source: Table by G. Duckert

The replication also examines the relationship between the leverage-based capital shortfall and SRISK, where regulatory capital adequacy is measured based on total assets rather than RWAs:

$$\text{Leverage-based capital shortfall} = \text{k} * \text{TA}_\text{S} - \text{Capital}_\text{S} \quad . \tag{3}$$

Under this metric, *k* corresponds to the prudential capital ratio used by V-Lab, set at 5.5% (Acharya, Engle, and Pierret 2014, 49). The results show that banks with high SRISK values also exhibit high leverage-based shortfalls (see Figure 3). The metrics are strongly positively correlated, reflecting their shared underlying concepts. SRISK incorporates factors such as firm size and leverage (Brownlees and Engle 2017, 49), while the leverage-based metric captures the capital gap relative to a minimum leverage ratio (Acharya, Engle, and Pierret 2014, 49).

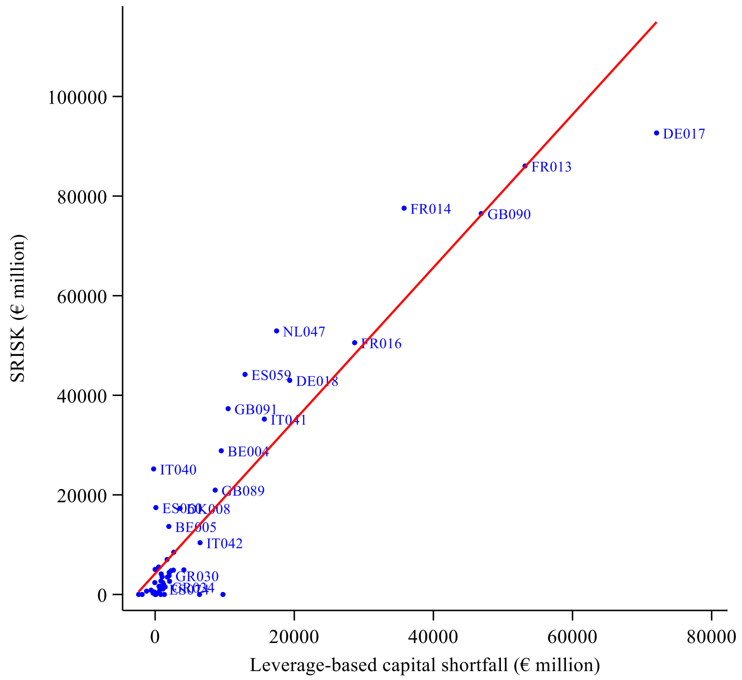


Figure 3. Leverage-based capital shortfall based on 2011 EBA stress test data disclosed in July 2011 (Eq. (3)) vs. SRISK based on V-Lab data with reference date: 31.12.2010.
Source: Figure by G. Duckert

Large, systemically important banks, characterized by high asset volumes or leverage, therefore tend to rank highly on both measures. This leverage-based metric also addresses concerns about the IRB approach for determining risk weights, as it does not rely on RWAs (Schäfer 2011, 7). Our findings align with those of Acharya, Engle, and Pierret (2014, 49) (see Appendix D), who observed a strong positive rank correlation of 0.679, statistically significant at the 1% level. We find an even stronger correlation of 0.692, highly significant at the 0.1% level (see Table 1). These results emphasize the influence of institution size and the relevance of the leverage-based approach for systemic risk analysis, given the remarkably close connection between the metrics. The EBA's subsequent capital exercise necessitates an adjusted metric, adding buffers for stressed sovereign exposures and raising the recommended capital ratio to 9% (EBA 2012, 5):

$$\text{EBA overall shortfall} = \max(0,[0.09 * \text{RWA} - \text{T1C}]) + \text{BuffSov} \quad . \tag{4}$$

Despite this adjustment, many banks still exhibited a capital shortfall of zero (see Figure 4). Among these were banks with some of the largest capital shortfalls according to SRISK.

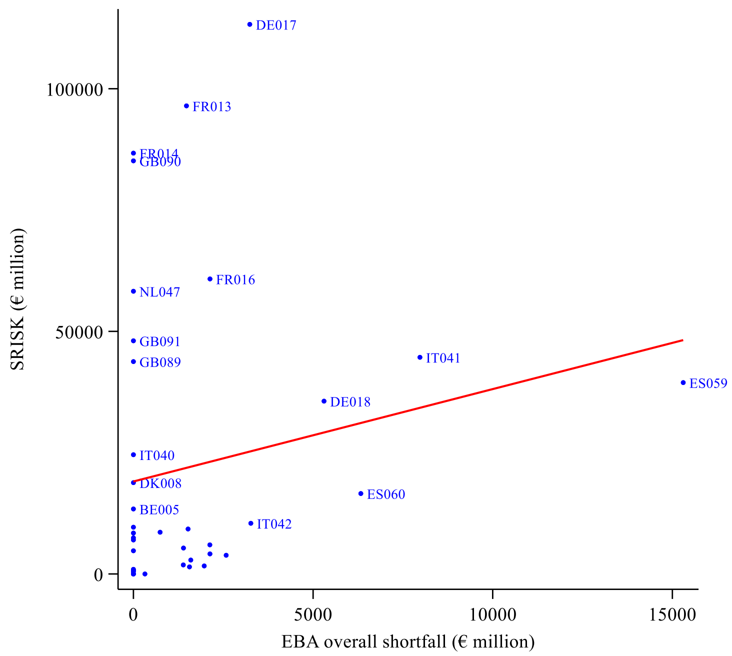


Figure 4. EBA overall shortfall based on 2011 EBA Capital Exercise data disclosed in December 2011 (Eq. (4)) vs. SRISK based on V-Lab data with reference date: 30.09.2011.
Source: Figure by G. Duckert

Furthermore, a weak positive correlation suggests that the two metrics assess similar banks as risk-prone to some extent, though this relationship is less pronounced than for the leverage-based capital shortfall. Rank correlations support this observation: while the original study found a non-significant correlation of 0.133, we identify a correlation of 0.238, also not statistically significant (see Table 1). This limited correlation may reflect heightened awareness of capital adequacy following the stress test or the fact that the variables in the shortfall metric were not subjected to a stress scenario, enabling banks to meet requirements despite the addition of BuffSov. The overall shortfall can also be traced back to the disclosed capital shortfall, reinforcing concerns about RWAs as a calculation basis. Nevertheless, the replication confirms the distribution of banks and the positive correlation observed in the original study by Acharya, Engle, and Pierret (2014) (see Appendix C). Overall, the analyses of shortfall metrics validate the findings of Acharya, Engle, and Pierret (2014), demonstrating that market-based risk models like SRISK and regulatory approaches provide complementary perspectives on banking risks. While regulatory metrics focus on compliance with standards, SRISK captures systemic risk dimensions. The strong positive correlation between SRISK and the leverage-based shortfall further underscores the value of alternative metrics in regulatory and systemic risk analysis.

4.3 Regression analysis of risk factors in the replication

Before presenting the comparison and main findings from the regression analysis, it is essential to reflect on the type of regression employed. The benchmark methodology utilizes a cross-sectional regression, which examines the empirical relationship between variables at a specific point in time. This approach differs from time-series regression, which studies relationships over a period (Wang and Cheng 2020, 65). The regression model is specified as follows:

$$\begin{aligned} &\text{Six-month realized volatility} \\ &= \beta_0 + \beta_1\,(\text{Book-to-market-ratio}) + \beta_2\,(\text{V-Lab risk weight}) \\ &+ \beta_3\,(\text{EBA Tier 1 leverage ratio}) + \beta_4\,(\text{EBA risk weight}) + \varepsilon \end{aligned} \tag{5}$$

The dependent variable, six-month realized volatility, is computed using the following formula:

$$RV_{i,t,W} = \sqrt{\frac{1}{W}\sum_{t+1}^{t+1+W} Var_{i,t}}\ , \tag{6}$$

and relies on V-Lab variance data for bank i at time t, June 30, 2011, over W = 130 days, consistent with the original study. Additionally, the following independent variables need to be considered: (a) the book-to-market ratio, calculated using market capitalization and book values of equity extracted from Bloomberg; (b) the V-Lab risk weight, which incorporates the prudential capital ratio *k* from V-Lab of 5.5% and the *LRMES* on June 30, 2011:

$$\text{V-Lab risk weight} = (1 - (1 - \mathrm{k}) * \mathrm{LRMES})^{-1} \quad ; \tag{7}$$

(c) the EBA Tier 1 leverage ratio (T1LVGR), defined as the ratio of Tier 1 capital to total assets:

$$\text{EBA Tier 1 leverage ratio} = \frac{\mathrm{T1C_S}}{\mathrm{TA_S}} \quad ; \tag{8}$$

and (d) the EBA risk weight, defined as the ratio of risk-weighted assets to total assets:

$$\text{EBA risk weight} = \frac{\mathrm{RWA_S}}{\mathrm{TA_S}} \quad . \tag{9}$$

To account for heteroskedasticity, White's heteroskedasticity-consistent standard errors (robust standard errors) are applied in the regressions.

Our analysis focuses on the full-length cross-sectional regression, incorporating all independent variables. While the original regression (see Appendix E) includes data from 53 banks, our replication is based on 44 banks. This reduction in our sample is primarily due to insufficient book-to-market ratio data for eight banks and missing total assets for one bank.

Furthermore, when examining the adjusted R^2 values, the benchmark regression demonstrates a higher value of 44.1%, compared to our research's adjusted R^2 of 28.1% (see Table 2), indicating that the benchmark regression explains a greater proportion of the variance in the dependent variable. Nevertheless, a relatively low adjusted R^2 does not invalidate the results, especially when independent variables remain statistically significant (Ozili 2023, 2).

When comparing model 6 of our regression with that of Acharya, Engle, and Pierret (2014), three of the four independent variables – V-Lab risk weight, EBA T1LVGR, and EBA risk weight – demonstrate consistent trends in both direction and magnitude across the two studies. Despite these similarities, the results reveal very slight variations in the magnitude and statistical significance of the coefficients.

