Financial Risk and Management:
An International Journal
Volume 1, Number 2, 2022

The Effectiveness of Stress Testing Models for Financial Stability

Khansa Rizky Febrianti Pramono[1]

[1] Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

Abstract

Article history:
Received: July 24, 2022
Revised: August 7, 2022
Accepted: September 29, 2022
Published: December 30, 2022

Keywords:
Bank Behavior,
Climate Risk,
Financial Stability,
Stress Testing,
Systemic Risk.

Identifier:
Zera Open
Page: 63-74
https://zeraopen.com/journal/frmij

This study examines the effectiveness of stress testing models in promoting financial stability, addressing the central question of whether current methodologies adequately capture systemic risk in an increasingly complex economic environment. Using a systematic literature review approach, the study synthesizes peer-reviewed evidence published between 2017 and 2021 to identify advances, limitations, and emerging challenges in stress testing frameworks. The results highlight significant sensitivity to modelling assumptions, scenario design constraints, and difficulties in capturing non-linear or climate-related risks, all of which influence the accuracy and interpretability of stress test outcomes. The article discusses these findings by comparing modelling approaches, evaluating empirical insights on bank behavior, and assessing the integration of emerging risk factors. The study concludes that while stress testing has become more sophisticated and influential, its effectiveness depends on continuous refinement and the incorporation of new sources of systemic vulnerability.

1. Introduction

Stress testing has become a central tool in assessing the resilience of financial systems amid heightened uncertainty, rapid digitalization, and increasing market interconnectedness. Initially developed as a regulatory response to systemic failures, stress testing has evolved into a forward-looking analytical framework that enables regulators and financial institutions to evaluate the capacity of banks to withstand severe yet plausible economic shocks. Recent studies highlight that stress testing now influences both supervisory decision-making and internal risk-management practices, underscoring its dual role in promoting stability and shaping strategic choices within the banking sector (Acharya et al., 2018). Given this growing significance, understanding the effectiveness of stress testing models has become a crucial topic within contemporary financial stability research.

Despite considerable methodological advancements, scholars point out that stress testing frameworks still face important limitations. One persistent challenge is model uncertainty, which can significantly affect the reliability of stress test outcomes. Gross and Población (2019) demonstrate that variations in modeling assumptions, such as macroeconomic projections, credit risk sensitivities, and shock transmission mechanisms, can produce widely divergent stress test results. Their findings reveal that the effectiveness of stress testing is contingent not only on the quality of data and scenarios but also on the robustness of the underlying modeling frameworks. This raises important concerns regarding how supervisory authorities interpret stress test results and how banks adjust their strategies in response.

In addition to model uncertainty, the emergence of new and complex risk channels requires stress testing models to evolve beyond traditional macroeconomic scenarios. Among these emerging risks, climate-related shocks have received increasing attention for their potential to destabilize financial systems. Battiston et al. (2021) argue that climate risks introduce non-linear, long-term, and highly uncertain dynamics that conventional stress testing models are not fully equipped to capture. Complementing this perspective, Albano et al. (2021) propose advanced scenario-construction techniques that incorporate climate uncertainties and emphasize the importance of integrating both physical and transition risks. These contributions highlight the growing need for innovative scenario design to ensure stress tests remain forward-looking and relevant.

Beyond methodological aspects, governance and transparency also shape the effectiveness of stress testing frameworks. Goldstein and Leitner (2018) show that the disclosure of stress test results can enhance market discipline by reducing informational asymmetries. However, they caution that excessive transparency may encourage banks to strategically manage their balance sheets to meet supervisory thresholds, potentially undermining the intended risk-reduction effects. This tension illustrates the delicate balance regulators face when designing stress testing regimes.

Despite progress, existing research remains fragmented, with studies often evaluating isolated components of stress testing rather than providing an integrated assessment of effectiveness. A systematic literature review is therefore needed to synthesize recent evidence, identify methodological strengths and weaknesses, and highlight implications for regulators and financial institutions. By examining peer-

reviewed studies published between 2017 and 2021, this article aims to provide a comprehensive evaluation of the effectiveness of stress testing models in promoting financial stability.

2. Literature Review

Recent scholarship on stress testing highlights its central role as a macroprudential tool designed to anticipate vulnerabilities in banking systems. Early supervisory frameworks have gradually evolved into more integrated approaches that incorporate macro-financial linkages, scenario-based analysis, and forward-looking risk assessments. Acharya et al. (2018) show that stress tests influence banks' capital planning and lending behavior, suggesting that these exercises have both supervisory and behavioral implications. Complementary studies also emphasize the importance of systemic risk measurement in informing stress testing design. Brownlees and Engle (2017), for instance, introduce the SRISK framework to quantify capital shortfalls under systemic stress, illustrating how market-based systemic risk measures complement regulatory stress test outputs.

