



EUROPEAN CENTRAL BANK

EUROSYSTEM

## Working Paper Series

David Aikman, Romain Angotti,  
Katarzyna Budnik

Stress testing with multiple scenarios:  
a tale on tails and reverse stress  
scenarios

No 2941

## **Abstract**

This paper proposes an operational approach to stress testing, allowing one to assess the banking sector's vulnerability in multiple plausible macro-financial scenarios. The approach helps identify macro-financial risk factors of particular relevance for the banking system and individual banks and searches for scenarios that could push them towards their worst outcomes. We demonstrate this concept using a macroprudential stress testing model for the euro area. By doing so, we show how multiple-scenario stress testing can complement single-scenario stress tests, aid in scenario design, and evaluate risks in the banking system. We also show how stress tests and scenarios can be optimized to accommodate different mandates and instruments of supervisory and macroprudential agencies.

**Keywords:** macroprudential stress test, multiple scenarios, reverse stress testing, financial stability, banking sector risks, systemic risks

**JEL Classification:** E37, E58, G21, G28

## Non-technical summary

In the wake of the global financial crisis (GFC), the world saw a surge in regulatory reforms, including introducing stress testing as a vital tool for regulators. However, why is stress testing so important? Firstly, it offers a forward-looking perspective that empowers regulators to act swiftly. By examining the financial health of institutions under challenging economic conditions, regulators gain the ability to demand protective measures in advance. Secondly, it simplifies complex information about institutions with intricate business models, hidden exposures, and complex balance sheets into easy-to-understand metrics. Doing so boosts transparency, aids in shaping effective policies, and enhances regulatory credibility.

Real-life stress tests often rely on complex macro-financial scenarios that aim to reflect risks to the financial sector. However, there is a catch: such scenarios may not always be realistic or severe and include truly most relevant risks at any time. The likelihood of their realisation is often unknown, and their impact may be hard to interpret in the policy context. They can leave aside dangerous possibilities and create an illusion of safety, or they can consider very implausible developments and create a false sense of alarm. These pitfalls have led to scrutiny of many past stress tests from regulators and markets.

Our paper introduces a fresh approach to assessing the banking sector's resilience by considering multiple plausible economic scenarios. We mitigate fallacies of missing possible realities or focusing on implausible ones by exploring all scenarios consistent with past economic shocks. This broad perspective allows us to pinpoint the weakest points of the banking system and individual banks and provides a comprehensive view of bank vulnerabilities.

We illustrate this concept using a semi-structural model for the euro area, covering individual euro area economies and banks and their interactions. This model, commonly used for macroprudential stress testing, allows us to go beyond assessing bank solvency and examine variables crucial to macro-financial supervisors, such as lending to the non-financial private sector.

Multiple scenario distributional stress testing can add relevant information to post-stress test management actions. We provide examples demonstrating how distributional stress testing can yield various at-risk measures, distinguish between institution-specific and system-wide risks, and offer insights into the evolving risks within the banking system.

Furthermore, multiple scenarios permit identifying the most relevant macro-financial risks in a process known as reverse stress testing. Reverse stress testing helps identify risks to the banking system and design scenarios that align with supervisory and macroprudential agency goals. It also helps incorporate economic narratives consistent with policymakers' expectations.

In conclusion, our paper advocates for incorporating multiple-scenario and reverse stress testing approaches alongside conventional single-scenario stress testing. These methods enrich scenario design, validate supervisory exercise results, and shed light on the diverse distribution of risks among individual institutions within the banking system. Notably, advancements in stress test infrastructure make these approaches increasingly feasible.

# 1 Introduction

The aftermath of the global financial crisis (GFC) ushered in a wave of regulatory reforms, including the introduction of stress testing into the regulatory toolkit. Regulatory stress testing offers essential benefits. First, it provides a forward-looking perspective that empowers regulators to take prompt action. By scrutinising the financial solvency of institutions during stylised but challenging economic conditions, regulators can demand adequate protective measures ahead of time. The second advantage of stress testing is its ability to distil complex information about institutions with intricate business models, latent exposures, and cryptic balance sheets into easily understandable metrics (see, e.g., Schuermann [2014]). Accordingly, regulatory stress tests increase transparency within the financial sector and help shape and communicate effective policies.