Table 2: Cross-sectional regression estimates of six-month realized volatility following the 2011 EBA stress test disclosure in July 2011. Dependent variable: six-month realized volatility (Eq. (6)). Independent variables: book-to-market ratio with reference date 30.06.2011, V-Lab risk weight (Eq. (7)) with V-Lab reference date 30.06.2011, EBA T1LVGR (Eq. (8)) at the end of the EBA stress scenario, EBA risk weight (Eq. (9)) at the end of the EBA stress scenario. White's heteroskedasticity-consistent standard errors are reported in parentheses. Sample size: 44.

		1	2	3	4	5	6
	Constant	4.655***	1.593	6.525***	5.370***	3.275	0.713
		(0.396)	(1.866)	(0.735)	(1.082)	(1.776)	(2.010)
	Book-to-market	0.118	0.0327	0.0506	0.113	-0.0437	-0.101
		(0.168)	(0.172)	(0.156)	(0.169)	(0.158)	(0.152)
V-Lab	Risk weight		1.700			1.837	2.445*
			(1.013)			(0.920)	(0.911)
EBA	T1LVGR			-34.14**		-35.21**	-58.00***
				(11.65)		(11.26)	(14.46)
EBA	Risk weight				-1.256		4.767*
					(1.766)		(2.039)
	R^2	0.0117	0.0751	0.183	0.0237	0.257	0.348
	Adj. R^2	-0.0119	0.0300	0.143	-0.0239	0.201	0.281
	F-test	0.495	1.666	4.589	0.498	4.610	5.210

* p<0.05, ** p<0.01, *** p<0.001

Source: Table by G. Duckert

In the benchmark study, the book-to-market ratio is statistically significant at the 1% level, with a positive coefficient of 0.04. This suggests that higher book-to-market ratios, typically indicating undervalued firms or financial stability, are associated with increased volatility. However, our replication finds the variable to be statistically insignificant, with a small negative coefficient of -0.101. This discrepancy, which represents the most notable deviation, is likely attributable to the reduced sample size, mainly due to the missing book-to-market values.

In the original analysis, the V-Lab risk weight is significant at the 1% level, with a coefficient of 2.99, showing a positive relationship with realized volatility. This effect is marginally smaller in our research, where the risk weight remains significant with a slightly reduced coefficient of 2.445. Both studies confirm that market-based measures effectively capture aspects of volatility.

The T1LVGR emerges as a critical variable, consistently negative and statistically significant across both studies. The benchmark study reports a coefficient of -62.44, significant at the 5% level, while our replication finds a coefficient of -58.00, significant at the 0.1% level. This alignment highlights the strong calming effect of higher capital adequacy on market volatility.

Finally, the EBA risk weight shows contrasting results between the two regressions. In the benchmark study, it is statistically insignificant, with a coefficient of 3.56, indicating little explanatory power for volatility. Conversely, our research identifies the variable as significant at the 5% level, with a positive coefficient of 4.767. This suggests that, within our research, regulatory risk weights play a more substantial role in explaining volatility, possibly reflecting heightened investor sensitivity to EBA-defined metrics in the context of our sample.

From model 4 to model 6, our replication reveals a shift in the sign of the EBA risk weight coefficient, from negative to positive. This supports the findings of Acharya, Engle, and Pierret (2014, 51), indicating that regulatory risk weights can be more informative when multiple influential factors are incorporated into the regression.

To conclude, our replication confirms the core findings of the benchmark study, demonstrating consistency in the direction and magnitude of three out of four independent variables. The primary divergence arises in relation to the book-to-market ratio, highlighting the impact of sample size on regression results. Overall, our analysis shows that market-based and regulatory risk factors effectively explain realized volatility, and that these results are consistently replicable.

5 Extension results

This section presents the results of the extended analysis for the 2021 stress test year. The overarching objective was not only to verify the findings of the replication and the original study but also to assess these results in the context of changing regulatory frameworks, economic developments, and market-oriented risk measures. The findings in Section 4 largely supported the conclusions of Acharya, Engle, and Pierret (2014). Thus, the analysis of later years seeks to determine whether these conclusions remain valid or whether significant changes have occurred. Accordingly, the results are presented and discussed in line with the main study objectives and incorporate metrics defined in Section 4, along with necessary adjustments for comparability. This section is organized as follows: Subsection 5.1 briefly updates the metrics applied in the 2021 analysis. Subsection 5.2 covers the extension for 2021, including a brief introduction, the underlying bank sample, and an analysis of the relationship between regulatory and market-based capital shortfall metrics, followed by the regression analysis of risk factors.

5.1 Adjustments to metrics and variables in the extension study

The extended analysis employs methodologies consistent with the original study and replication, applying data from the EBA and V-Lab. However, structural changes in the 2021 EBA stress test datasets necessitate adjustments to both shortfall metrics and regression variables, as certain components familiar from the 2011 EBA stress test are no longer published, renamed, or slightly modified.

Adjustments to the disclosed and absolute capital shortfall metrics account for changes in data availability and definitions of key components, ensuring methodological consistency despite structural differences. Since RWAs are no longer published, they are replaced by the total risk exposure amount (TREA) at the end of each stress scenario. TREA, which aggregates credit, market, and operational risks (EBA 2021a, 17–18), corresponds to RWAs under Basel II (Basel Committee on Banking Supervision 2004, 12–13). Reflecting regulatory changes under Basel III, Common Equity Tier 1 is adopted as a proxy, aligning closely with the characteristics of Core Tier 1 capital under Basel II, thus preserving key principles of loss absorption and financial resilience (Baker, Cummings, and Jagtiani 2017, 257).

The leverage-based capital shortfall, also adjusted for the CET1 definition, requires a new data source for total assets, as these are not disclosed in the 2021 stress test. Therefore, actual total assets from the respective base year of the stress test serve as the best estimate. This proxy appears optimal for two reasons. First, insights from the 2011 stress test indicate that total assets under a stress scenario change only marginally over time compared to the base year. Second, the only variable in recent stress tests derived from total assets is the total leverage ratio exposure, which can be assumed constant (EBA 2021a, 154). The corresponding actual total asset values are sourced from Bloomberg for all sample banks.

Adjustments to the independent regression variables focus on the T1LVGR and EBA risk weight. Tier 1 capital from the stress test results remains the primary input for T1LVGR, while TREA replaces RWAs in the calculation of the EBA risk weight. To ensure coherence with shortfall metrics, actual total assets are used as the denominator for both variables.

These methodological adjustments, detailed further in Appendix F, ensure that the analysis remains rigorous and consistent under evolving regulatory and reporting frameworks.

5.2 Extension 2021

The 2021 stress test was chosen to examine the validity of the methods and conclusions of the original study by Acharya, Engle, and Pierret (2014) and its replication under altered conditions, as it marks a significant milestone in the evolution of the European banking landscape. Since the GFC, systemic stress tests have been conducted and refined regularly, shifting their focus from the recapitalization needs of individual banks to the monitoring of capital adequacy compliance and the formulation of regulatory measures (Baudino et al. 2018, 6).

In addition, the European Single Supervisory Mechanism (SSM) was introduced during this period, enhancing banking supervision within the Eurozone and strengthening the resilience of the European banking system (Vodenska et al. 2021, 1). Supervisory authorities were also equipped with a broader toolkit to dynamically address risks. One example is the countercyclical capital buffer, which adjusts capital requirements based on financial cycles (Davydiuk 2017, 1). A major influence on the banking system has been the establishment of the Basel Accord. Following the GFC, it introduced stricter quality requirements for capital, supplemented by the leverage ratio as a backstop mechanism and in response to the limitations of risk-weighted capital requirements (Baker, Cummings, and Jagtiani 2017, 257). Basel III also introduced a variety of capital buffers, applicable to all banks as well as specific buffers for systemically significant institutions, addressing the structural dimension of systemic risk posed by large, interconnected banks (Anderson et al. 2018, 40).

Between the 2011 and 2021 stress tests, several macroeconomically relevant major events occurred. These include the Eurozone debt crisis of the early 2010s, which saw increases in sovereign risk as banks holding sovereign debt from affected countries experienced significant balance sheet deterioration (Samarakoon 2017, 116). Banks have also been grappling with prolonged low interest rates, including periods of negative rates, prompting adjustments to their business models. Geopolitical uncertainties in the past decade, such as Brexit, the US-China trade war, and Russia's annexation of Crimea, have significantly affected the real economy and financial markets in Europe (Hanisch 2020, 81).

The 2021 stress test took place in the wake of the global COVID-19 pandemic, referred to as the first substantial challenge to the regulatory system developed post-GFC (Duncan et al. 2022, 1). The stress test itself is designed to assess the resilience of the banking system amid COVID-19, incorporating a general three-year macroeconomic downturn scenario that includes and evaluates the effects of a prolonged pandemic situation (EBA 2021b, 12–13). The market disruptions and persistent nature of the pandemic make the 2021 stress test an especially suitable basis for further analysis, particularly as surprising results were observed in the U.S. banking system, where no banking crisis or financial distress occurred, and banks withstood the shock (Duncan et al. 2022, 1; Berger and Demirgüç-Kunt 2021, 4).

By applying Acharya, Engle, and Pierret’s (2014) methodology to the 2021 stress test, this analysis aims to assess whether their findings hold under vastly different macroeconomic and regulatory conditions and whether the results from the U.S. banking system can be confirmed.

5.2.1 Bank sample in the 2021 extension

For 2021, the starting point is the bank sample included in the 2021 stress test, encompassing 50 European banks (EBA 2021b, 10). Banks without publicly available market data were excluded by first identifying matches based on names, resulting in a sample of 31 banks. Subsequently, tickers and ISINs were compared, which revealed no changes. Finally, the banks were investigated further, incorporating insights from the replication. First, the Nordea Bank was found in the selection, correctly listed as a Finnish bank in the EBA data, and thus remained in the sample. Second, the name Bank Pekao SA in the V-Lab dataset was mapped to Polska Kasa Opieki SA in the EBA stress test 2021, as the V-Lab name is a well-known abbreviation (Bloomberg 2024). This increased the sample size to 32 banks, which serves as the basis for conducting the analyses. Appendix G contains a list of all banks included.

