Scholar Publishing Group
International Journal of Business Management and Economics and Trade
https://doi.org/10.38007/IJBMET.2026.070205
ISSN 2790-0959 Vol. 7, Issue 2:35-50

Evaluation of the Effectiveness of Financial Risk Management Tools Under the New Regulatory Environment

Minxia Liu

State Street Corporation, SSGA Trust Company, Boston, 02114, MA, US

Keywords: Financial Risk Management; CRR3; CRd6; Basel Iii; Stress Testing; Liquidity Coverage Ratio; Expected Shortfall; Model Risk

Abstract: The Regulatory Environment in the Post-2023 Period has Changed the Standards for Assessing Financial Risk Management Tools. Capital and liquidity ratios are no longer sufficient on their own because the revised credit, market, operational, credit valuation adjustment and governance rules now require institutions to demonstrate resilience under combined risks, model limitations and severe but plausible scenarios. Evaluate the effectiveness of capital buffers, value-at-risk and expected shortfall, supervisory stress testing, liquidity coverage tools, hedging and risk-transfer instruments, and model-risk governance in this study. Based on the requirements of CRR3/CRD6 and recent empirical studies, descriptive data from the European Banking Authority's risk dashboards, and the 2023 and 2025 EU-wide stress tests are integrated for this analysis. The five indices of performance are: loss absorption capacity, tail-risk sensitivity, liquidity survival, operational usability and governance strength. The results show that regulatory capital and liquidity tools are still effective first-line defences: EU/EEA liquidity coverage and net stable funding ratios remain significantly above the 100% minimum, and the adverse end-point aggregate CET1 ratio has risen from 10.4% in the 2023 exercise to 12.1% in the 2025 exercise. It is not effective in all circumstances. Static balance-sheet assumptions, model risk, foreign-currency liquidity gaps, regulatory arbitrage, episodic investment in risk-management capacity and weak integration of cyber, climate and counterparty channels reduce the reliability of individual tools. Therefore, the paper puts forward an all-encompassing, evidence-based system that integrates dynamic multi-scenario stress testing, model-risk controls, granular liquidity analysis, transparent hedging limits and board-level accountability. Therefore, under good coordination among all the tools, the new regulatory environment will be relatively stable; otherwise, these individual compliance indicators might be mistaken for real safety indicators.

1. Introduction

In the mid-2020s, financial institutions have increased the reported capital and liquidity since the global financial crisis, but the risk environment for them has also become more complex. A rapid

increase in interest rates raised the duration and valuation mismatch; geopolitical risks added commodity, sovereign and foreign-exchange risks; non-bank financial institutions enhanced participation in market-based finance; and digital dependence increased cyber, third-party, and operational concentration risks. At the same time, policies for climate change and natural disasters have also begun to affect credit risk, collateral value and the business model. Therefore, the goals of financial risk management have changed. The problem is no longer whether the bank can meet the regulatory ratio; instead, it is how well the collection of risk tools can identify, contain and report losses in the event of multiple risk channels occurring simultaneously.

The revised Capital Requirements Regulation, often referred to as CRR3, has introduced some key components of the final Basel III reforms in the EU. Change the measure of credit, market, operating and credit valuation adjustment risk, and introduce an output floor to limit excessive dispersion between internal models and standardised risk-weighted assets [1]. CRD6 has strengthened supervisory powers and governance expectations for sanctions, third-country branch requirements and the treatment of environmental, social and governance risks [2]. Together, these measures are promoting the development of prudential supervision and governance structure. They also change the economics of risk-management instruments because the capital benefit of hedging, securitisation, internal modelling or portfolio restructuring must be evaluated after considering revised risk weights, floors and supervisory overlays.

Therefore, the government has introduced new rules. The previous measures generally have a single indicator, such as Value-at-Risk (VaR), Liquidity Coverage Ratio (LCR), credit derivatives or stress tests. The above method will only find local optima and ignore interactions. A hedge that reduces market VaR may introduce counterparty and liquidity risk; an increase in high-quality liquid assets lowers short-term funding risk but reduces income; a capital buffer absorbs losses but may be difficult to access if management fears a market stigma; and a stress test increases readiness without changing the scenarios covered in its design. Therefore, recent macroprudential studies have shifted to system-level models for stress testing that include feedback loops, non-bank institutions, and real-financial linkages, rather than the previous first-round bank loss assessments [3].

Based on the above, this paper will investigate six categories of financial risk control instruments: regulatory and management capital buffers; VaR and expected shortfall (ES); supervisory and internal stress testing; LCR and net stable funding ratio (NSFR) management; hedging and credit-risk transfer; and model-risk governance. The structure of the evaluation is as follows: Introduction, current-situation analysis, problem identification, solution strategy and conclusion. It has three research problems. First, which tools will continue to offer stable risk reduction in the new regulatory system? Second, where do measurement error, behavioural responses and regulatory interactions reduce their effectiveness? Third, what all-encompassing governance system can enhance both the resilience and operating efficiency?

The paper makes the following contributions. Effectiveness refers to all three conditions simultaneously: it should meet regulatory requirements, be economically feasible and simple to operate. It then uses the same framework for verifiable public data, such as quarterly EU/EEA capital and liquidity indicators and the two consecutive EU-wide stress tests. Finally, it shows an implementation plan that links risk appetite, data lineage, scenario design, model validation, liquidity and collateral management, hedging controls and board accountability. Empirical analysis is intentionally descriptive because the public data do not support a causal identification strategy at the level of individual banks. However, regulatory mapping, published evidence and cross-tool comparison can also be used to provide a practical reference for determining whether the current toolkit is performing as expected.

2. Current Status of Financial Risk Management under the Revised Regulatory Framework

2.1 Regulatory restructuring and changes in incentives

The new environment is a low-risk environment that does not require models for decision-making. Under the previous Basel arrangements, the internal model generated different risk-weighted assets for the same portfolio. The output floor reduces this dispersion by linking model-based requirements to a certain proportion of the standardised calculation. There are still internal models, but now they are no longer used to directly set capital; instead, they support risk differentiation, pricing, limit setting and supervisory dialogue within a bounded capital system. Therefore, the performance of an internal rating or market-risk model should be judged by its forecasting accuracy, decision-making utility and stability, rather than by how much capital reduction it has achieved.