The widespread adoption of stress tests created the need for severe, plausible, and meaningful scenarios. The main regulatory stress tests, therein the Comprehensive Capital Assessment Reviews led by the Federal Reserve, and the EU-wide stress tests led by the European Banking Authority (EBA),<sup>1</sup> involve single hypothetical adverse macro-financial scenarios. Three-year future scenarios project economic activity, asset prices, and labour markets that mirror financial system risks during scenario design and could lead to substantial losses for financial institutions. However, these scenarios' severity, plausibility, and relevance are often elusive and sensitive to challenge by regulators and markets.

This paper proposes to assess the banking sector's vulnerability by examining multiple plausible stress scenarios. It first develops a stochastic interpretation of a stress test and then shows how multiple scenarios deliver a complete picture of the banking system's resilience and individual banks. The multiple-scenario stress test describes the full probability distribution of bank outcomes, free of the fallacies of overlooking plausible events or raising false alarms. It allows accurate identification of risks, informs about their distribution between banks, and can separate between their systemic and idiosyncratic components. Furthermore, the approach facilitates reverse stress testing, that is, the search for sufficiently severe scenarios to lead to desirably adverse bank outcomes. We illustrate the merits of multiple scenario stress testing by employing a macroprudential stress testing model for the euro area.

Our approach has two components: knowledge of plausible futures and policy preferences. Each future scenario is jointly described by risk factors and the corresponding financial outcomes. The scenario probability space provides complete information on future scenarios and their plausibility. Policy pref-

---

<sup>1</sup>The annual Comprehensive Capital Assessment Reviews have been conducted annually since 2010, and the EU-wide stress tests, biannually since 2011. For a history and comparison of supervisory stress testing in different jurisdictions, see Cihak [2004] for early takeaways, and Borio et al. [2014], Baudino et al. [2018] or Pliszka [2021] for post-GFC assessments.

erences describe the focus of policymakers and its desired severity. Policy preferences can also inform about the fitting scenario narrative, which can benefit the relevance of a stress test.

Breaking stress tests into two components has broad advantages. It distinguishes between the "identification" and impacts of relevant risks and the "choice" of a sufficiently severe scenario. It recognises different challenges of both steps, where for the first, it is to find the best model and data, and for the latter, it is to link stress testing choices and later policy actions optimally. This framework describes well the practice of stress testing, providing a common framework for solvency, liquidity stress tests, or even single-factor sensitivity analysis. Last, it allows us to coherently describe and interrelate distributional, reverse stress testing with systemic and tail risk indicators and issues of scenario selection.

We employ a semi-structural macro-micro Bank Euro Area Stress Test (BEAST) model detailed in Budnik et al. [2023] to derive a complete scenario space. The BEAST represents the individual euro area economies and banks and their two-way interactions. Banks entering the model cover around 70% of the euro area banking sector. As a stress test model, the BEAST emphasises the impact of macro-financial scenarios on credit risk, net interest, net fee and commission income, and, in the second line, market risk and remaining profitability items. As a macroprudential infrastructure, it incorporates banks' behavioural reactions and related amplification mechanisms between the banking system and the real economy, bank solvency, and funding costs.

Relevantly, the BEAST model permits stochastic simulations of macro-financial scenarios. The model's behavioural and risk parameter equations are estimated using a combination of macroeconomic and bank-level data. Information on historical shocks, their correlations, and the uncertainty of model parameters can be explored to build different combinations of future macro-financial outcomes.<sup>2</sup> Each simulated scenario is plausible in at least two senses. It is internally consistent thanks to the disciplined construction of a semi-structural model (Flood and Korenko [2015]), and statistically plausible, as it relies on estimated distributions of shocks and parameters (Studer [1999], Breuer et al. [2009], Breuer and Csiszár [2013], Breuer and Csiszár [2016]).

First, we use model simulations to demonstrate the possibilities offered by distributional stress testing. We look at the distributions of predicted bank outcomes to extract information on evolving tail risks to the banking system and individual banks without taking a position on the type of these risks. We also show that distributional stress testing can help establish individual banks' exposure to systemic and idiosyncratic risks. Finally, we elaborate on how distributional stress testing can be employed to validate

---

<sup>2</sup>See Budnik et al. [2021b] and Budnik et al. [2021d] for the discussion of the application of the model to stochastic simulations in the context of assessing the benefits of resilience-building regulatory policies.

and interpret other stress tests by comparing our results with the 2023 EBA/SSM stress test results.