5.2.2 Capital shortfall metrics in the 2021 extension

The initial investigation of the 2021 extension examines capital shortfall metrics derived from regulatory and market-based approaches. Unlike in 2011, data on capital buffers for sovereign debt exposures is unavailable, leading to the omission of the EBA overall shortfall metric this year. As explained in Subsection 5.1, adjustments were introduced to the underlying metrics to reflect changes in the EBA dataset. In contrast, the publicly available V-Lab market data remained unchanged, with only the reference date updated. This ensures a consistent basis for subsequent analyses and comparisons with the original study and the replication.

The first finding, while comparing disclosed capital shortfall and SRISK, indicates that most banks report a regulatory shortfall of zero in 2021 (see Figure 8), with only one outlier bank showing some dispersion.

However, SRISK reveals significant capital shortfalls for certain banks, exceeding 50 billion euros in some cases. This divergence highlights a key issue in assessments: while regulatory metrics suggest that banks can withstand stress scenarios and meet the capital requirements, market-based metrics indicate insufficient capitalization for systemic risks in several instances.

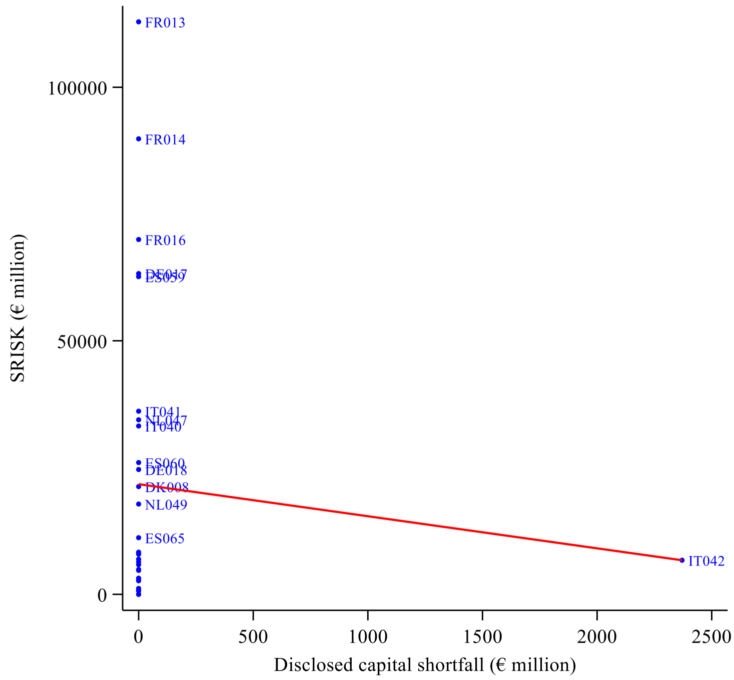


Figure 5. Disclosed capital shortfall based on 2021 EBA stress test data disclosed in July 2021 (see Appendix F, Eq. (10)) vs. SRISK based on V-Lab data with reference date: 31.12.2020.
Source: Figure by G. Duckert

Second, a weak negative correlation between the two metrics is observed, with a rank correlation of -0.029 (see Table 4), which is not statistically significant.

Table 3: Rank correlations between capital shortfall metrics based on 2021 EBA stress test data disclosed in July 2021 vs. SRISK based on V-Lab data with reference date: 31.12.2020.

Disclosed capital shortfall	Absolute capital shortfall	Leverage-based capital shortfall
-0.0292	-0.5979***	0.8707***

* p<0.05, ** p<0.01, *** p<0.001

Source: Table by G. Duckert

This suggests statistical independence, and it can essentially be concluded that no meaningful correlation exists, as the result is heavily influenced by one outlier. Nonetheless, it can still be stated that banks with high SRISK values tend to show no regulatory shortfall.

To explore the frequent zero values in regulatory metrics, the absolute capital shortfall is analyzed. A strong negative correlation with SRISK is identified, confirming the previous statement and further extending it by showing that banks classified as undercapitalized under SRISK appear well-capitalized in regulatory terms (see Figure 9).

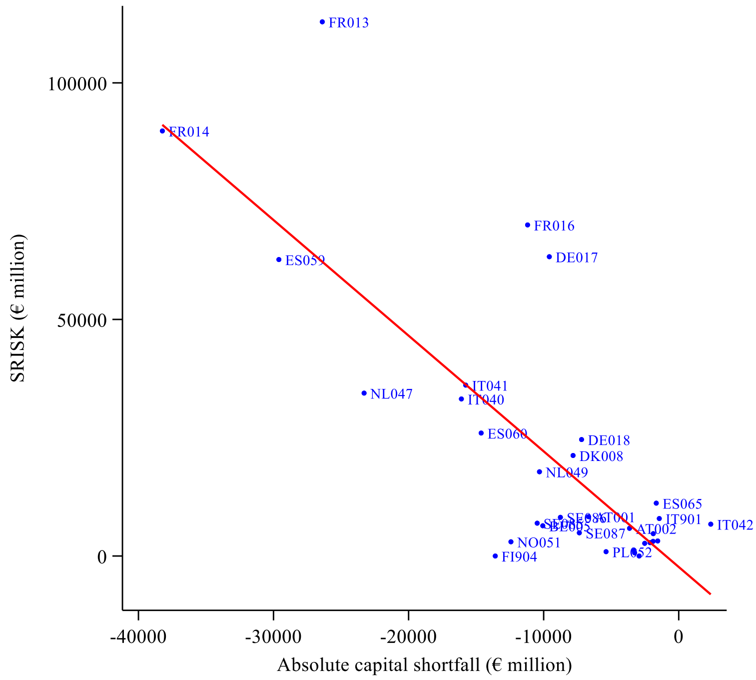


Figure 6. Absolute capital shortfall based on 2021 EBA stress test data disclosed in July 2021 (see Appendix F, Eq. (11)) vs. SRISK based on V-Lab data with reference date: 31.12.2020.
Source: Figure by G. Duckert

This indicates that banks hold excess capital beyond regulatory requirements, as evidenced by a statistically significant rank correlation coefficient of -0.598 (see Table 4). The results reinforce the discrepancy between market-based and regulatory metrics, particularly for large, systemically important banks. While regulatory metrics suggest adequate capitalization, SRISK identifies significant vulnerabilities.

Both comparisons – SRISK vs. disclosed and absolute capital shortfalls – corroborate the findings from the original study and the replication, particularly that banks deemed safest under regulatory stress tests are often those with the highest SRISK shortfalls during stress events (Acharya, Engle, and Pierret 2014, 49). A comparison of the disclosed capital shortfall evolution from 2011 to 2021 reveals that most banks still show zero shortfalls. However, rank correlations suggest a trend towards statistical independence between SRISK and the regulatory metric.

Overall, the results and conclusions remain consistent over the years 2011, both in the replication and the original study, and in 2021. The persistence of these results highlights the limitations of regulatory metrics in insufficiently accounting for systemic risk and their dynamic nature and underpins doubts about the lack of transparency of risk-weighted metrics. Capital shortfall calculations under Basel III, benefiting from stricter RWA requirements, higher CET1 ratios, and countercyclical buffers, have improved bank capital positions (Basel Committee on Banking Supervision 2010, 27), likely explaining the low regulatory shortfalls observed in 2021. However, questions remain regarding the EBA's scenario selection, its potentially conservative assumptions, and how bank behavior may influence stress test outcomes, given that tests are conducted regularly, and banks can adjust accordingly. Lastly, consistent results across 2011 and 2021 suggest that, despite regulatory changes, the fundamental issue of underestimating systemic risk for large, interconnected banks remains unresolved.

The relationship between SRISK and the leverage-based capital shortfall reveals a highly significant and strong positive correlation, with a rank correlation coefficient of 0.871 (see Table 4). This indicates that banks identified as undercapitalized by SRISK are similarly flagged under the leverage-based framework, with capital shortfalls reaching up to 70 billion euros in the leverage-based measure and exceeding 100 billion euros in SRISK (see Figure 10). Both metrics highlight vulnerabilities among large, systemically important institutions, suggesting that total assets align more closely with market-based assessments than RWAs. These findings underscore the systemic importance of large banks, which dominate both metrics and highlight the potential underestimation of systemic risk in traditional, RWA-based frameworks.

The 2021 findings align with those of 2011, confirming not only the replication but also the conclusions of Acharya, Engle, and Pierret (2014). They argue that leverage-based and market-based measures capture bank risk and systemic risk more effectively, as they better account for interconnectedness and changing risk dynamics compared to RWA-based metrics (Acharya, Engle, and Pierret 2014, 52). The shift towards total assets also aligns with regulatory changes under Basel III, which introduced the leverage ratio to address the limitations of RWA-based approaches and curb excessive leverage (Basel Committee on Banking Supervision 2010, 4). Additionally, the use of leverage-based metrics supports the call for simpler, more transparent methods that more accurately reflect systemic risks (Danielsson et al. 2012, 26). These results reflect broader trends in regulatory stress testing, making it unsurprising that the EBA's 2021 stress test further incorporated the leverage ratio alongside other measures (EBA 2021a, 15).