Market-risk management will also be affected by the basic assessment of the trading book. The revised framework has more strictly separated trading and banking books, increased the weight given to expected shortfall, introduced liquidity horizons, and applied additional capital treatment for non-modelable risk factors. Data sufficiency and modelability are thus central governance problems. Although a position may have been hedged recently, if the relevant risk factor does not have frequent and verifiable prices, the position will still be capital-intensive. Generally speaking, the incentive will be positive as it is connected to observable market depth; however, institutions may choose a standardised approach or more liquid but crowded hedges and thus increase basis risk in a crisis.

Operating Risk is treated the same way. The revised standardised measure replaces several previous ways and links a company's size to an operating-risk capital charge; supervisors will still look at loss data, controls and scenario analysis. Therefore, one can no longer manage cyber incidents, outsourced technology failures, fraud and legal risks solely by means of statistical loss distributions. Therefore, the toolset should combine quantitative capital, incident taxonomies, third-party maps, recovery tests and decision rights. CRD6 governance requirements also promote this by increasing the proportion of management-board responsibilities, risk culture and supervisory intervention.

Regulations Changes and Their Main Impacts on Tool Design are shown in Table 1. As shown in the table, regulations are factors that affect the economic development of all areas and are not external constraints. A tool is effective only if it can reduce the targeted exposure without creating an unknown offset risk, is still operable under stress, and generates information that can be challenged by independent control functions.

Table 1. Regulatory Changes and Implications for Financial Risk Management Tools

Regulatory area	Main change	Affected tools	Effectiveness implication
Credit risk and output floor	Revised standardised risk weights, constrained internal models and phased output floor	Internal ratings, collateral, guarantees, securitisation and capital allocation	Capital relief must be tested after floor effects; model accuracy and pricing value become more important than RWA reduction alone.
Market risk	Expected shortfall, liquidity horizons, stricter	VaR/ES, sensitivities, hedging, limit systems	Tail sensitivity improves, but data scarcity, basis

Regulatory area	Main change	Affected tools	Effectiveness implication
	trading-book boundary and non-modellable risk-factor charge	and valuation controls	risk and model complexity can offset the benefit.
Operational risk	Single standardised measurement supported by supervisory review of loss experience and controls	Scenario analysis, insurance, incident management, cyber resilience and outsourcing controls	Capital becomes more comparable, while actual resilience depends on prevention, recovery testing and governance.
CVA and counterparty risk	More risk-sensitive treatment of counterparty spread risk and derivatives exposures	Collateral, netting, central clearing, CVA hedging and counterparty limits	Hedges must be evaluated jointly with funding, wrong-way and concentration risks.
Governance and ESG risk	Stronger supervisory powers, sanctions, transition planning and management-body accountability	Risk appetite, stress testing, model validation, disclosure and strategic planning	Effectiveness requires evidence of use in decisions, not merely documented policies or regulatory submissions.

Table analysis: As shown in the table, the regulatory package change both the numerator and the decision scope of risk management. Capital, model and hedge benefits are becoming more conditional on standardised floors, data quality and governance evidence. The first kind of pattern is to introduce linked controls for individual quantitative indicators. Market-risk reform increases the sensitivity of the tail but raises modelability and data requirements; operational-risk reform improves comparability but does not replace recovery capacity; and the output floor reduces unnecessary RWA variability and weakens the incentive to use internal models solely for capital optimisation. Therefore, the good institution should maintain high-quality analysis even after the standardization of regulatory capital.

2.2 Status of core quantitative tools

VaR is still used widely in daily trading limits, risk reports and back-testing. It determines a loss upper bound that will not be exceeded at a certain confidence level over an extended period. Given a loss variable L and a confidence level α, VaR is defined as follows:

$$VaR_{\alpha}(L) = inf\{\ell \in \mathbb{R} : P(L \leq \ell) \geq \alpha\} \quad (1)$$

L is the portfolio loss, α is the confidence level, and ℓ is a candidate loss threshold in Equation (1). VaR is more convenient for practical use as it can present a portfolio as a single comparable number and support limits, capital allocation and exception testing. A key deficiency of this metric is its inability to quantify the scale of potential losses exceeding the quantile threshold. It may be unstable if the correlation changes, returns are discontinuous, or historical samples do not include the relevant stress period. Under the revised market-risk framework, VaR is still suitable for management purposes but is no longer an all-encompassing indicator of tail risk.

Expected shortfall addresses the above deficiency by estimating the mean loss in the tail beyond VaR. The measure of a continuous loss distribution is given by:

$$ES_{\alpha}(L) = E[L \mid L \geq VaR_{\alpha}(L)] \quad (2)$$

Equation (2) shows that expected shortfall is the expected loss in the case of a large loss, i.e., a loss exceeding VaR. ES is consistent under the standard assumptions, and because it is more sensitive to the shape of the loss tail, it is better suited for capitalising trading-book risk. However, it is not inherently more accurate. Extreme observations are rare; tail dependence cannot be estimated, and liquidity-adjusted losses may differ significantly from mark-to-market changes. Huang and Moreira have shown that the Design Parameters for liquidity regulation can affect the equilibrium bank capital and the interbank market, and even simple ratios have broad macroprudential effects [4].

Stress tests add specific macro-economic, market, credit and operational scenarios to VaR and ES. The purpose is counterfactual analysis; thus, in some cases, the company will determine which portfolios, income paths and capital layers would fail. Calice and Savoia have found that after the financial crisis, stronger capital requirements and stress tests have changed the off-balance-sheet behaviour of US banks; however, they have also pointed out responses in the form of regulatory arbitrage by undercapitalised institutions [5]. This is a good result for evaluation. A tool may reduce the reported exposure and shift the underlying risk to commitments, derivatives, securitisation or entities outside the original measurement scope.