A central area of debate in the literature concerns the limitations and uncertainty embedded in stress testing models. Gross and Población (2019) demonstrate that stress test results are highly sensitive to assumptions regarding credit risk dynamics, macroeconomic conditions, and model structures. This concern is echoed in broader systemic risk literature. Engle (2018) argues that risk models often underestimate the scale of systemic interactions during crises, particularly when relying on historical data that fail to capture non-linear dynamics

and tail events. Furthermore, Gross et al. (2020) show that expected credit loss models used in stress testing can exhibit significant variability depending on the modelling approach, highlighting the challenge of ensuring consistency in top-down credit risk assessments.

Another growing strand of literature examines the integration of emerging risks into stress testing frameworks. Climate-related financial risks, in particular, pose methodological challenges due to their long-term, uncertain, and potentially irreversible nature. Battiston et al. (2021) emphasize the need for stress testing models to incorporate transition and physical climate risks, while Albano et al. (2021) propose advanced scenario-construction techniques to better capture such complexities. Beyond climate risks, Adrian et al. (2019) introduce the "vulnerable growth" concept, underscoring how macro-financial conditions may amplify downside risks and cause non-linear stress amplification, an insight relevant for designing stress scenarios that capture extreme yet plausible dynamics.

Lastly, the literature points to governance and transparency as critical determinants of stress testing effectiveness. Goldstein and Leitner (2018) find that disclosure of stress test outcomes can enhance market discipline, though excessive transparency may incentivize banks to manage their portfolios strategically rather than reduce underlying risk. Taken together, recent studies suggest that while stress testing frameworks have become more sophisticated, their effectiveness ultimately depends on robust modelling practices, appropriate scenario design, and supervisory interpretation. These insights provide an essential conceptual foundation for

assessing the strengths and limitations of stress testing models in promoting financial stability.

3. Methods

The study employs a systematic literature review (SLR) approach to synthesize peer-reviewed research published between 2017 and 2021 on the effectiveness of stress testing models for financial stability. The review process followed a structured search strategy using reputable academic databases, including Scopus, Web of Science, and Google Scholar. Keywords such as "stress testing", "financial stability", "systemic risk", "scenario design", "model uncertainty", and "macroprudential regulation" were combined to identify relevant studies. Inclusion criteria were limited to peer-reviewed journal articles, high-quality institutional publications, and working papers from reputable organizations directly focused on stress testing methodologies, systemic risk assessment, or macroprudential tools. Articles outside the time range, not written in English, or unrelated to the evaluation of stress testing effectiveness were excluded.

The screening process involved three stages: initial identification through keyword searches, abstract review to determine relevance, and full-text evaluation to ensure methodological and thematic alignment. Data extraction focused on key themes including modeling approaches, scenario design, treatment of emerging risks, transparency and governance issues, and identified limitations within stress testing frameworks. This qualitative synthesis approach allowed the analysis to integrate findings from diverse studies while mapping patterns, methodological similarities,

and conceptual gaps across the literature. The overall objective of the methods used was to ensure transparency, replicability, and a comprehensive understanding of the strengths and weaknesses of current stress testing models.

4. Results and Discussion

The synthesis of peer-reviewed studies published between 2017 and 2021 indicates that stress testing remains a central macroprudential tool for evaluating banking system resilience, yet its effectiveness depends heavily on modelling robustness, scenario design, and supervisory interpretation. Acharya et al. (2018) demonstrate that stress tests influence banks' capital strategies and lending behavior, supporting the view that stress testing has both supervisory and behavioral implications. Brownlees and Engle (2017) further show that market-based measures of systemic risk, such as SRISK, can complement supervisory stress test outputs by capturing real-time vulnerabilities during market stress. Together, these studies suggest that while stress testing has improved risk identification, its practical effectiveness hinges on how well models capture systemic interactions and institutional behavior.

A key finding emerging from the literature concerns the sensitivity of stress test outcomes to modelling assumptions. Gross and Población (2019) illustrate that even minor variations in macroeconomic projections or credit risk parameters can generate significantly different stress test results, raising concerns about model reliability. This issue is reinforced by the work of Gross et al. (2020), who show that expected credit loss estimates under IFRS 9 and CECL frameworks vary widely

depending on the modelling approach, particularly in top-down stress testing frameworks. These studies highlight that model uncertainty remains a fundamental limitation: stress tests often rely on assumptions that may not hold during episodes of severe financial disruption, potentially underestimating loss severity or misrepresenting risk correlations. Engle (2018) similarly emphasizes that risk models frequently fail to capture nonlinear dynamics and tail-risk amplification, suggesting that traditional approaches may underestimate systemic spillovers.

Peer-reviewed empirical studies also reveal how stress tests perform under real-world supervisory settings. Gambetta et al. (2019), examining macro stress testing across European banks, find that stress tests significantly influence banks' risk profiles and capital adjustments, providing evidence that the exercise affects institutional behavior beyond compliance. However, they also note substantial differences in how banks respond, reflecting heterogeneity in internal risk management capabilities. Pederzoli and Torricelli (2017) further argue that macroprudential stress tests often struggle to fully capture systemic risk due to their limited integration of network effects and cross-institution contagion channels. Their analysis of the EBA's 2014 exercise highlights persistent gaps between systemic risk measures and actual stress test outputs, underscoring shortcomings in scenario severity and risk transmission modelling.