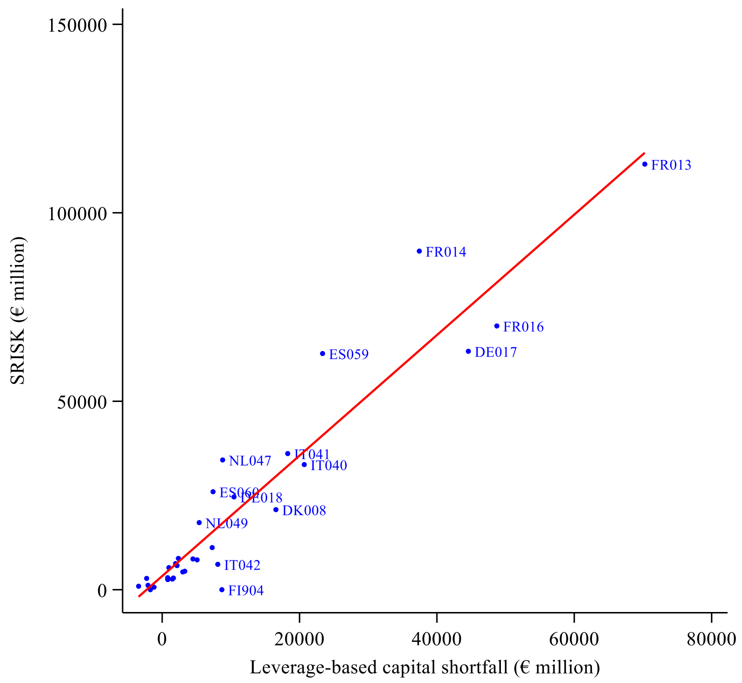


Figure 7. Leverage-based capital shortfall based on 2021 EBA stress test data disclosed in July 2021 (see Appendix F, Eq. (12)) vs. SRISK based on V-Lab data with reference date: 31.12.2020.
Source: Figure by G. Duckert

5.2.3 Regression analysis of risk factors in the 2021 extension

This subsection evaluates the relative importance of regulatory and market-based factors in predicting risks during stress scenarios, using a cross-sectional regression based on the 2021 EBA stress test data. The analysis requires recalculating the realized volatility (see Equation (6)) based on t, June 30, 2021, over a period of $W = 131$ days. Furthermore, the regression methodology and the treatment of standard errors remained consistent, with only slight adjustments to the regulatory risk factors T1LVGR and EBA risk weight (see Section 5.1).

With regard to R^2, a positive development of the values from 36.7% in the first model to 55.8% in the sixth model can be observed (see Table 5). Adjusted R^2 shows a similar trend, indicating that the inclusion of additional variables leads to an improved explanation of realized volatility.

Table 4: Cross-sectional regression estimates of six-month realized volatility following the 2021 EBA stress test disclosure in July 2021. Dependent variable: six-month realized volatility (Eq. (6)). Independent variables: book-to-market ratio with reference date 30.06.2021, V-Lab risk weight (Eq. (7)) with V-Lab reference date 30.06.2021, EBA T1LVGR (see Appendix F, Eq. (13)) at the end of the EBA stress scenario, EBA risk weight (see Appendix F, Eq. (14)) at the end of the EBA stress scenario. White's heteroskedasticity-consistent standard errors are reported in parentheses. Sample size: 32.

		1	2	3	4	5	6
	Constant	1.700***	1.653***	1.150***	1.277***	0.757	0.788
		(0.0941)	(0.399)	(0.165)	(0.136)	(0.401)	(0.403)
	Book-to-market	0.168**	0.169**	0.262***	0.186***	0.281***	0.234**
		(0.0489)	(0.0501)	(0.0487)	(0.0461)	(0.0489)	(0.0659)
V-Lab	Risk weight		0.0226			0.164	0.172
			(0.183)			(0.162)	(0.161)
EBA	T1LVGR			7.841***		8.636***	3.629
				(2.135)		(2.149)	(4.251)
EBA	Risk weight				0.977**		0.696
					(0.297)		(0.551)
	R^2	0.367	0.367	0.519	0.532	0.537	0.558
	Adj. R^2	0.346	0.324	0.486	0.500	0.487	0.492
	F-test	11.75	5.775	15.28	13.03	11.51	10.02

* p<0.05, ** p<0.01, *** p<0.001

Source: Table by G. Duckert

In model 6, considering all factors, the adjusted R^2 reaches 49.2%, one of its highest values, demonstrating a strong explanation of the variance in the dependent variable. In comparison, the replication shows an adjusted R^2 of 28.1% (see Table 2), and the original study reports 44.1% (see Appendix E). This increase reflects the incorporation of variables that better capture the changing dynamics of financial markets and stress test outcomes over time.

The book-to-market variable shows a small but relatively consistent value across all models, ranging from 0.168 to 0.281, and remains statistically significant at the 1% level. This constancy and significance at the 1% level are also observed in the original study but are not maintained in the replication. Nevertheless, this reflects a strong finding suggesting that the book-to-market ratio reliably explains differences in volatility in the stress test year 2021. The similar results in the 2021 regression and in Acharya, Engle, and Pierret's (2014) original study, despite not being observed in the replication, suggest that firm valuation remains a key determinant of market responses to stress tests, although its influence appears to have diminished somewhat over time. While the V-Lab risk factor shows a noticeable increase in its coefficient in the 2011 replication, achieving statistical significance at the 5% level in model 6, its impact diminishes in the 2021 analysis. In the original study, the V-Lab risk weight was a prominent factor, with a positive coefficient of 2.99 in model 6, statistically significant at the 1% level. By 2021, however, this factor becomes statistically insignificant across all models, with considerably smaller coefficients between 0.023 and 0.172. This decline suggests a reduced reliance on market-based risk measures, as regulatory frameworks like Basel III introduced more standardized and transparent methods for assessing bank risk. The lack of significance in the 2021 analysis further implies that the V-Lab risk weight no longer meaningfully impacts realized volatility.

A striking result is the change in sign for the T1LVGR risk factor, which was negative with relatively high absolute values in both the replication and the original study, whereas the 2021 regression shows consistently positive and lower coefficients. In the original 2011 study, model 6 highlighted T1LVGR as significant at the 5% level, with a large negative coefficient of -62.44, reflecting the market's reliance on capital adequacy as a stabilizing factor in early regulatory stress testing. By 2021, the coefficient drops to 3.629 and lacks statistical significance, signaling a much weaker relationship. This shift may seem counterintuitive, as T1LVGR now shows a slight positive influence on realized volatility rather than the calming effect it exhibited in earlier analyses. However, it could also reflect the general adaptation to leverage ratios across the banking sector following regulatory reforms like Basel III, making this metric less distinguishing among banks in recent years.

A key finding of Acharya, Engle, and Pierret's (2014) study was that the EBA risk weight only contributes information about realized volatility when other, more important factors are considered. While this finding was largely supported by the replication, the 2021 analysis reveals entirely different results. In model 4, the coefficient is 0.977 and significant at the 1% level, but in model 6, it decreases to 0.696 and becomes statistically insignificant. Hence, the EBA risk weight loses part of its explanatory power and is less relevant in 2021 than in 2011. Finally, the constant decreases across the models and is not significant in model 6, which was already indicated in the replication and the original study.

The comparison provides valuable insights into how the determinants of six-month realized volatility have shifted over the past decade. While certain variables remain influential over time, the results also reveal significant changes in their impact and importance. In conclusion, regarding the key findings of the original study, we cannot fully support the statement that the EBA risk weight only contributes new information when other risk factors are accounted for (Acharya, Engle, and Pierret 2014, 46), as we specifically observe in model 4 of the 2021 regression that the coefficient is positive and, when considered alone, explains more than in model 6. However, we can make a similar assertion, as we observe that additional variables provide substantial explanatory power, as the adjusted R^2 increases across models. Furthermore, the T1LVGR changes its direction of influence and remains positive in its coefficients across all models, which contradicts the original conclusion of Acharya, Engle, and Pierret (2014). Ultimately, T1LVGR and book-to-market remain the driving factors of realized volatility, with the latter having a statistically significant impact, providing new insights into risk dynamics. These findings underscore the dynamic nature of financial markets and the necessity for stress testing methodologies to adapt to evolving conditions.

Overall, the results confirm central insights from the literature. The relatively strong relationship between T1LVGR and volatility, suggesting that higher leverage tends to be associated with higher volatility, is consistent with findings that leverage not only creates volatility but also drives systemic risk (Phelan 2016, 219). Additionally, the significant relationship between book-to-market and volatility could indicate economies of scale, as larger banks typically have higher exposure to systemic risks, leading to higher volatility (Moch 2018, 253). The less consistent influence of the regulatory EBA risk weight could also reflect the criticism that such indicators do not always adequately capture market-based risks, especially under stress (Baudino et al. 2018, 24).

5.2.4 Results of the 2021 extension in the context of the COVID-19 pandemic

The results of the preceding analysis must be interpreted in light of the extraordinary economic and regulatory conditions brought about by the COVID-19 pandemic. In 2021, global uncertainty reached unprecedented levels, placing structural challenges and immediate pressures on banks. These exceptional circumstances likely influenced the observed trends and outcomes, shaping the financial landscape and regulatory responses during this period.

Under Basel III, banks are encouraged to maintain capital buffers exceeding minimum requirements to address systemic vulnerabilities. Key measures such as the capital conservation buffer, the countercyclical buffer – which can be released during stress periods – and the systemic risk buffer for systemically important banks have collectively improved the banking system's resilience compared to Basel II standards (Giese and Haldane 2020, 205). Consequently, the COVID-19 crisis built upon a more resilient financial system, which may explain why the regulatory capital shortfalls observed in the 2021 EBA stress test were predominantly negligible or nonexistent. Nevertheless, market-based metrics such as SRISK continue to reveal notable capital shortfalls, particularly among systemically important banks. These results align with the broader context of the pandemic, which saw a sharp increase in systemic risks during the early stages of COVID-19 (Rizwan, Ahmad, and Ashraf 2020, 6).

Although these risks diminished over time due to policy responses, the pandemic underscored persistent discrepancies between regulatory and market-based assessments. Importantly, while banks in general were not exposed to significant capital or liquidity shortages, they faced rising credit risks that posed challenges across the sector (Didier et al. 2021, 10).

Furthermore, the COVID-19 pandemic was an exogenous shock, unlike the endogenously driven GFC, and macroprudential frameworks were in place to address such systemic shocks (Rizwan, Ahmad, and Ashraf 2020, 2). These included coordinated regulatory and supervisory measures by the Basel Committee on Banking Supervision, as well as substantial policy interventions at the national level to mitigate the economic fallout. Additionally, the ECB implemented the Pandemic Emergency Purchase Programme (PEPP), which indeed had a calming effect on the financial sector. Prior to this, significant systemic risk contributions were observed for European banks (Borri and Giorgio 2022, 2). A critical success factor in absorbing the shock – one with even greater impact than the actual securities purchase programs – were the ECB’s speeches, which restored market confidence (Delatte and Guillaume 2020, 14).