European stress-testing practice has evolved from static, bank-only exercises to a large number of top-down, macro-micro and system-wide models. The European Central Bank's recent methodological papers show that the granular data, model validation, policy-impact analysis and interconnectedness modules have all been expanded [6]. These developments enhance the capacity for contagion and second-round tests, but most public supervisory exercises still have simplifications, such as a fixed balance sheet. Therefore, the internal stress test should also cover model management operations, deposit behaviour, collateral call, market liquidity and the feedback mechanism of capital depletion on credit supply.

The human-capital channel is often overlooked. Schneider, Strahan and Yang have found that supervisory stress testing raised the demand for quantitative risk management skills, but the response focused on test preparation and poor test results rather than wide-ranging, sustained investment in risk capability [7]. Therefore, even a technically advanced framework will perform poorly if the expertise is not continuous, the first-line owner is ambiguous, and the validation team lacks independence. Therefore, the model is also organizational; it should be known, discussed, preserved and related to decisions outside of the annual regulatory cycle.

2.3 Observed capital and liquidity conditions

The EBA Risk Dashboard shows the top EU/EEA institutions for a quarter in a row. By the end of 2024, the weighted-average LCR and NSFR were both significantly above the 100% regulatory requirements, at 163.4% and 127.1%, respectively [8]. The proportion of capital was also relatively large. The above aggregates indicate that the regulatory buffers have effectively increased measurable balance-sheet resilience, but averages may mask specific vulnerabilities in currency, business model, legal entity and depositor type. The bank has a relatively high overall LCR, but it is dollar-illiquid due to short-term wholesale funding or considerable intraday collateral requirements. The construction of the assessment system will be reasonably practical and convenient to operate.

3. Research Problems and Evaluation Framework

3.1 Why effectiveness is difficult to measure

First, accounting resilience, regulatory resilience and economic resilience are not the same. Regulatory capital ratios use the specified definitions for eligible capital and risk-weighted assets. The above definitions enhance comparability but fail to cover losses resulting from poorly modelled risks, temporarily excluded or transferred to counterparties. Economic resilience asks whether the institution can continue critical services, meet cash obligations and avoid disorderly deleveraging after realistic losses. Although the bank meets the lower bound, there is a risk of a bank run, concentrated collateral demand or operational suspension. Therefore, the evaluation cannot determine that a tool is effective simply because the reported value exceeds the threshold.

The second issue is endogeneity. The cause of portfolio adjustment is rule-driven. Higher risk weights may encourage genuine de-risking, but they may also promote hedging with imperfect counterparties, asset sales in less transparent markets, synthetic transfers or balance-sheet structures designed to reduce measured exposure. The output floor reduces one type of model arbitrage but may increase the marginal benefit of other balance-sheet adjustments. Observed improvements in a ratio may therefore reflect retained earnings, portfolio changes, changes in the denominator or regulatory restatement. A reasonable assessment will identify the cause of the increase in a ratio rather than assuming a lower risk.

The third is the model and scenario uncertainty. Risk models are built based on a limited number of past observations, and the high-impact, rare, and atypical occurrences of risk cannot be predicted. Correlation increases under stress; market depth falls; customer behaviour changes; and policy responses alter loss paths. A single severe case may be informative but fails to show the main vulnerability. Model risk is relatively high when multiple tools use the same data, assumptions or vendor infrastructure. For instance, the same default probabilities may feed pricing, limits, expected credit loss, capital and stress tests, creating a false diversification in the reports.

The fourth problem is inconvenience. Capital and liquidity buffers are for the use of stress, but management may avoid drawing on them because investors, rating agencies or supervisors will interpret a lower ratio as a weakness. Hedges are economically feasible but relatively expensive to maintain, difficult to unwind or unavailable promptly when required. Stress test results are relatively late in the decision-making cycle. Therefore, the whole assessment should be timely, understandable, actionable and economically feasible to maintain the control environment.

3.2 Evaluation dimensions, data and method

A normal descriptive Design will be employed. Regulatory documents and recent empirical studies have identified the intended mechanisms of all these tools. Public EBA data offer observable indices of capital, liquidity and stress-resistance. It should not be hoped that only regulation will solve the problem; one needs to ascertain whether the evidence supports the goal of using tools and whether the known defects are serious enough to invalidate the conclusion. Table 2 is the evaluation indicators and related evidence.

Table Analysis: Table 2 divides effectiveness into Outcomes, Sensitivity, Usability and Governance. Therefore, a relatively high regulatory ratio will not be used. Loss absorption and liquidity survival test whether the balance sheet can withstand a shock, and tail sensitivity tests if the tool recognises the relevant shock in the first place. It should also be feasible to operate and govern properly, or an accurate prediction that is too late for use or application will have little value. Finally, Proportionality Criterion will determine whether the increase in complexity is an operational risk or a management cost. Therefore, the framework can be used by large-model-

intensive banks and small institutions that need an easier way.

Table 2. Evaluation Dimensions, Indicators and Decision Criteria

Dimension	Core question	Primary indicators	Decision criterion
Loss absorption	Can the institution remain above binding requirements after severe losses?	CET1 depletion, leverage-ratio depletion, capital headroom and breach count	Higher post-stress ratios and credible, releasable buffers indicate greater effectiveness.
Tail-risk sensitivity	Does the tool respond to nonlinear, concentrated and illiquid losses?	ES/VaR relationship, stress loss distribution, scenario coverage and model exceptions	The tool should increase measured risk when tails thicken or liquidity deteriorates.
Liquidity survival	Can cash obligations and collateral calls be met without disorderly asset sales?	LCR, NSFR, currency-specific ratios, deposit concentration and collateral capacity	Ratios above minimums must be supported by monetisable assets and operational access.
Operational usability	Can results be produced, understood and acted on within the decision horizon?	Data latency, limit escalation, management actions, recovery time and test frequency	A useful tool changes pricing, limits, funding or contingency actions before losses accelerate.
Governance robustness	Are assumptions, models and risk transfers independently challenged?	Validation findings, overrides, model inventory, hedge effectiveness and board reporting	Effectiveness requires transparent ownership, independent challenge and remediation evidence.
Cost and proportionality	Does risk reduction justify capital, liquidity, data and compliance costs?	Implementation cost, earnings drag, model complexity and duplicative reporting	Controls should be risk-sensitive and proportionate without creating blind spots.