Another significant stream of research focuses on emerging risks, particularly climate-related financial vulnerabilities. Battiston et al. (2017) show that climate shocks propagate through financial networks in nonlinear ways, generating systemic effects that traditional stress tests fail to capture. Their climate network model

reveals how transition risks can amplify through balance-sheet interconnections, affecting multiple sectors simultaneously. This concern aligns with more recent research by Battiston et al. (2021), who argue that climate risks require fundamentally different modelling techniques, given their long-term, uncertain, and path-dependent nature. Albano et al. (2021) expand on this by proposing improved climate scenario construction frameworks capable of capturing complex feedback loops. These studies collectively highlight that stress test effectiveness now depends on integrating new categories of systemic risks that traditional models were not designed to address.

Finally, the literature also addresses the broader evolution of stress testing frameworks. Kapinos et al. (2018) provide an overview of how stress testing practices have advanced over time, emphasizing the increasing emphasis on macro-micro integration, data granularity, and scenario realism. Adrian et al. (2019) introduce the concept of "vulnerable growth", illustrating how macro-financial environments characterized by high leverage and low volatility can create hidden fragilities that stress tests may fail to detect. Meanwhile, Goldstein and Leitner (2018) highlight a key governance tension: while disclosure improves market discipline, excessive transparency can prompt banks to adjust portfolios strategically to pass the tests, potentially undermining the broader goal of systemic risk reduction.

Overall, the reviewed evidence suggests that stress testing has become more sophisticated and more widely adopted, yet its effectiveness is still constrained by model uncertainty, scenario design limitations, and incomplete integration of emerging risks. Although stress tests remain essential for financial stability

assessment, their reliability depends on continued methodological refinement and improved supervisory frameworks.

5. Conclusion

This systematic literature review shows that stress testing has become an indispensable tool for assessing financial stability, but its effectiveness continues to depend on the robustness of underlying models, the quality of scenario design, and the integration of emerging systemic risks. Evidence across peer-reviewed studies demonstrates that stress tests influence bank behavior, capital planning, and supervisory expectations, reinforcing their importance as both regulatory and risk-management instruments. However, the review also highlights significant limitations arising from model uncertainty, data sensitivity, and the challenge of capturing non-linear interactions within financial systems, which can weaken the reliability or interpretability of stress test outcomes.

A recurring theme in the literature is the need for stress testing frameworks to evolve in response to new categories of systemic risk. Climate-related financial risks, in particular, introduce long-term, uncertain, and network-driven vulnerabilities that traditional stress testing models are not fully equipped to capture. Similarly, empirical findings show that supervisory transparency and disclosure influence how banks adjust their portfolios, which can either strengthen or undermine the overall objective of risk reduction. These insights underscore the importance of continuously improving the methodological foundations of stress

testing, including scenario generation techniques, macro-micro model integration, and approaches for capturing systemic contagion.

Overall, the findings suggest that although stress testing has advanced considerably in recent years, it remains a work in progress. Its effectiveness ultimately relies on the balance between model sophistication and practical usability, the capacity to incorporate evolving risk factors, and the consistent application of supervisory judgment. Continued refinement, informed by empirical research and real-world supervisory experience, will be essential for ensuring that stress testing remains a reliable pillar of financial stability policy.

References

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

Adrian, T., Boyarchenko, N., & Giannone, D. (2019). Vulnerable growth. *American Economic Review, 109*(4), 1263-1289.

Albano, C. M., McCarthy, M. I., Dettinger, M. D., & McAfee, S. A. (2021). Techniques for constructing climate scenarios for stress test applications. *Climatic Change, 164*(3), 33.

Battiston, S., Dafermos, Y., & Monasterolo, I. (2021). Climate risks and financial stability. *Journal of Financial Stability, 54*, 100867.

Battiston, S., Mandel, A., Monasterolo, I., Schütze, F., & Visentin, G. (2017). A climate stress-test of the financial system. *Nature Climate Change, 7*(4), 283-288.

Brownlees, C., & Engle, R. F. (2017). SRISK: A conditional capital shortfall measure of systemic risk. *The Review of Financial Studies, 30*(1), 48-79.

Engle, R. (2018). Systemic risk 10 years later. *Annual Review of Financial Economics, 10*(1), 125-152.

Gambetta, N., García-Benau, M. A., & Zorio-Grima, A. (2019). Stress test impact and bank risk profile: Evidence from macro stress testing in Europe. *International review of economics & finance, 61*, 347-354.

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

Gross, M. M., Laliotis, D., Leika, M., & Lukyantsau, P. (2020). Expected credit loss modeling from a top-down stress testing perspective. IMF Working Paper No. 20/111.

Gross, M., & Población, J. (2019). Implications of model uncertainty for bank stress testing. *Journal of Financial Services Research, 55*(1), 31-58.

Kapinos, P. S., Martin, C., & Mitnik, O. A. (2018). Stress testing banks: whence and whither?. *Journal of Financial Perspectives, 5*(1).

Pederzoli, C., & Torricelli, C. (2017). Systemic risk measures and macroprudential stress tests: an assessment over the 2014 EBA exercise. *Annals of Finance, 13*(3), 237-251.