Stress tests also proved vital during the COVID-19 pandemic. Early in the crisis, regulators conducted ad hoc assessments to evaluate the immediate system-wide impact of COVID-19. These tests, adapted to account for the pandemic's unique challenges, contributed to the banking sector's overall resilience by providing timely insights into vulnerabilities and enabling appropriate policy responses (Ellis, Sharma, and Brzeszczyński 2022, 18).

In conclusion, the results underscore the effectiveness of post-GFC reforms in strengthening the financial sector (Acharya and Steffen 2020, 9). The resilience demonstrated during the COVID-19 pandemic aligns with findings from the U.S. banking system, where banks weathered the crisis without triggering systemic instability.

6 Conclusion

This section concludes the thesis by summarizing the analyses and key findings, addressing the research questions, and reflecting on the constraints of the study. Subsection 6.1 highlights the main results and provides answers to the research questions. Subsection 6.2 discusses the study’s limitations and their implications for interpretation. Finally, Subsection 6.3 offers a forward-looking perspective on the future of stress testing in an evolving financial landscape.

6.1 Main findings and research questions

The primary aim of this thesis was to replicate and extend the analyses conducted by Acharya, Engle, and Pierret (2014) to assess the robustness of their findings and evaluate their relevance under evolving regulatory and market conditions. This study was guided by two main objectives: first, to investigate the relationship between regulatory and market-based capital shortfall metrics, and second, to examine the role of risk factors in predicting realized volatility under stress scenarios.

The analysis of the replication under 2011 stress test data enables us to answer the first research question: *Do the original conclusions of Acharya, Engle, and Pierret's (2014) analysis remain valid when replicated using 2011 EBA stress test data and updated V-Lab benchmark data?*

In general, the replication confirmed the validity of the original study´s conclusions. In particular, consistency in direction and strength of relationships between shortfall metrics aligns with the original findings. Regulatory shortfall metrics continue to exhibit negative correlations with SRISK, supporting the critiques of RWA-based measures and the IRB approach, which tend to underestimate systemic vulnerabilities for large financial institutions (Acharya, Engle, and Pierret 2014, 42). In contrast, leverage-based shortfall metrics maintain strong positive correlations with market-based metrics, reinforcing the statement that leverage-based metrics are more effective in identifying systemic risks (Acharya, Engle, and Pierret 2014, 52).

The replication's regression closely mirrors the original analysis, validating the conclusions about discrepancies between regulatory and market-based risk factors. The T1LVGR and risk weights align with the direction and magnitude of factors observed in the original study, while the book-to-market variable did not exhibit the expected direction in the 2011 replication. However, the findings that size and leverage play a critical role in the systemic importance of banks align with the conclusions of Acharya, Engle, and Pierret (2014). Overall, our results confirm the original study's findings and emphasize the robustness of its methodology in identifying systemic risks.

Extending the analysis to the 2021 stress test enables us to answer the second research question: *Do the replicated findings and the original conclusions of Acharya, Engle, and Pierret's (2014) study remain valid when extended using 2021 EBA stress test data?*

This investigation reveals both continuity and evolution in the results. Relationships and correlations remained largely stable across the periods, indicating a consistent connection between capital shortfall metrics. Market-based SRISK consistently identified significant vulnerabilities overlooked by traditional RWA-based approaches, supported by the key finding that banks deemed safe under regulatory metrics across years exhibited substantial SRISK shortfalls.

In contrast, the regression result for the later period exhibited notable differences. The explanatory power of the regression from 2021 increased considerably compared to 2011, likely reflecting the impact of Basel III reforms, such as stricter capital requirements and the inclusion of leverage ratios. Key drivers of realized volatility in the later periods were the T1LVGR and book-to-market ratio, with the latter also demonstrating statistical significance. Conversely, the EBA risk weight showed no significant influence on volatility and lost explanatory power when controlled for other variables.

Overall, the 2021 results confirm persistent limitations in regulatory metrics, reinforcing the conclusions from 2011. Market-based and leverage-based approaches remain essential for accurately capturing systemic risks, particularly for large, interconnected banks.

To conclude, the analyses across 2011 and 2021 demonstrate that while regulatory advancements, particularly the transition from Basel II to Basel III, have enhanced the assessment of systemic vulnerabilities over time, market-based metrics like SRISK remain indispensable. The resilience observed during the COVID-19 pandemic underlines the importance of continually refining both regulatory and market-based frameworks, while also highlighting the progress made in addressing the evolving risks of an increasingly interconnected financial system. Combining these approaches provides a comprehensive framework for identifying systemic risks in large institutions and evolving financial landscapes.

6.2 Limitations

Despite its contribution, our study faces limitations that should be considered when interpreting the results. First, a significant number of institutions had to be excluded from the replication analysis due to missing data, particularly for the book-to-market ratio. This limitation reduced the sample size in our regression from 53 to 44 banks and may have introduced a small sample bias, as the exclusion of smaller or less transparent banks distorts the results (Caballero 1994, 52). Consequently, the findings may not fully represent the broader banking sector.

Second, the original study focused on a specific period under Basel II, whereas the later analysis reflected the transition to Basel III. This temporal discrepancy limits the direct comparability of findings across periods and highlights the shift in the objectives of regulatory stress tests, away from purely recapitalization requirements for individual banks toward ensuring capital adequacy across the banking sector (Baudino et al. 2018, 6). Moreover, the 2011 analysis was conducted when the IRB approach and internal models were extensively applied for RWA calculation, often resulting in overly optimistic risk assessments (Le Leslé and Avramova 2012, 7). Thus, the findings for 2011 may not entirely reflect the systemic vulnerabilities that stricter regulatory frameworks would have identified. Lastly, adjustments to metrics, such as replacing RWAs with TREA and using proxies for total assets, were necessary to ensure consistency across the periods. Nonetheless, these adjustments introduced a degree of approximation that could affect the precision of our results (Shepard, Camilli, and Williams 1985, 77).

These limitations highlight the need for further research to address data gaps, refine methods and metrics, and investigate the dynamics of systemic risk in diverse regulatory contexts.

6.3 Future outlook

Stress testing will remain a critical tool for ensuring financial stability, as global markets face growing complexities and interconnected risks. Key trends in stress testing are anticipated to address emerging challenges, improve methodologies, and expand the scope of risk assessment. One major focus will be on climate-related financial risks, with regulators incorporating scenarios to evaluate the impact of extreme weather events and the transition to a low-carbon economy. These stress tests aim to ensure that financial institutions can withstand environmental challenges (EY 2023, 5). Similarly, cybersecurity risks are moving to the forefront of stress testing, as regulators plan tests to assess banks' resilience to cyberattacks and digital disruptions (EY 2023, 5). The evolving macroeconomic environment will make it necessary to include future shock scenarios in stress tests arising from geopolitical tensions, inflation, and global recessions. For instance, scenarios such as severe stress in real estate market and corporate debt defaults will test the resilience of institutions against economic volatility, as the U.S. Federal Reserve is already doing (Federal Reserve 2024, 11). Lastly, the shift towards forward-looking frameworks will enable stress tests to consider arising trends such as decentralized finance, cryptocurrencies, and central bank digital currencies. These frameworks aim to proactively anticipate vulnerabilities and ensure institutions are prepared for rapidly changing market conditions (The Financial Times Limited 2024).

In conclusion, this thesis underscores the continued importance of stress testing as a cornerstone of financial stability. By combining rigorous methodologies with forward-looking perspectives, stress tests can adapt to emerging challenges and address existing concerns, ensuring the resilience of financial institutions and safeguarding global markets.

References

Acharya, Viral, Robert Engle, and Diane Pierret. 2014. "Testing Macroprudential Stress Tests: The Risk of Regulatory Risk Weights." *Journal of Monetary Economics* 65 (July):36–53. https://doi.org/10.1016/j.jmoneco.2014.04.014.

Acharya, Viral, and Sascha Steffen. 2020. "Stress Tests' for Banks as Liquidity Insurers in a Time of COVID." *CEPR Press VoxEU*, March, 1–9.

Anderson, Ron, Chikako Baba, Jon Danielsson, Udaibir S Das, Heedon Kang, and Miguel Segoviano. 2018. "Macroprudential Stress Tests and Policies: Searching for Robust and Implementable Frameworks." *IMF Working Paper WP/18/197*, February, 1–79.

Baker, Colleen, Christine Cummings, and Julapa Jagtiani. 2017. "The Impacts of Financial Regulations: Solvency and Liquidity in the Post-Crisis Period." *Journal of Financial Regulation and Compliance* 25 (3): 253–70. https://doi.org/10.1108/JFRC-02-2017-0027.

Baltagi, Badi H. 2021. *Econometric Analysis of Panel Data*. 6th ed. 2021. Springer Texts in Business and Economics. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-53953-5.

Basel Committee on Banking Supervision. 2004. "Basel II: International Convergence of Capital Measurement and Capital Standards." https://www.bis.org/publ/bcbs107.pdf.

Basel Committee on Banking Supervision. 2010. "Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems." https://www.bis.org/publ/bcbs189.pdf.

Baudino, Patrizia, Roland Goetschmann, Jérôme Henry, Ken Taniguchi, and Weisha Zhu. 2018. *Stress-Testing Banks: A Comparative Analysis*. FSI Insights on Policy Implementation, no 12 (November 2018). Basel: Bank for International Settlements, Financial Stability Institute.

Beckert, Jens. 2010. "Die Finanzkrise Ist Auch Eine Vertrauenskrise." *Gesellschaftsforschung* 1:9–13.

Berger, Allen N., and Asli Demirgüç-Kunt. 2021. "Banking Research in the Time of COVID-19." *Journal of Financial Stability* 57 (December):1–20. https://doi.org/10.1016/j.jfs.2021.100939.