Next, we show how to solve the reverse stress testing problem elegantly in a multivariate dynamic setup. As postulated by Breuer and Csiszár [2013], we systematically inspect alternative scenarios to identify those that push the banking system or individual banks below the pre-set severity threshold. Sufficiently severe scenarios are then separated from the complete space of plausible realities and form a space reflecting the uncertainty of adverse developments.<sup>3</sup> We then discuss how to tailor adverse scenarios to the ambitions of policymakers. We disentangle scenarios corresponding to a lower percentile of the system-wide solvency distribution, with bank solvency being the variable in the eye of most supervisory agencies. Then, we show how to deliver relevant scenarios for macroprudential agencies, which may wish to put more emphasis on bank lending or the role of amplification mechanisms.

Lastly, we show how to align reverse stress testing with the desire to incorporate an economic narrative consistent with policymakers' expectations. To this end, we inspect the distribution of sufficiently severe scenarios to select the subset closest to the postulated narrative. The selected scenarios reflect narratives commonly associated with hypothetical stress tests. However, in contrast to the latter, their probability can be evaluated.

This paper touches on several streams of literature. The first and broadest is the practice of designing stress testing scenarios (Cihak [2004], Quagliariello [2009]). Our approach blends the probabilistic scenario design with the merits of the hypothetical approach, which casts a "hypothetical adverse situation triggered by the materialisation of risks" (ESRB [2021]).

The second is stress testing based on multiple scenarios and the closely linked literature on distributional forecasts. Both literatures postulate looking at many alternative futures. Distributional forecasting evaluates the entire distribution of plausible outcomes, while multiple scenario stress tests seek regions of the distribution that speak of vulnerabilities in the system or institution (Studer [1999]). Multiple scenario stress testing traditionally has posed two problems: how to arrive at a distribution of plausible multivariate scenarios and how to identify its areas of severity. Regarding the first challenge, Studer [1999] and Breuer et al. [2009] propose to describe the joint distributions of risks applying the Mahalanobis distance<sup>4</sup>, Breuer and Csiszár [2013] and Breuer and Csiszár [2016] use entropy balls<sup>5</sup> and McNeil and

---

<sup>3</sup>The definition of sufficiently severe scenarios as a space, rather than single realisations, links to the notion of mixed scenarios introduced by Breuer and Csiszár [2013].

<sup>4</sup>More plausible scenarios are those whose distance from the multivariate mean is below a certain threshold. Scenarios with a decreasing plausibility level form ellipsoids of scenario spaces with increasing distance from the mean of the multivariate risk distributions. A Mahalanobis ellipsoid would be a good choice for normal or t-student risk distributions.

<sup>5</sup>Entropy balls require the a priori knowledge of the reference distribution of risks. The plausibility of scenarios is then judged by the distance to the reference distributions, with increasing distance marking decreasing probability of realisation.

Smith [2012] employ the concept of half-space trimming<sup>6</sup>. Regarding the definition of severity, Studer [1999], Breuer et al. [2009], Breuer and Csiszár [2013], Breuer and Csiszár [2016] filter scenario distributions looking for those leading to the lowest expected asset payoff and McNeil and Smith [2012] for those offering the minimum net asset value.

Our additions to the discussion of multiple scenario stress tests are two-fold. First, we argue for the use of models with structural elements in the generation of plausible scenarios. Models combining structural and empirical identification permit the generation of complex but internally consistent scenarios, with scenario spaces becoming more robust to missing events not present in the estimation sample. They can also accommodate different distributions of risk factors and scenario variables, as well as the uncertainty of the model. Second, we suspend the distinction between risk factors and financial outcomes while assessing the plausibility of a scenario.<sup>7</sup> It allows for a more straightforward description of scenario severity, also in the presence of non-linearities or amplification mechanisms.

Third, our work connects to macro-financial at-risk measures and tail-based measures of systemic risks. Following the Adrian et al. [2019] growth-at-risk article, there has been renewed interest in applying the value-at-risk analysis to financial stability problems. For example, Cont et al. [2020] introduces liquidity-at-risk for joint solvency-liquidity stress testing, and Budnik et al. [2022] uses bank lending-at-risk to track the evolution of the macroprudential stance. Huang et al. [2009], Adrian and Brunnermeier [2016], Acharya et al. [2017] construct systemic risk indices by looking at tail risks for future bank solvency. Our approach shares with this literature its forward-looking orientation and permits the derivation of various at-risk measures and indices.