A typical case of a regulatory index whose actual effect depends on both calculation and enforcement is the liquidity coverage ratio. It is the amount of high-quality liquid assets divided by the net cash outflows over a 30-day stress period.

$$LCR = \frac{HQLA}{NCO_{30}} \times 100\% \tag{3}$$

HQLA in Equation (3) refers to unencumbered assets expected to retain their value and be monetised under stress, and NCO30 is the total expected net cash outflows over 30 calendar days. A ratio of over 100% indicates full coverage of the regulatory stress horizon. Efficiency is also reflected in the place of assets, central bank membership, settlement facilities, currency convertibility and the credibility of runoff assumptions. Therefore, although the aggregate LCR and NSFR data are strong indicators of buffer construction, they do not prove that all entities can access their buffers under all stress conditions.

4. Evaluation Results and Identified Weaknesses

4.1 Liquidity tools: strong aggregate buffers with composition risk

As shown in Figure 1, the weighted-average LCR and NSFR for EU/EEA banks in Q4 2023 and Q4 2024 are presented below. These values are directly from the EBA Risk Dashboard. The dashed line is the 100% regulatory minimum. During this time, there was continuous repayment of central bank funds, fluctuating interest rates and the introduction of the new prudential package.

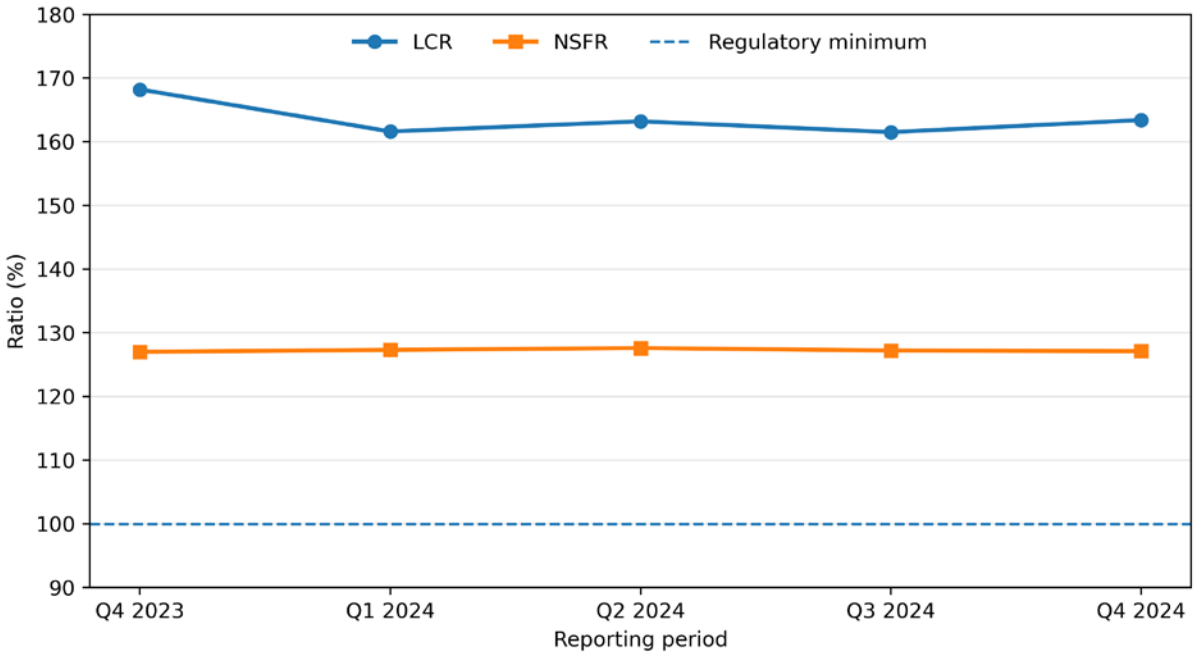

Figure 1. EU/EEA weighted-average liquidity ratios, Q4 2023-Q4 2024

Source: European Banking Authority, Risk Dashboard Q4 2024 [8]. Values: LCR 168.2, 161.6, 163.2, 161.5 and 163.4%; NSFR 127.0, 127.3, 127.6, 127.2 and 127.1%.

Figure analysis: The chart shows that both regulatory liquidity tools were well above the minimum requirements at all times. LCR responds more sensitively to short-term changes in liquid assets and runoff, thus fluctuating more significantly than NSFR; in Q1 2024, it fell by 6.6 percentage points and then increased slightly. The NSFR was still in a narrow 0.6-percentage point range and stable. Together, they form a strong buffer: the bank has enough short-term funds and good long-term funding plans. Aggregate stability does not indicate whether there are currency-specific shortfalls, legal-entity trapping, intraday needs or concentration among uninsured and operational deposits. The above indicators do not include the regulatory ratios.

The evidence supports a relatively high, but conditional, efficacy rating for LCR and NSFR management. The ratios have created a relatively large buffer and reduced the demand for unstable short-term financing. They also gave the upper-level institutions the relevant indices. However, the fixed 30-day horizon of LCR may promote boundary management, and the assumed runoff rate does not replicate digital bank-run behaviour. Therefore, the regulatory calculation of institutions should be supplemented with survival horizons by currency and entity, deposit-behaviour segmentation, intraday liquidity stress, collateral-concentration tests, and reverse stress scenarios that identify the outflow rate at which the contingency funding plan fails.