Bloomberg. 2024. "Bank Polska Kasa Opieki SA." November 16, 2024. https://www.bloomberg.com/profile/company/BKPKF:US.

Bolton, Patrick, Morgan Despres, Luiz Awazu Pereira da Silva, Frédéric Samama, and Romain Svartzman. 2020. "The Green Swan - Central Banking and Financial Stability in the Age of Climate Change," January, 1–107.

Bookstaber, Rick, Jill Cetina, Greg Feldberg, Mark Flood, and Paul Glasserman. 2013. "Stress Tests to Promote Financial Stability: Assessing Progress and Looking to the

Future." *Journal of Risk Management in Financial Institutions* 7 (1): 1–13. https://doi.org/10.69554/SVZG1421.

Borio, Claudio, Mathias Drehmann, and Kostas Tsatsaronis. 2014. "Stress-Testing Macro Stress Testing: Does It Live up to Expectations?" *Journal of Financial Stability* 12 (June):3–15. https://doi.org/10.1016/j.jfs.2013.06.001.

Borri, Nicola, and Giorgio Di Giorgio. 2022. "Systemic Risk and the COVID Challenge in the European Banking Sector." *Journal of Banking & Finance* 140 (July):1–13. https://doi.org/10.1016/j.jbankfin.2021.106073.

Brownlees, Christian, and Robert F. Engle. 2017. "SRISK: A Conditional Capital Shortfall Measure of Systemic Risk." *Review of Financial Studies* 30 (1): 48–79. https://doi.org/10.1093/rfs/hhw060.

Caballero, Ricardo J. 1994. "Small Sample Bias and Adjustment Costs." *The Review of Economics and Statistics* 76 (1): 52–58. https://doi.org/10.2307/2109825.

Čihák, Martin. 2007. "Introduction to Applied Stress Testing." *IMF Working Paper WP/07/59*, March, 1–74.

Covas, Francisco. 2020. "Challenges in Stress Testing and Climate Change," October, 1–9.

Crisan, Lucian Ciprian. 2014. "The Sress Test - A New Challenge for the Banking Union." In *Proceedings of the 12th International Academic Conference, Prague*, 255–60. Prague, Czech Republic.

Danielsson, Jon, Kevin R James, Marcela Valenzuela, and Ilknur Zer. 2012. "Dealing with Systemic Risk When We Measure It Badly," February, 1–35.

Davydiuk, Tetiana. 2017. "Dynamic Bank Capital Requirements." *2017 Meeting Papers*, Society for Economic Dynamics, , no. 1328, 1–79. https://doi.org/10.2139/ssrn.3110800.

Delatte, Anne-Laure, and Alexis Guillaume. 2020. "Covid 19: A New Challenge for the EMU?" *CEPII Working Paper*, no. 2020–08 (July), 1–25.

Dent, Kieran, Ben Westwood, and Miguel Segoviano. 2016. "Stress Testing of Banks: An Introduction." *Bank of England Quarterly Bulletin 2016 Q3*, September, 130–43.

Didier, Tatiana, Federico Huneeus, Mauricio Larrain, and Sergio L. Schmukler. 2021. "Financing Firms in Hibernation during the COVID-19 Pandemic." *Journal of Financial Stability* 53 (April):1–14. https://doi.org/10.1016/j.jfs.2020.100837.

Duncan, Elizabeth, Akos Horvath, Diana Iercosan, Bert Loudis, Alice Maddrey, Francis Martinez, Timothy Mooney, et al. 2022. "COVID-19 as a Stress Test: Assessing the Bank Regulatory Framework." *Journal of Financial Stability* 61 (August):101016. https://doi.org/10.1016/j.jfs.2022.101016.

EBA. 2011a. "2011 EU-Wide Stress Test: Aggregate Report." https://www.eba.europa.eu/risk-and-data-analysis/risk-analysis/eu-wide-stress-testing/eu-wide-stress-testing-2011.

EBA. 2011b. "2011 EU-Wide Stress Test: Methodological Note - Additional Guidance." https://www.eba.europa.eu/risk-and-data-analysis/risk-analysis/eu-wide-stress-testing/eu-wide-stress-testing-2011.

EBA. 2011c. "Results of the 2011 EU-Wide Stress Test under the Adverse Scenario."

EBA. 2012. "Final Report on the Implementation of Capital Plans Following the EBA's 2011 Recommendation on the Creation of Temporary Capital Buffers to Restore Market Confidence."

EBA. 2021a. "2021 EU-Wide Stress Test: Methodological Note." https://www.eba.europa.eu/risk-and-data-analysis/risk-analysis/eu-wide-stress-testing/stress-tests-2021.

EBA. 2021b. "2021 EU-Wide Stress Test: Results." https://www.eba.europa.eu/risk-and-data-analysis/risk-analysis/eu-wide-stress-testing/stress-tests-2021.

ECB. 2013. "A Macro Stress Testing Framework for Assessing Systemic Risks in the Banking Sector." *Occasional Paper Series*, no. 152, 1–80. https://doi.org/10.2139/ssrn.2337894.

Ellis, Scott, Satish Sharma, and Janusz Brzeszczyński. 2022. "Systemic Risk Measures and Regulatory Challenges." *Journal of Financial Stability* 61 (August):1–22. https://doi.org/10.1016/j.jfs.2021.100960.

EY. 2023. "2024 Global Financial Services Regulatory Outlook." https://www.ey.com/en_nl/insights/banking-capital-markets/how-firms-can-respond-to-the-2024-regulatory-landscape.

Federal Reserve. 2024. "2024 Supervisory Stress Test Methodology," March, 1–92.

Ferrari, Stijn, Patrick Roy, and Crisitina Vespro. 2011. "Stress Testing Credit Risk: Modelling Issues." *Financial Stability Review*, no. 9 (January), 105–20.

Giese, Julia, and Andy Haldane. 2020. "COVID-19 and the Financial System: A Tale of Two Crises." *Oxford Review of Economic Policy* 36 (Supplement_1): 200–214. https://doi.org/10.1093/oxrep/graa035.

Greenlaw, David, Anil K. Kashyap, Kermit L. Schoenholtz, and Hyun Song Shin. 2012. "Stressed Out: Macroprudential Principles for Stress Testing." *Chicago Booth Research Paper*, no. 12–08, 1–51. https://doi.org/10.2139/ssrn.2004380.

Hanisch, Max. 2020. "Iran, Russland, Hongkong: Geopolitische Risiken belasten deutsche Wirtschaft." *DIW Wochenbericht*, no. 6, 80–85. https://doi.org/10.18723/DIW_WB:2020-6-1.

Hirtle, Beverly, and Andreas Lehnert. 2015. “Supervisory Stress Tests.” *Annual Review of Financial Economics* 7 (1): 339–55. https://doi.org/10.1146/annurev-financial-111914-042040.

Hirtle, Beverly, Til Schuermann, and Kevin J. Stiroh. 2009. “Macroprudential Supervision of Financial Institutions: Lessons from the SCAP.” *Federal Reserve Bank of New York Staff Reports*, no. 409, 1–15. https://doi.org/10.2139/ssrn.1515800.

Kalhoefer, Christian, and Guenter Lang. 2019. “The More the Merrier? Detecting Impacts of Bank Regulation After the Global Financial Crisis.” *Credit and Capital Markets – Kredit Und Kapital* 52 (2): 191–212. https://doi.org/10.3790/ccm.52.2.191.

King, Gary. 1995. “Replication, Replication.” *PS: Political Science & Politics* 28 (3): 444–52. https://doi.org/10.2307/420301.

Lall, Ranjit. 2009. “Why Basel II Failed and Why Any Basel III Is Doomed.” *GEG Working Paper*, no. 2009/52, 1–31.

Lane, Philip R. 2012. “The European Sovereign Debt Crisis.” *Journal of Economic Perspectives* 26 (3): 49–68. https://doi.org/10.1257/jep.26.3.49.

Le Leslé, Vanessa, and Sofiya Avramova. 2012. “Revisiting Risk-Weighted Assets: Why Do RWAs Differ Across Countries and What Can Be Done About It?” *IMF Working Paper WP/12/90*, March, 1–47.

Moch, Nils. 2018. “The Contribution of Large Banking Institutions to Systemic Risk: What Do We Know? A Literature Review.” *Review of Economics* 69 (3): 231–57. https://doi.org/10.1515/roe-2018-0011.

Moreno de Alborán de Cominges, Álvaro. 2016. “Bankia: Its Valuation Using Different Methodologies.” https://elischolar.library.yale.edu/ypfs-documents2/2345.

Nordea Bank Abp. 2023. “Annual Report 2023.” https://www.nordea.com/en/doc/annual-report-nordea-bank-abp-2023.pdf.

Nosek, Brian A., and Timothy M. Errington. 2020. “What Is Replication?” *PLOS Biology* 18 (3): 1–8. https://doi.org/10.1371/journal.pbio.3000691.

Ozili, Peterson K. 2023. “The Acceptable R-Square in Empirical Modelling for Social Science Research.” *MPRA Paper No. 116496*, February, 1–9.

permanent tsb Group Holdings p.l.c. 2024. “Group History.” November 16, 2024. http://www.irishlifeandpermanent.ie/who-we-are/group-history.aspx.

Phelan, Gregory. 2016. “Financial Intermediation, Leverage, and Macroeconomic Instability.” *American Economic Journal: Macroeconomics* 8 (4): 199–224. https://doi.org/10.1257/mac.20140233.

Rizwan, Muhammad Suhail, Ghufran Ahmad, and Dawood Ashraf. 2020. “Systemic Risk: The Impact of COVID-19.” *Finance Research Letters* 36 (October):1–7. https://doi.org/10.1016/j.frl.2020.101682.

Rutkowski, Felix, Alexander Schäfer, and Isabel Schnabel. 2018. “Zehn Jahre nach Ausbruch der Finanzkrise – viel erreicht, noch viel zu tun.” *Vierteljahrshefte zur Wirtschaftsforschung* 87 (3): 9–25. https://doi.org/10.3790/vjh.87.3.9.