Fourth, we add to the literature on reverse stress testing. This literature deals with the complexity of reverse stress testing with multivariate scenarios by offering different approaches to identify candidate scenarios. Henry [2021], Glasserman et al. [2015], Flood and Korenko [2015] look only at a subset of plausible scenarios. Henry [2021] operates with a grid of scenarios derived by arbitrary scaling economically consistent risk factors, Glasserman et al. [2015] inspects risk factors in the vicinity of large historical losses, Flood and Korenko [2015] randomly checks scenarios selected from a mesh of the distribution of (elliptically distributed) risk factors. Kopeliovich et al. [2015], Kapinos and Mitnik [2016], Grundke and Pliszka [2018], Traccucci et al. [2019] reduce the dimensionality of the risk factor space by applying the principal component analysis (PCA). Our approach is closest to McNeil and Smith [2012]

---

<sup>6</sup>The half-space trimming to construct relevant scenario sets similarly asks for the a priori knowledge the distributions of risks.

<sup>7</sup>Breuer and Summer [2018] start to jointly assess risk factors and financial outcomes and label such scenarios as generalised. However, in contrast to us, they still evaluate the probability of generalised scenarios based solely on the distribution of risk factors.



and Breuer and Summer [2018], who systematically search for relevant scenarios in the full distribution of outcomes. However, we combine their idea with that of Budnik et al. [2021c] and Sarychev [2014] and show how to select scenarios that are plausible and severe and incorporate a postulated narrative.

This paper is structured as follows. Section 2 introduces the main concepts developed in the paper. Section 3 provides a high-level overview of the macroprudential stress test model. Section 4 presents the simulation setup. Section 5 presents the results corresponding to the full distribution-based stress test for the solvency of the general banking system. Section 7 introduces the reverse stress testing for system-wide bank solvency. Section 8 looks beyond system-wide bank solvency. Finally, Section 9 concludes the paper.

## 2 Exposition

This section introduces multiple scenario stress testing along with its different variants. It nests the types of stress tests illustrated in the following parts of the paper in a common framework and facilitates linking our contributions to the earlier literature. However, practice-oriented readers can skip this section and move directly to the model-based examples.

The section starts with an exposition of a standard real-life stress test and gradually expands to the most general exposition of a stress test. Next, it builds on this description to establish the definitions of a distributional stress test, stress test severity and plausibility, and a reverse stress test. Last, it explains the principles of selecting scenarios with a postulated scenario narrative and when multiple variables are in the focus.

### 2.1 Basic notation

Each stress test can be described as a projection of variables of interest while emphasising the realisation of risks. A stress test employs information  $\mathcal{X}$  available at the time  $t$  of carrying out the exercise. The information set can include current and historical information on the balance sheets of banks, information on current and past macro-financial conditions, and regulatory policies. This chapter assumes that a stress test covers a horizon of  $H$  periods.

The realisation of risks commonly takes the form of future sequences of risk factors. Without loss of generality, we assume that there are  $R$  risk factors and the vector  $s_{t+h} = (s_{1,t+h}, \dots, s_{R,t+h})^\top$  describes their realization in the  $h - th$  period of the stress test horizon. The vector  $S = (s_t, \dots, s_{t+H})$  summarises the evolution of risk factors on the full stress test horizon. It is also convenient to denote any set of such

risk sequences as  $\mathcal{S}$ .

The results of a stress test include economic and bank-level variables. We assume that there are  $N$  projected variables in a stress test, with their realisation in the period  $t+h$  contained in the vector  $y_{t+h} = (y_{1,t+h}, \dots, y_{N,t+h})^\top$ . The single outcome of a stress test, which we will refer to as a single scenario, is  $Y = (y_t, \dots, y_{t+H})$ . A batch of feasible stress test results  $Y$  should be indicated by  $\mathcal{Y}$ .

Finally,  $M(\cdot)$  will be a single mapping function, and  $\mathcal{M}(\cdot)$  any family of mapping functions, such that:

$$M : (\mathcal{X}, \mathcal{S}) \rightarrow \mathcal{Y} \quad (1)$$

Such mapping functions should include, at minimum, models projecting risk factors into bank balance sheets, accounting principles, and regulatory rules.