The Components of HQLA are also required. Sovereign securities are relatively liquid, but if

yields rise, there will be mark-to-market losses and margin calls or hedging expenses. Central-bank reserves are immediately liquid but may be unevenly distributed across subsidiaries. Covered bonds and other Level 2 assets may be less liquid after multiple banks sell simultaneously. Therefore, the real risk-management problem is not how much HQLA is reported, but how quickly it can be monetised after haircuts, operational delays and encumbrance constraints.

4.2 Capital buffers and stress testing: improved resilience, incomplete comparability

Stress-test effectiveness can be judged by the reduction in total CET1 capital. A simple positive depletion measure is set as the starting ratio minus the adverse end-point ratio:

$$D_{CET1} = CET1_{start} - CET1_{end} \tag{4}$$

Equation (4) is the percentage-point reduction in the CET1 ratio over the scenario period. A relatively small depletion indicates stronger earnings, lower losses or a lower risk-weighted asset level, and a high end-point ratio shows more reserves. Given the actual circumstances, the relevant rules and samples of a particular system will vary. Therefore, it is more suitable as one part of a whole evaluation rather than a numerical rank of exercise scores.

The 2023 EU-wide stress test started with a fully-loaded aggregate CET1 ratio of 15.0% and ended at 10.4% in the adverse scenario [9]. The aggregate CET1 ratio under CRR3 in the 2025 exercise started at 15.8% and ended at 12.1% in the adverse scenario [10]. Figure 2 shows these official starting and ending points. Because the exercises have different starting dates, scenarios and regulatory bases, they cannot be used to conduct a controlled experiment on the effect of regulation.

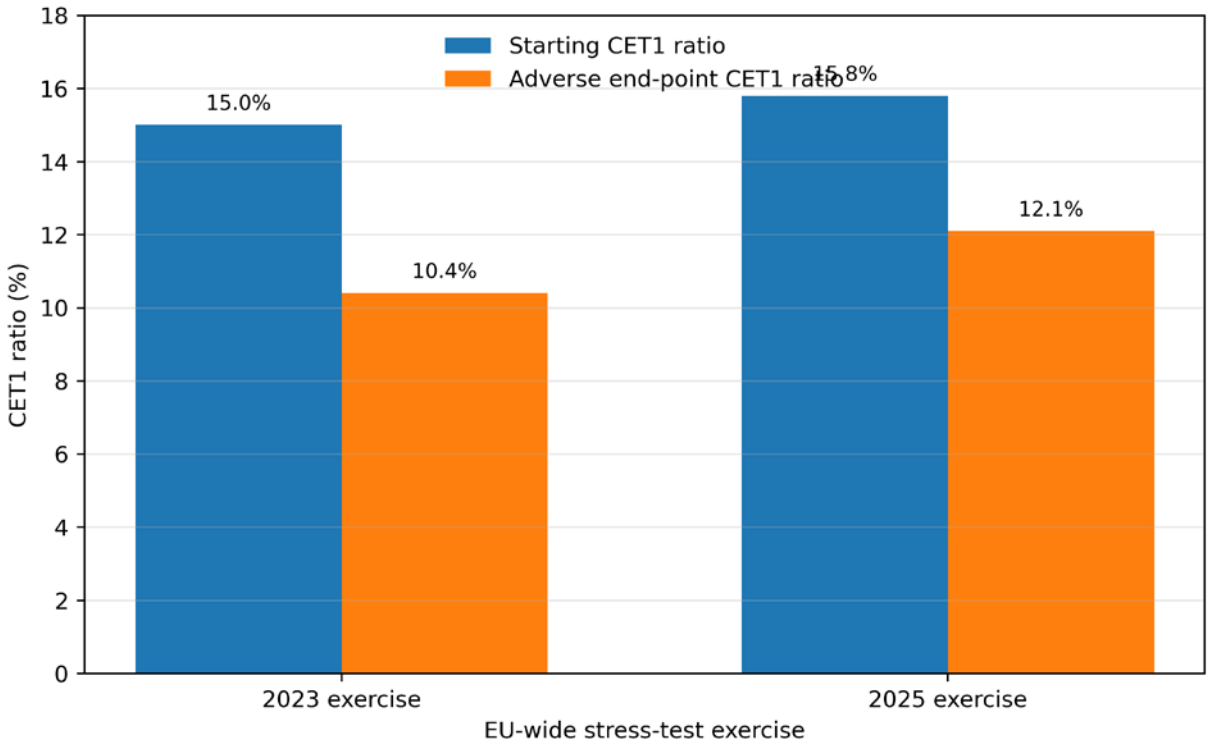

Figure 2. Aggregate CET1 ratios in the 2023 and 2025 EU-wide stress tests

Source: European Banking Authority, 2023 EU-wide stress test results [9] and 2025 EU-wide stress test results [10]. The 2023 data are fully loaded; the 2025 data are transitional and restated under CRR3.

Figure analysis: The 2025 exercise has a higher starting position and a higher adverse end-point CET1 ratio than the 2023 exercise. The corresponding falls in 2023 and 2025 will be 4.6 percentage points and 3.7 percentage points, respectively. This pattern is in line with improved capital resilience and stronger income generation; however, it should not be taken as a direct cause-and-effect relationship because of differences in scenario severity, accounting treatment, sample composition and the CRR3 restatement. Therefore, the main effectiveness conclusion is relatively strong but limited: capital buffers and stress testing together show considerable aggregate loss-absorption capacity, but supervisory and internal analysis still need to examine bank-level dispersion, interim troughs, management actions and risks outside the published scenario [11-13].

The data show that the capital buffer has served as the first line of defence. Banks entered both exercises with substantial common equity and had maintained a positive aggregate capital ratio after significant losses. Stress tests can determine the maximum extent of the institutions' losses from credit, market and income risks and which institutions are most exposed. Provide reasons to improve the quality of data, scenario governance and capital planning. According to the 2025 data, the entire system has sufficient headroom for absorbed losses under normal circumstances and will not be affected by an unexpected shock [14-16].