Samarakoon, Lalith P. 2017. “Contagion of the Eurozone Debt Crisis.” *Journal of International Financial Markets, Institutions and Money* 49 (July):115–28. https://doi.org/10.1016/j.intfin.2017.03.001.

Schäfer, Dorothea. 2011. “Banken: Leverage Ratio Ist Das Bessere Risikomaß.” *DIW Wochenbericht*, 11–17.

Schmieder, Christian, Claus Puhr, and Maher Hasan. 2011. “Next Generation Balance Sheet Stress Testing.” *IMF Working Paper WP/11/83*, April, 1–42.

Shepard, Lorrie A., Gregory Camilli, and David M. Williams. 1985. “Validity of Approximation Techniques for Detecting Item Bias.” *Journal of Educational Measurement* 22 (2): 77–105. https://doi.org/10.1111/j.1745-3984.1985.tb01050.x.

Sorge, Marco. 2004. “Stress-Testing Financial Systems: An Overview of Current Methodologies.” *BIS Working Paper*, no. 165, 1–36. https://doi.org/10.2139/ssrn.759585.

Steffen, Sascha. 2014. “Robustness, Validity and Significance of the ECB’s Asset Quality Review and Stress Test Exercise.” *SAFE White Paper*, no. 23, 1–45. https://doi.org/10.2139/ssrn.2517361.

Taskinsoy, John. 2013. “Rigorous Capital Requirements Under Basel III: Possible Impact on Turkey’s Financial Sector.” *Journal of WEI Business and Economics* 2 (1): 1–30.

Taskinsoy, John. 2019. “Typology of Stress Testing: Microprudential vs. Macroprudential Stress Testing of Risk Exposures.” *SSRN Electronic Journal*, March, 1–30. https://doi.org/10.2139/ssrn.3361528.

Taskinsoy, John. 2022. “Stress Testing Financial Systems: Macro and Micro Stress Tests, Basel Standards and Value-at-Risk as Financial Stability Measures,” February, 1–55. https://dx.doi.org/10.2139/ssrn.4032869.

The Financial Times Limited. 2024. “Banks’ Use of AI Could Be Included in Stress Tests, Says Bank of England Deputy Governor,” October 31, 2024. https://www.ft.com/content/d4d212a8-c63a-4b00-9f4c-e06ed59f9279.

Vodenska, Irena, Hideaki Aoyama, Alexander P. Becker, Yoshi Fujiwara, Hiroshi Iyetomi, and Eliza Lungu. 2021. “From Stress Testing to Systemic Stress Testing: The Importance of Macroprudential Regulation.” *Journal of Financial Stability* 52 (February):1–25. https://doi.org/10.1016/j.jfs.2020.100803.

Wang, Xiaofeng, and Zhenshun Cheng. 2020. "Cross-Sectional Studies." *Chest* 158 (1): 65–71. https://doi.org/10.1016/j.chest.2020.03.012.

Appendices

Appendix A

EBA metrics and published codes

This appendix contains all necessary information, date, and scenario codes required for the unique identification of variables underlying the metrics. These codes are provided by the EBA with the publication of stress test results in the respective year. For the analyses, we distinguish between capital shortfall measures and regression measures. Additionally, the subscript "S" in the variables indicates that they are considered at the end of the stress scenario. The subscript "B" denotes Actuals sourced from Bloomberg, as these variables are used as proxies. The Actuals, as well as the variables from the capital exercise for the overall shortfall metric, are not subject to any stress scenario.

		Metric	Variable	Information Code	Date Code	Scenario Code
Stress Test 2011	Capital shortfall measure	Disclosed capital shortfall	RWA_S	30025	20121231	105
			$Capital_S$	30035	20121231	105
		Absolute capital shortfall	RWA_S	30025	20121231	105
			$Capital_S$	30035	20121231	105
		Leverage-based capital shortfall	TA_S	30029	20121231	105
			$Capital_S$	30035	20121231	105
		EBA overall shortfall	RWA	100900	20110930	-
			T1C	100600	20110930	-
			BuffSov	101000	20110930	-
	Regression measure	Tier 1 leverage ratio	$T1C_S$	30036	20121231	105
			TA_S	30029	20121231	105
		EBA risk weight	RWA_S	30025	20121231	105
			TA_S	30029	20121231	105
Stress Test 2021	Capital shortfall measure	Disclosed capital shortfall	$TREA_S$	213107	202312	3
			$CET1_S$	213106	202312	3
		Absolute capital shortfall	$TREA_S$	213107	202312	3
			$CET1_S$	213106	202312	3
		Leverage-based capital shortfall	TA_B	-	-	-
			$CET1_S$	213106	202312	3
	Regression measure	Tier 1 leverage ratio	$T1C_S$	213110	202312	3
			TA_B	-	-	-
		EBA risk weight	$TREA_S$	213107	202312	3
			TA_B	-	-	-

Source: Table by G. Duckert

Appendix B

List of banks for the 2011 replication study

This list contains all banks, their codes, and data sources underlying the replication sample.

Bank Code	Bank Name	Stress Test 2011	V-Lab	Capital Exercise 2011
AT001	Erste Group Bank AG	✓	✓	✓
AT002	Raiffeisen Bank International AG	✓	✓	✓
BE004	Dexia SA	✓	✓	
BE005	KBC Groep NV	✓	✓	✓
CY006	Marfin Popular Bank PCL	✓	✓	✓
CY007	Bank of Cyprus Plc	✓	✓	✓
DE017	Deutsche Bank AG	✓	✓	✓
DE018	Commerzbank AG	✓	✓	✓
DE027	Landesbank Berlin Holding AG	✓	✓	✓
DK008	Danske Bank A/S	✓	✓	✓
DK009	Jyske Bank A/S	✓	✓	✓
DK010	Sydbank A/S	✓	✓	✓
ES059	Banco Santander SA	✓	✓	✓
ES060	Banco Bilbao Vizcaya Argentaria SA	✓	✓	✓
ES061	Bankia SA	✓	✓	
ES063	Liberbank SA	✓	✓	
ES064	Banco Popular Espanol SA	✓	✓	✓
ES065	Banco de Sabadell SA	✓	✓	
ES069	Bankinter SA	✓	✓	
ES071	Banca Civica SA	✓	✓	
ES074	Banco Pastor SA	✓	✓	
FI012	Pohjola Bank Oyj	✓	✓	✓
FR013	BNP Paribas SA	✓	✓	✓
FR014	Credit Agricole SA	✓	✓	✓
FR016	Societe Generale SA	✓	✓	✓
GB089	HSBC Holdings PLC	✓	✓	✓
GB090	Barclays PLC	✓	✓	✓
GB091	Lloyds Banking Group PLC	✓	✓	✓
GR030	Eurobank Ergasias Services and Holdings SA	✓	✓	
GR031	National Bank of Greece SA	✓	✓	
GR032	Alpha Services and Holdings SA	✓	✓	
GR033	Piraeus Financial Holdings SA	✓	✓	
GR034	Agricultural Bank of Greece	✓	✓	
GR035	TT Hellenic Postbank SA	✓	✓	
HU036	OTP Bank PLC	✓	✓	✓
IE037	AIB Group PLC	✓	✓	✓
IE038	Bank of Ireland Group PLC	✓	✓	✓
IT040	Intesa Sanpaolo SpA	✓	✓	✓
IT041	UniCredit SpA	✓	✓	✓
IT042	Banca Monte dei Paschi di Siena SpA	✓	✓	✓
IT044	Unione di Banche Italiane SpA	✓	✓	✓
MT046	Bank of Valletta PLC	✓	✓	✓
NL047	ING Groep NV	✓	✓	✓
NL049	ABN AMRO Group NV	✓	✓	✓
NO051	DNB ASA	✓	✓	✓
PL052	PKO Bank Polski SA	✓	✓	✓
PT054	Banco Comercial Portugues, S.A.	✓	✓	✓
PT055	Banco Espirito Santo SA	✓	✓	✓
PT056	Banco BPI SA	✓	✓	✓
SE085	Skandinaviska Enskilda Banken AB	✓	✓	✓
SE086	Svenska Handelsbanken AB	✓	✓	✓
SE087	Swedbank AB	✓	✓	✓
SI057	Nova Ljubljanska Banka dd	✓	✓	✓

Source: Table by G. Duckert

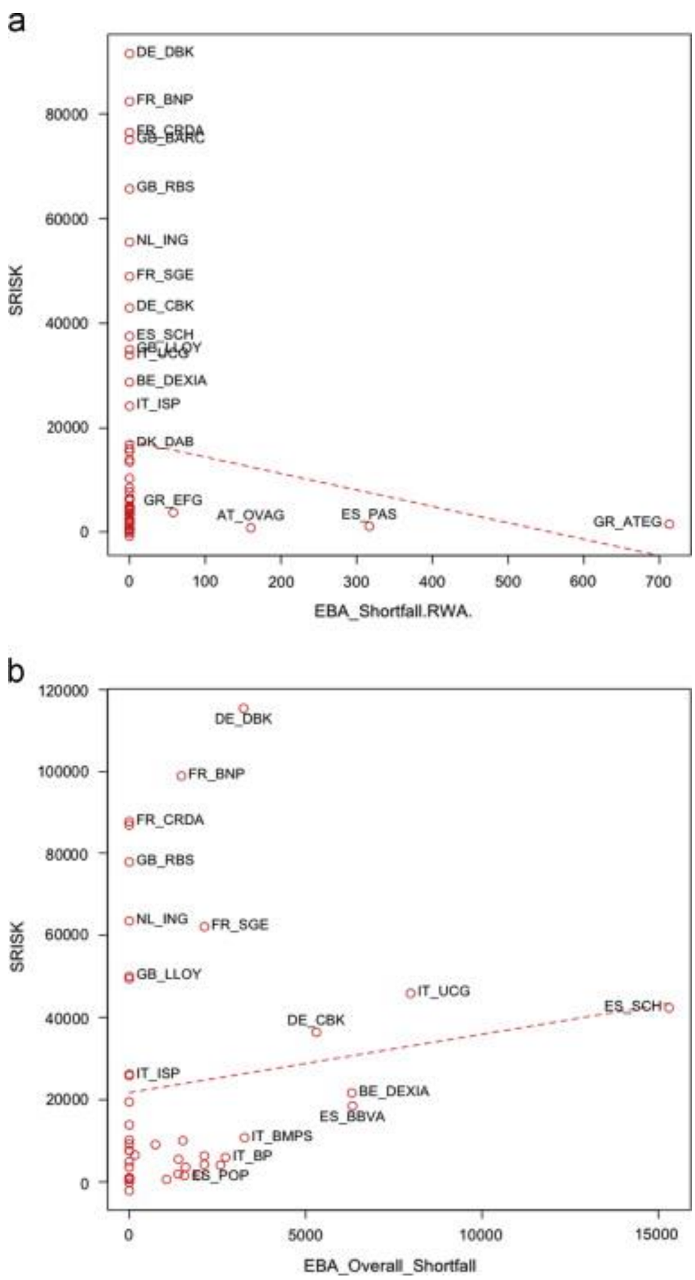
Appendix C

Original EBA capital shortfalls vs. SRISK

This appendix presents the original results of Acharya, Engle, and Pierret (2014):

(a) Disclosed capital shortfall based on 2011 EBA stress test data disclosed in July 2011 (Eq. (1)) vs. SRISK based on V-Lab data with download date: 31.12.2010.