A policymaker is usually interested only in a subset of variables in a stress test. We acknowledge this fact by introducing function  $C$ , which selects  $K \geq 1$  variables and horizons of interest from the full set of stress test results:

$$C : Y \rightarrow \mathbb{R}^K \quad (2)$$

Furthermore,  $C_k(\cdot)$  denotes the  $k$ -th variable selected by  $C(\cdot)$ .  $c^* = (c_1^*, \dots, c_K^*)^\top$  should be a vector  $K \times 1$  of threshold levels for the variables selected by  $C(\cdot)$ .<sup>8</sup>

Describing the relative severity of stress test scenarios involves the function  $\mathcal{C}$  and the corresponding linear order  $\leq_{\mathcal{C}}$ . The stress test results  $Y^* \in \mathcal{Y}$  should be judged equally or less severe than  $Y^{**} \in \mathcal{Y}$  as long as:

$$C(Y^*) \leq_{\mathcal{C}} C(Y^{**}) \iff \mathcal{C}(Y^*) \leq \mathcal{C}(Y^{**}) \quad (3)$$

where we assume the convention that the least advantageous outcomes are first sorted.<sup>9</sup>

<sup>8</sup>For variables  $k$  for which no thresholds are applied,  $c_k = -/+ \infty$ .

<sup>9</sup>For instance, lower CET1 ratios proceed higher CET1 ratios. The sorting can always be ensured by multiplying a variable by  $+/-1$ .

## 2.2 Traditional supervisory stress test

The primary real-life benchmark for our analysis is the EU-wide EBA/SSM stress test.<sup>10</sup> The stress test follows a bottom-up approach that requires each participating bank to estimate the impact of macro-financial scenarios on a unified set of risks (credit, market, counterparty and operational risk) on their balance sheets. In the process, bank risk and solvency estimates are challenged by supervisory authorities in order to ascertain their sufficient level of stringency and comparability across the participating institutions.

The building blocks of the EU-wide stress test scenario design were set up in 2010 and have remained stable since then. The baseline macro-financial scenario is sourced from the most recent projections of the Eurosystem. The adverse macro-financial scenario is a hypothetical scenario with a narrative anchored on the financial stability risks identified by the General Board of the European Systemic Risk Board (ESRB). The narrative is translated into shocks to macroeconomic aggregates, such as confidence shocks to consumption and investment, and financial variables, such as shocks to bond yields and foreign exchange rates. These shocks are then fed into a multicountry model and translated into the final adverse macro-financial scenario for EU countries (EBA [2021]).

The process of refinement of the adverse macro-financial scenario involves expert insights and selective cross-country comparisons. There is no systematic assessment of the joint likelihood of the realisation of risks in the narrative. At the end of the refinement process, the overall adversity of the macro-financial outlook often decreases while becoming more evenly distributed among countries.

The relatively standard framework of the EBA/SSM stress test can be well represented by:

$$M(\mathcal{X}, \{S^0, S^1\}) = \{Y^0, Y^1\} \quad (4)$$

The information that feeds into the EBA/SSM stress test  $\mathcal{X}$  is the financial accounts of the banks, together with any historical data available to the banks and supervisors. The two sets of risk factors take the form of the baseline  $S^0$  and the adverse  $S^1$  macro-financial scenarios. Each of these consists of the projected values of GDP, house prices, interest rates, and other macro-financial variables and has a size of  $R \times H$ . Finally, the function  $M$  mapping risk factors in the evolution of bank balance sheets involves

---

<sup>10</sup>The stress test is run jointly by the European Banking Authority (EBA) and the ECB Banking Supervision (SSM). The EBA develops the overall methodology, designs templates and focuses on the most significant EU banks. The SSM stress test uses the EBA methodology and applies it to all banks under direct supervision, including those not otherwise covered by the EBA. For more information, visit <https://www.eba.europa.eu/risk-analysis-and-data/eu-wide-stress-testing>.

methodological assumptions, banks' models, and supervisory scrutiny. The methodological assumption involves, among others, constant balance sheet and various caps and floors.

The stress test concludes with two sets of results,  $X^0$  and  $X^1$ , corresponding to the baseline and the adverse stress test scenario, respectively. Each scenario has a size  $N \times H$ . The focus variable is the CET1 ratios at the system and bank-level in the adverse scenario and the end of the scenario horizon. Consequently, the selection function  $C(\cdot)$  extracts a simple vector of the CET1 ratios at  $t + H$ . Finally, the ultimate measure of scenario severity becomes evident only with the availability of the outcome of the EBA/SSM stress test.