The Scope of the stress test is limited. A static balance sheet fails to show the bank's growth model, portfolio migration, or other defensive measures, and thus cannot reveal the actual feedback effect. Public exercises may also underrepresent simultaneous deposit runs, cyber outages, market closures and non-bank deleveraging. Therefore, the internal tests should have multiple narratives, probabilistic sensitivity ranges and reverse stress tests. Distinguish between actions that are operationally feasible and those that require liquid markets, rapid asset sales or free-of-charge transfers of funds among institutions [17-19].

Another issue is capital-buffer usability. Management may keep an extra buffer above the requirement because the perceived cost of entering the combined buffer is high. This behaviour increases normal-time resilience but may make buffers effectively non-releasable and thus cause credit contraction under stress. Supervisors and the board should specify the conditions for using buffers and recovery. The aim is not to reduce capital at all times; rather, one needs sufficient capital to absorb losses without causing an avoidable stigma or sudden deleveraging [20-22].

4.3 Cross-tool assessment

Table 3 presents the supporting data for all the main-tool categories. The ratings are qualitative and cover all the items in Table 2. They differentiate a tool's primary protective role from the deficiencies in that role when used alone [23].

Table analysis: Cross-tool comparison shows that a single instrument is not adequate. Capital and liquidity buffers receive the highest ratings because they can absorb losses and cash outflows directly, and the public data show a large buffer. Stress testing also needs to cover these buffers in the future, but only if they are properly implemented throughout the organisation. VaR and ES are still suitable for high-frequency market-risk decisions, but they are susceptible to tail-data and liquidity assumptions. Hedging is the most transaction-specific type because the reduced risk cannot be offset by basis, counterparty or funding risk. Model-risk governance is not an additional line of defence, but the control mechanism that guarantees the reliability of all other tools.

Table 3. Evidence-based Evaluation of Main Financial Risk Management Tools

Tool	Primary contribution	Observed evidence	Main limitation	Assessment
Capital buffers	Absorb unexpected losses and support confidence	Positive aggregate CET1 ratios after severe EBA adverse scenarios	Usability stigma, denominator effects and possible credit-supply contraction	High, conditional on releasability and credible capital planning
VaR and expected shortfall	Measure market-loss distribution and support daily limits	ES improves tail sensitivity relative to VaR and aligns with revised market-risk capital	Sparse tail data, model instability, liquidity horizons and non-modellable factors	Moderate to high for measurement; insufficient as a stand-alone control
Supervisory and internal stress testing	Reveal vulnerabilities under coherent scenarios and test capital plans	Higher adverse end-point CET1 ratio in the 2025 exercise than in 2023	Scenario omission, static balance sheets, model overlap and episodic organisational use	High when multiple scenarios, feedback loops and credible actions are included
LCR and NSFR management	Maintain short-term liquidity and stable structural funding	EU/EEA averages remained far above 100% through Q4 2024	Currency, entity, intraday and depositor-concentration risk can be hidden	High at aggregate level; must be supplemented by granular survival metrics
Hedging and risk transfer	Reduce market, credit or counterparty exposure and stabilise earnings	Can lower sensitivities, losses and capital needs when transfer is effective	Basis, counterparty, funding, legal and regulatory-arbitrage risks	Moderate and highly transaction-specific
Model-risk governance	Controls assumptions, validation, overrides and common dependencies	Recent frameworks broaden validation and system-wide modelling	Skill shortages, fragmented inventories and remediation delays	Foundational; effectiveness depends on independence and sustained investment

4.4 Principal weaknesses under the new regulatory environment

The five weaknesses are as follows. First, measurement perimeters remain fragmented. Banking-book credit risk, trading-book market risk, treasury liquidity, counterparty risk, climate risk and operational resilience are often managed by different systems and committees. This Structure can achieve local optimisation. Treasury hedging will increase interest-rate risk and collateral requirements; synthetic credit transfer can free up capital but introduces concentrated counterparty

risk; cloud migration may reduce operating expenses at the cost of higher dependence on a few providers.

Second, the regulations are becoming more complicated. A high degree of complexity is more prone to errors in implementation, inconsistency in data definitions and a delay in fault rectification. Smaller institutions are too resource-intensive in terms of reporting and lack funds for decision-useful analysis; large institutions have standardised internal and management systems that produce conflicting signals. The output floor simplifies comparability at the aggregate level but adds another level in the transition period.

Third, tail and structural risks are not easily quantifiable. Climate change, cyberattacks, geopolitical sanctions, settlement disruption and non-bank fire sales spread through channels that are not well-represented in historical data. Do not return a false precision value in response. Given the uncertainty, a number of scenarios are used to show expert risks and exposure maps, along with decision thresholds.

Fourth, public ratios are outdated snapshots. They do not show how quickly the liquidity will be depleted, when the collateral will be called upon, the order of management actions or the feasibility of recovery options. Frequent observation and event-driven triggering can be used to link the regulatory measures with the present circumstances.

Fifth, incentives may reduce tool usage. The business unit is unwilling to accept the conservative limit; the model owner wishes to protect the complicated methods; senior management avoids using a buffer due to market signals. Thus, good governance links remuneration, career advancement and responsibility to risk-adjusted performance rather than short-term revenue or minimal compliance.

5. Strategies for Improving Effectiveness

5.1 Establish an integrated effectiveness scorecard

Consider all the tools at the institution level. A practical score card may combine absorption capacity, tail sensitivity, operational convenience, governance level and implementation cost. The effectiveness index of tool j can be shown as:

$$E_j = w_A A_j + w_T T_j + w_U U_j + w_G G_j - w_K K_j, \ \ \Sigma w \ = \ 1 \tag{5}$$

Aj is the loss-absorption capacity; Tj is the sensitivity to tail and interaction risk; Uj is the operational usability; Gj is the governance robustness; Kj is the cost or complexity. Weights w reflect the institution's risk appetite and business model, and sum up to one. The index is not to be substituted for regulatory ratios or model results. It is not to be regarded as good because of one good factor. For example, a high-accuracy model may receive a relatively low overall score if it is too slow for the decision-making process, cannot be independently verified, or exhibits excessive data dependence.