(b) EBA overall shortfall based on 2011 EBA Capital Exercise data disclosed in December 2011 (Eq. (4)) vs. SRISK based on V-Lab data with download date 30.09.2011.



Source: Figure taken from Acharya, Engle, and Pierret (2014, 50)

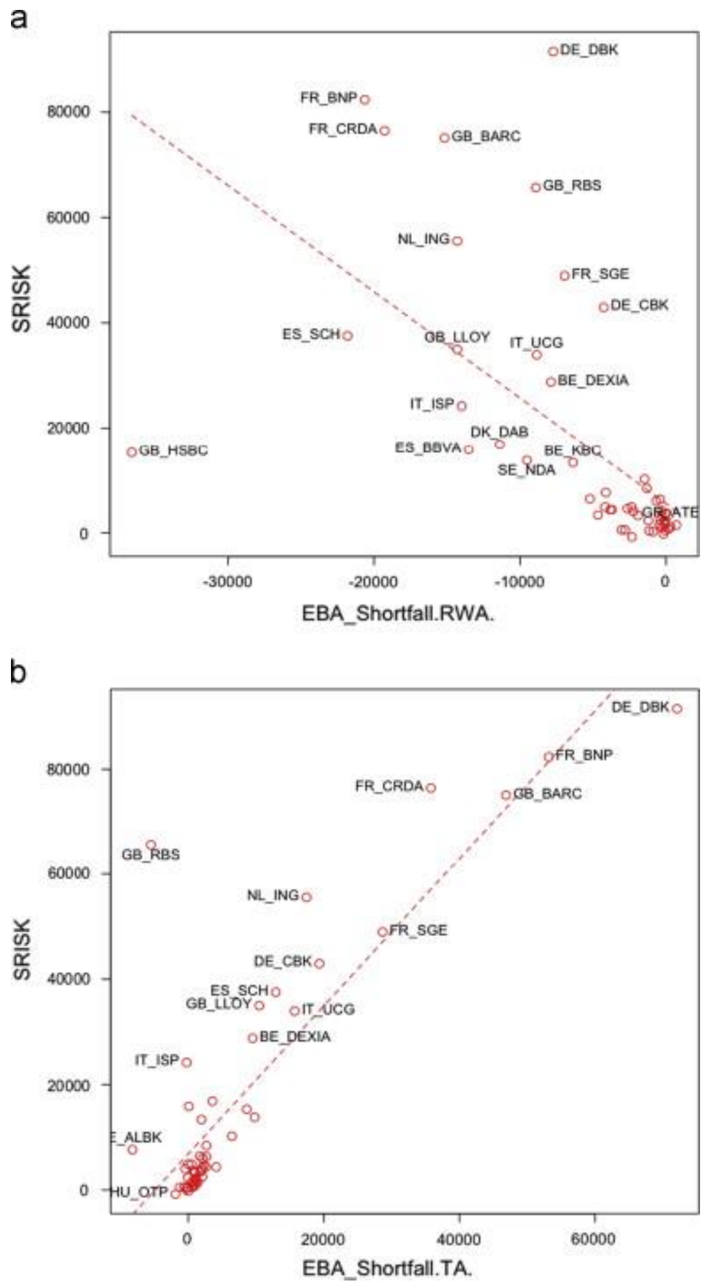
Appendix D

Original EBA risk-based and leverage-based capital shortfalls vs. SRISK

This appendix presents the original results of Acharya, Engle, and Pierret (2014):

(a) Absolute capital shortfall based on 2011 EBA stress test data disclosed in July 2011 (Eq. (2)) vs. SRISK based on V-Lab data with download date: 31.12.2010.

(b) Leverage-based capital shortfall based on 2011 EBA stress test data disclosed in July 2011 (Eq. (3)) vs. SRISK based on V-Lab data with download date 31.12.2010.



Source: Figure taken from Acharya, Engle, and Pierret (2014, 51)

Appendix E

Original realized volatility regressions

This appendix presents the original results of Acharya, Engle, and Pierret (2014):

Cross-sectional regression estimates of six-month realized volatility following the 2011 EBA stress test disclosure in July 2011. Dependent variable: six-month realized volatility (Eq. (6)). Independent variables: book-to-market ratio, V-Lab risk weight (Eq. (7)) with V-Lab download date 30.06.2011, EBA T1LVGR (Eq. (8)) at the end of the EBA stress scenario, EBA risk weight (Eq. (9)) at the end of the EBA stress scenario. White's heteroskedasticity-consistent standard errors are reported in parentheses. Sample size: 53.

		1	2	3	4	5	6
	Constant	4.39**	-0.12	6.34**	5.34**	1.70	0.12
		(0.27)	(1.82)	(0.83)	(0.88)	(1.89)	(1.90)
	Book-to-market	0.03**	0.03**	0.03**	0.03**	0.03**	0.04**
		(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.004)
V-Lab	Risk weight (Eq. (3))		2.50*			2.62**	2.99**
			(0.96)			(0.79)	(0.78)
EBA	*T1LVGR*, scenario end			−39.99*		−41.39*	−62.44*
				(16.82)		(19.02)	(26.39)
EBA	Risk weight, scenario end				−1.75		3.56
					(1.52)		(2.08)
	F-test	11.48**	10.2**	11.88**	6.43**	12.72**	11.25**
	Adj. R^2 (%)	16.78	26.14	29.50	17.28	40.34	44.10

* Statistical significance at the 5% level.
** Statistical significance at the 1% level.

Source: Table taken from Acharya, Engle, and Pierret (2014, 46)

Appendix F

List of adjusted formulas for the extension study

This appendix details the formulas for the extension studies of the 2021 EBA stress test year. $CET1_S$ replaces the previously used $Capital_S$, $TREA_S$ substitutes RWA_S, and actual total assets (TA_B) serve as proxies for the EBA-derived total assets at the end of the stress scenario. The subscript "S" denotes the end of the stress scenario, while "B" represents actual total assets sourced from Bloomberg.

Disclosed capital shortfall:

$$\text{Disclosed capital shortfall} = \max(0, [k' * \text{TREA}_S - \text{CET1}_S]) \tag{10}$$

Absolute capital shortfall:

$$\text{Absolute capital shortfall} = k' * \text{TREA}_S - \text{CET1}_S \tag{11}$$

Leverage-based capital shortfall:

$$\text{Leverage-based capital shortfall} = k * \text{TA}_B - \text{CET1}_S \tag{12}$$

EBA Tier 1 leverage ratio:

$$\text{EBA Tier 1 leverage ratio} = \frac{\text{T1C}_S}{\text{TA}_B} \tag{13}$$

EBA risk weight:

$$\text{EBA risk weight} = \frac{\text{TREA}_S}{\text{TA}_B} \tag{14}$$

Source: Formulas by G. Duckert

Appendix G

List of banks for the 2021 extension study

This list contains all banks, their codes, and data sources underlying the 2021 extension sample.

Bank Code	Bank Name	Stress Test 2021	V-Lab
AT001	Erste Group Bank AG	✓	✓
AT002	Raiffeisen Bank International AG	✓	✓
BE005	KBC Groep NV	✓	✓
DE017	Deutsche Bank AG	✓	✓
DE018	Commerzbank AG	✓	✓
DK008	Danske Bank A/S	✓	✓
DK009	Jyske Bank A/S	✓	✓
ES059	Banco Santander SA	✓	✓
ES060	Banco Bilbao Vizcaya Argentaria SA	✓	✓
ES065	Banco de Sabadell SA	✓	✓
ES069	Bankinter SA	✓	✓
FI904	Nordea Bank Abp	✓	✓
FR013	BNP Paribas SA	✓	✓
FR014	Credit Agricole SA	✓	✓
FR016	Societe Generale SA	✓	✓
HU036	OTP Bank PLC	✓	✓
IE037	AIB Group PLC	✓	✓
IE038	Bank of Ireland Group PLC	✓	✓
IT040	Intesa Sanpaolo SpA	✓	✓
IT041	UniCredit SpA	✓	✓
IT042	Banca Monte dei Paschi di Siena SpA	✓	✓
IT901	Banco BPM SpA	✓	✓
IT903	Mediobanca SpA	✓	✓
NL047	ING Groep NV	✓	✓
NL049	ABN AMRO Group NV	✓	✓
NO051	DNB ASA	✓	✓
PL052	PKO Bank Polski SA	✓	✓
PL902	Bank Pekao SA	✓	✓
PT054	Banco Comercial Portugues, S.A.	✓	✓
SE085	Skandinaviska Enskilda Banken AB	✓	✓
SE086	Svenska Handelsbanken AB	✓	✓
SE087	Swedbank AB	✓	✓

Source: Table by G. Duckert