### 2.3 Macroprudential stress test

The ECB macroprudential stress test looks at the effects of the same baseline and adverse macro-financial scenarios as in the EBA/SSM exercise under the alternative set of methodological assumptions.<sup>11</sup> The BEAST model replaces banks' own calculations and supervisory scrutiny and the stress test relaxes the assumption of the constant balance sheet in order to study dynamic adjustments of banks and amplification mechanisms therein the feedback loop between the real economy and the banking sector. It also removes several other caps and floors in the supervisory exercise, such as zero write-offs and recovery rates.

Finally, the macroprudential stress test also recognises the uncertainty of the model parameters. The known, empirically informed distributions of model parameters  $\mathcal{P}$  (see Section 3 and Appendix A) allow stochastic simulations under varying parametric assumptions.

The corollary of equation (1) for the ECB macroprudential stress test is as follows:

$$\mathcal{M} : (\mathcal{X}, \{S^0, S^1\}) = \{\mathcal{Y}^0, \mathcal{Y}^1\} \quad (5)$$

where  $\mathcal{M}$  condenses numerous mapping functions depending on the realisations of  $\mathcal{P}$ , and each of  $\mathcal{Y}^0$  and  $\mathcal{Y}^1$  is now a scenario space rather than a single realisation.

The communication of the macroprudential stress test involves a richer set of variables than the EBA/SSM stress test. In addition to the system-wide CET1 ratio and the distribution of solvency rates

---

<sup>11</sup>There is a slight difference in the specification of baseline and adverse macro-financial scenarios in the EBA/SSM and macroprudential stress test. In the latter exercise, they consist of exogenous i.i.d. shocks and take into account future (and already known) policy paths. The shocks replicate the path of macro-financial variables in the EBA/SSM stress test when there are no amplification mechanisms in the stress test (see also Chapter 8 in Budnik et al. [2023]).

between banks,  $C(\cdot)$  covers the reduction in own funds, the evolution of lending to the non-financial private sector, and the strength of the amplification mechanism expressed in terms of GDP. In contrast to the supervisory exercise, the emphasis is almost equally on the baseline and adverse scenarios. There is more consideration of the dynamics of variables over the entire horizon (not only at its end), and the outcome variables are reported with uncertainty ranges.

## 2.4 Multiple scenarios and distributional stress test

Following the two real-life examples of stress tests, we now introduce the general multiple-scenario stress test. Such a stress test considers the probability of an entire space of plausible scenarios  $\mathcal{Y}$ . This probability space  $(\Omega, \mathcal{F}, P)$  is described by  $\Omega$  as a sample space,  $\mathcal{F}$  as a  $\sigma$ -algebra of events, and  $P$  as a probability in  $\mathcal{F}$ .  $\omega \in \Omega$  is a basic event and  $Y = Y(\omega)$  a single set of stress test results.

In our application, the plausible scenario space  $\mathcal{Y} = \mathcal{M}(\mathcal{X}, \mathcal{S})$  is derived by repeating the drawing of the shocks  $\mathcal{S}$  and the values of model parameters  $\mathcal{P}$  from their estimated distributions. Put differently, a basic event  $\omega \in \Omega$  is just a pair of a  $H$  long sequence of  $N$  shocks describing scenario  $S$  and a single parameterisation of a mapping function  $M$ .  $\mathcal{M}$  acknowledges the theoretical interdependencies between macro-financial and bank-level variables and accommodates possibly incorrect specifications of any single mapping function. Such scenarios are plausible in at least two senses. They are statistically plausible by relying on past data. They are also economically plausible, as they preserve internal consistency and meet reasonability criteria coded in the structure of a semi-structural model.<sup>12</sup>

$\phi$  is a density function, and  $\Phi$  is the cumulative distribution function described in the scenario space  $\mathcal{Y}$ :

$$\int_a^b \phi_{\mathcal{Y}}(Y) dY = P(a \leq Y \leq b)$$

$$\Phi_{\mathcal{Y}}(Y) = \int_{-\infty}^Y \phi(\mu) d\mu \tag{6}$$

It is worth noticing that the scenario probability distribution is described in the result space  $\mathcal{Y}$  and not in the shock space  $\mathcal{S}$ . It has several advantages that we explore further. First, it can accommodate irregular shock-generating processes and flexibly combine them with model uncertainty. Second, it facilitates exploring scenario space in the presence of non-linearities in  $\mathcal{M}$ . Such non-linearities in the

<sup>12</sup>The model structure ingrains long-term stylised facts about the behavior of macrofinancial variables, ensures dynamic homogeneity of all relevant macroeconomic and bank-level variables, and theory-based relationships between variables.

regulatory framework (for instance, in risk weight formulas) and amplification mechanisms are likely to emerge in most stress tests.<sup>13</sup> Last, it permits the introduction of an unequivocal definition of scenario severity (Section 2.1) that fits a wide range of models and stress tests.