5.2 Integrate data lineage, risk taxonomy and model inventories

The first construction priority is a general risk-data architecture. Trace the source transaction of every material material to the regulatory report, management indicator and stress test result. Common identifiers for counterparties, collateral, legal entities, currencies and products are required. Data lineage should identify transformations, overrides and reconciliations, and material data-quality problems should be linked to risk limits and capital-planning uncertainty. The structure will reduce redundant controls and determine whether different tools are based on the same vulnerability assumption.

The unified model inventory should include regulatory models, management models, vendor

tools, spreadsheets, machine-learning applications and material-calculation engines. Classify according to impact, complexity, data sensitivity and substitutability. Validation frequency can then be risk-based and not uniform. Independently test and benchmark the high-impact model, carry out sensitivity analysis, monitor the results, and set clear conditions for restriction or obsolescence. Aggregate model risk at the level of common dependencies, such as macroeconomic scenarios, probability-of-default estimates, and market-data vendors.

5.3 Replace single-scenario testing with a dynamic scenario library

Internal Stress Tests should use a scenario library, not an annual narrative. The library should include macroeconomic recession, rapid deposit outflow, sovereign spread widening, commercial real estate deterioration, commodity shock, cyber outage, major counterparty default, central clearing margin escalation and non-bank fire sales. Combine scenarios in a way that makes sense. Each scenario should contain the path, not just the destination; otherwise, the time of income and expenditure, liquidity outflow, management actions, etc., will affect life.

Dynamic balance-sheet modelling should supplement the regulatory constant-balance-sheet view. The types of management actions are automatic, pre-approved, conditional or aspirational. Asset sales should have liquidity haircuts and market-capacity constraints. Capital issuance should reflect realistic execution windows. Model deposit pricing and retention according to customer segments. Reverse stress testing should identify a combination of losses and outflows that exceeds the risk appetite or is unrecoverable. Convert the above results into an early-warning threshold and a pre-arranged response.

Scenario Governance Requires Independent Audit. Business units present changes in exposure; the risk function is responsible for methodology and aggregation; finance reconciles earnings and capital impacts; treasury validates funding and collateral assumptions; and internal audit tests the entire process. Boards should present a short comparison of scenarios, vulnerable concentrations, uncertainty ranges and management measures, etc., instead of just a single pass/fail rate.

5.4 Strengthen liquidity, collateral and foreign-currency management

Add operating liquidity indicators to the regulatory LCR and NSFR. Institutions need survival horizons by legal entity and significant currency, intraday payment and settlement profiles, counterbalancing capacity after haircuts, and collateral mobilisation times. The Deposit Model should separate insured retail, uninsured retail, operational corporate, non-operational corporate, financial-institution and digital-platform balances. Adjust behaviour assumptions following market changes and challenge them with extreme but realistic digital outflows.

Contingency Funding Plans should specify actions, responsible parties, preconditions and implementation time. Distinguish between the central bank's capacity and market funding, and identify assets that are legally or operationally trapped. Regular mobilisation tests should confirm that securities can be pledged, settlement instructions work and documentation is current. Monitor the liquidity impact of hedges and derivatives, as well as variation margin, initial margin, downgrade triggers and wrong-way risk at the institution. It is now an operational survival framework rather than an LCR report.

5.5 Control hedging and credit-risk transfer on a total-risk basis

An All-encompassing Risk Indicator for Hedging Selection is Proposed. Before approving the transaction, management needs to quantify how these factors will change after reaching the output floor: market sensitivity, expected shortfall, counterparty exposure, collateral requirements, funding

costs, accounting volatility, legal enforceability and capital requirements. Test the Effect of Hedging Under Normal and Stressed Conditions. Set concentration limits for protection providers and central counterparties, and test the failure of the largest hedge counterparty under a replacement-cost scenario.

Credit-risk transfer can increase the size of the portfolio and lending, but it requires a large-scale and long-term transfer by the institution. Retained tranches, implicit support, early-termination clauses and links to protection sellers can reduce the economic benefits. Strengthen governance by adding transaction-level stress testing, an independent legal review, transparent investor disclosures and periodic comparisons of regulatory capital relief with economic risk reduction. The two are significantly different; thus, the economy will bear the risk and set a price.

5.6 Embed operational, cyber and climate risks in the core framework

Operational Resilience should be linked to financial risk and not be treated as an independent compliance program. Cyber- or third-party failures will result in payment failures, deposit flight, trading losses, collateral defaults and damage to reputation. Critical-business-service mapping should thus connect processes, applications, providers, locations, data and financial exposures. The test scenarios should consider both direct losses and secondary liquidity or market effects. Recovery objectives should be confirmed by practical exercises of provider failure and degraded manual operation.

Include climate and environmental risks in the sector and counterparty analysis, collateral valuation, transition plan assessment and long-term scenarios. Given the high uncertainty, there will not be a single, exact capital amount. Identify concentrations, vulnerable business models, data gaps and strategic options in the institutions. The reduced short-term risk appetite will raise exposure limits and strengthen due diligence; strategically, there will be a portfolio rearrangement, new clients and products, etc. The rules are the same: publicly stated reasons, independent review, and a clear link to the decision.

5.7 Align governance, skills and proportionality

Inform the board of directors and senior management that the risk tool will be used, not only built. Risk appetite statements should link capital, liquidity, concentration, market, counterparty and operational limits to escalation and action. Limit breaches and model overrides should be reported with ageing, ownership and remediation status. Compensation should consider risk-adjusted performance and control failures. The Chief Risk Officer needs sufficient powers, resources and access to the board to question business decisions.