The probability space of  $C(\mathcal{Y})$  can be measured using a marginal density function (or the corresponding cumulative probability function) together with the following:

$$\phi_{C(\mathcal{Y})}(C(Y)) = \int_{\mathcal{Y} \setminus C(\mathcal{Y})} \phi_{\mathcal{Y}}(Y) d\mu(Y \setminus C(Y)) \quad (7)$$

The most straightforward application of multiple-scenario stress testing is the distributional stress test. The distributional stress test inspects the  $K$ -variate distribution of  $C(\mathcal{Y})$ , often emphasising the lower tails of variables. Most often,  $C(\mathcal{Y})$  is analysed by looking at the marginal distributions of individual variables. However, outcome variables can also be beneficially analysed jointly, as illustrated in Section 5.

## 2.5 Reverse stress test

Reverse stress testing seeks to find plausible scenarios that put the system under sufficient stress. After Cihak [2004], we consider two approaches to reverse stress testing: one based on threshold values of outcome variables and one considering the plausibility of an adverse event.

Reverse stress testing based on thresholds seeks for scenarios defined as:

$$\mathcal{Y}^* : \mathcal{E}(\mathcal{Y}) < c^* \quad (8)$$

The condition gives a result space with the probability of realisation described by the conditional probability density function:

$$\phi_{\mathcal{Y} | (\mathcal{E}(\mathcal{Y}) < c^*)}(Y | \mathcal{E}(Y) < c^*) = \int_{-\infty}^{c^*} \int_{\mathcal{Y} \setminus C(\mathcal{Y})} \phi_{\mathcal{Y}}(Y) d\mu(\mathcal{E}(Y)) d\mu(Y \setminus C(Y)) \quad (9)$$

The worst-case scenario approach sets the level of plausibility of stress scenarios  $p^*$ . We can for-

---

<sup>13</sup>The challenges resulting from the craggy pass-through of risk factors to outcome variables have been identified already by Glasserman et al. [2015], who address them by applying nonparametric estimates of the outcome variables.

mulate it analogously to the threshold scenario approach by identifying the infimum of a scenario subset described by the scenario selection criteria  $\mathcal{C}$ :

$$c^* = \inf\{\mathcal{C}(\mathcal{Y}) : \Phi_{\mathcal{C}(\mathcal{Y})}^{-1}(p^*)\} \quad (10)$$

## 2.6 Multi-variate severity criteria

A practical challenge for reverse stress testing is to deal with multivariate severity criteria. The most straightforward way to interpret multiple criteria would be to apply them sequentially, variable by variable. In such a case, the order  $\leq_{\mathcal{C}}$  would correspond to the Leontieff preference (or loss) function of the policymaker with equal weights. It would give the following corollary of equation (11):

$$\mathcal{Y}^* : C_1(\mathcal{Y}) < c_1^*, \dots, C_K(\mathcal{Y}) < c_K^* \quad (11)$$

However, as noted in Breuer et al. [2009], the sequential application of criteria yields the artefact of "dimensional dependence". The plausibility of selected scenarios negatively depends on the number of criteria  $K$ .

The alternative proposed by Breuer et al. [2009] is to interpret  $\mathcal{C}(\mathcal{Y})$  as a Mahalanobis distance of  $Y \in \mathcal{Y}$  from the postulated and known distribution  $\Gamma$ :

$$\mathcal{C}^\Gamma : \mathcal{Y} | \Gamma \rightarrow \mathbb{R} \quad (12)$$

Distance from the distribution  $\Gamma$  offers severity metrics independent of the number of criteria. However, there are at least two weaknesses to this approach. First, it asks for knowledge of  $\Gamma$ , which can be increasingly difficult to specify for the increasing number of variables and in cases when they follow atypical, e.g., bimodal, distributions (Flood and Korenko [2015]). Second, while  $\Gamma$  can express the statistical or theoretical relationship between different variables of interest, it must not map policy preferences about their relative importance.