Sustained human capital investment should continue after the regulatory test cycle. Quantitative Modelers, Data Engineers, Treasury Specialists, Cyber experts, Validators and other professionals need to convert the technical results into business decisions for the institution. Rotate between the first-line and independent-control roles for learning, but ensure autonomy in execution. Training should address the model's deficiencies in scenario reasoning and decision-making under uncertainty, not just procedural compliance.

Finally, Proportionality should reduce complexity without changing the results. A small bank can use a standardised model, conservative assumptions and a few scenarios if it has good data control, clear ownership and a credible contingency plan. A large institution has a complicated model and presents increasing accuracy and decision value. The supervisor will establish the data, prevent duplicate uploads, enhance buffer utilisation, and enable open comparison of regulatory and management editions.

6. Conclusion

Evaluate the effectiveness of financial risk management tools in a new regulatory environment based on the framework and recent European data in this study. Based on the above analysis, capital cushions and LCR/NSFR management/stress tests can also provide support. By the end of 2024, EU/EEA liquidity ratios had been well above the minimum requirements, and the aggregate adverse end-point CET1 ratio in the 2025 EU-wide stress test was higher than in the 2023 exercise. According to the above, the bank's capital and liquidity of the banking system have been strengthened.

The evidence does not support the conclusion that compliance indicators are sufficient. VaR and expected shortfall are still subject to tail-data, liquidity and modelability problems. Stress tests can exclude some circumstances and use a simplified balance sheet. Capital and liquidity buffers may be unavailable for use due to stigma, entity restrictions or operational constraints. Hedging and credit-risk transfer shift risk to counterparties, funding and basis channels, etc. Governance of model risk and continuous human expertise ensure the ongoing reliability of the quantitative toolset.

The first policy and management implication is that efficiency will be realised through cooperation. Institutions should link regulatory ratios with dynamic scenario tests, granular liquidity and collateral analysis, total-risk hedge assessment, integrated model inventories, operational-resilience mapping and board-level decision rights. A tool will be considered effective only if it can identify the relevant risks, present timely and understandable information, remain functional in a crisis, and reduce the overall economic vulnerability after considering costs and secondary risks. The new regulatory environment under this norm will be relatively favourable; however, whether it truly benefits the people will depend on how well it is implemented and whether institutions will establish an all-encompassing risk management system or simply treat risk management as a list of compliance checks.

References

[1] *Łuczak P. Credit Risk Management in Commercial Banks[J]. Przegląd Prawno-Ekonomiczny, 2025 (3): 11-21.*

[2] *Van Cleynenbreugel P. Administrative design obligations in European Union secondary legislation: the emerging contours and constitutional cracks of an implicit legislative template[J]. 2026.*

[3] *Aikman, D., Beale, D., Brinley-Codd, A., Covi, G., Hüser, A.-C., & Lepore, C. (2023). Macroprudential Stress-Test Models: A Survey. IMF Working Paper WP/23/173. International Monetary Fund.*

[4] *Han, X. (2026). Research on Standardized Pathways for Automotive Manufacturing Process Quality Systems Oriented Toward Defect Prevention. Advances in Computer and Communication, 7(2).*

[5] *Zhang, Z. (2026). Research on Performance Monitoring and Data-driven Optimization Mechanisms for AI Systems Oriented Toward Decision Support. Advances in Computer and Communication, 7(2).*

[6] *Ma, X. (2026). Research on the Design and Application of Observability Architectures for High-reliability Distributed Systems in Cloud-native Environments. Engineering Advances, 6(2).*

[7] *Wang, Z. (2026). Earnings Quality of Resource Enterprises Under Commodity Price Volatility Transmission.*

[8] *Ma, W. (2026). Reliable Data Infrastructure Supported by Automated Data Pipelines. Engineering Advances, 6(2).*

[9] Jin, C. (2026). Research on the Optimization Strategy of Retail Enterprise Product Portfolio Based on Association Rule Analysis. Advances in Computer and Communication, 7(2).
[10] Wang, Z. (2026). Data-driven Pathways for Optimizing Supply Chain Operational Efficiency in Physical Industries. Advances in Computer and Communication, 7(2).
[11] Han, X. (2026). Research on Adaptive Optimization Methods for Multi-Objective Manufacturing Process Parameters in Complex Assembly Processes. Procedia Computer Science, 279, 366-374.
[12] Wang, N. (2026). Construction and Application of Interpretable Enterprise Performance Prediction Model Based on Multi-Source Operational Data.
[13] Ma, X. (2026). Distributed Fault Root Cause Localization and Self-Healing Strategy Generation Based on Causal Inference. Procedia Computer Science, 281, 753-760.
[14] Ma, W. (2026). Cloud Data Platform Governance Model under Infrastructure Automation.
[15] Wang, Z. (2026). Research on Data Analysis and Quality Prediction of Manufacturing Processes Based on Improved Deep Learning Algorithms. Procedia Computer Science, 281, 1328-1337.
[16] Chi, C. (2026). Research on Mortgage Asset Cash Flow Simulation and Risk Transmission Mechanisms across Structural Tranches Based on Loan-Level Data.
[17] Chi, C. (2026). Quantitative Evaluation of Tranche-Level Returns and Loss Distributions of Structured Credit Assets under Stress Scenarios.
[18] Su, J. (2026). Design and Implementation of Android High Reliability Communication Module Based on Hierarchical Architecture.
[19] Qian, X. (2026). Market Data–Driven Demand Forecasting and Inventory Coordination Mechanisms in Retail Supply Chains.
[20] Yang, Y. (2026). The Integration and Application of Ai Technology in Marketing Data Analysis and Business Decision-Making.
[21] Sun, L. (2026). Resource Scheduling and Elastic Scaling Optimization of Distributed Advertising Systems in Cloud-Native Environments.
[22] Wang, N. (2026). Research on Integrated Analysis Method of Sales and Operations Data for Management Decision Support.
[23] Han, X. (2026). Research on Manufacturing Deviation Vehicle Performance Transmission Mechanism and Intelligent Compensation Control for Automotive Engineering. Procedia Computer Science, 281, 594-602.