Our solution is to specify  $C(\cdot)$  as a linear combination of individual criteria with weights  $w$ . Such  $C^w(\cdot)$  will map all scenarios in  $\mathbb{R}$ :

$$\mathcal{C}^w : \mathcal{Y} \rightarrow \mathbb{R} \quad (13)$$

The weighting approach resembles the proposal of Kopeliovich et al. [2015] or Grundke and Pliszka [2018], with the difference that  $w$  is set by the policy maker rather than selected by the Principal Component Analysis (PCA). An attractive feature of policy weights is that their application does not require approximately normally distributed stress test variables. It is essential, provided that modelling of systemic events can lead to asymmetric or thick-tailed outcome distributions.

Then, the application of  $C^w(\cdot)$  involves a multidimensional nonparametric sorting rank algorithm of Sarychev [2014]. It can be summarised as follows:

$$\begin{aligned} i : C_k(\mathcal{Y}) &\rightarrow \mathbb{N} \\ i(C_k(\mathcal{Y})) &= \text{rank}((C_k(\mathcal{Y}), \leq)) \\ \mathcal{C}^w(\mathcal{Y}) &= \sum_{k=1}^K w_k i(C_k(\mathcal{Y})) \end{aligned} \quad (14)$$

Stress test outcomes are first sorted separately for each variable  $k$  to create an ordered set  $(C_k(\mathcal{Y}), \leq)$ . The rank of each element of  $(C_k(\mathcal{Y}), \leq)$  is then aggregated into the joint criteria with weight  $w_k$ .

## 2.7 Incorporating economic narrative

Occasionally, policymakers aim to create a scenario that exemplifies a particular economic narrative. The reverse stress test reduces the elements of the economic narrative necessary to ensure a sufficient severity of the scenario. However, such a space of plausible and severe scenarios can still nest many alternative realities, with some being more relevant than others, e.g. in stress test communication.

To this end, we adapt the method to select plausible scenarios with a narrative postulated by Budnik et al. [2021c]. We should comb through the scenario space to identify futures that emphasise risks and vulnerabilities in the narrative. The first step is translating the desired narrative into criteria for evaluating single scenarios. Let  $\mathcal{N}$  be a narrative function of the same type as  $\mathcal{C}$  and described in any scenario space  $\mathcal{Y}$ . The function selects a subset of scenario variables that express the criteria and weights them with  $v$ :



$$\mathcal{N}^v : \mathcal{Y} \rightarrow \mathbb{R} \quad (15)$$

The criteria ingrained in  $\mathcal{N}$  can, for instance, map the presence of credit crunch in a scenario by assigning a non-zero weight to a cumulative change in euro area lending in the medium-term horizon<sup>14</sup>.

One can define the cumulative probability function  $F$  described in  $\mathcal{N}(\mathcal{Y})$  as follows:

$$\int_a^b f_{\mathcal{N}(\mathcal{Y})}(\mathcal{N}(Y)) d\mathcal{N}(Y) = P(a \leq \mathcal{N}(Y) \leq b) \quad (16)$$

$$F_{\mathcal{N}(\mathcal{Y})}(\mathcal{N}(Y)) = \int_{-\infty}^{\mathcal{N}(Y)} f(\mu) d\mu$$

The selection of scenarios along with the equation (11) can be combined by setting the desired severity threshold  $p^* = p^C \times p^N$  where  $p^C$  steers only the severity of the scenario and  $p^N$  governs the choice of the narrative:

$$\mathcal{Y}^* : \mathcal{N}(\mathcal{C}(\mathcal{Y})) < \inf\{F_{\Phi_{\mathcal{C}(\mathcal{Y})}^{-1}(p^{\mathcal{C}})}(p^N)\} \quad (17)$$

The scenarios selected based on the severity criteria ( $p^{\mathcal{C}}$  or  $c^*$ ) are later treated as equivalently probable but having different relevance in terms of a narrative.<sup>15</sup>

### 3 Model

The BEAST is a widely utilised financial stability "workhorse model" deployed for risk and policy analysis. Since 2018, the model has built the backbone of the ECB macroprudential stress test (Budnik [2019], Budnik et al. [2019], Budnik et al. [2021a], Budnik et al. [2021a]). The model's main advantage in the current application is that it achieves a balance between maintaining the heterogeneity of individual banks and simultaneously offering an aggregate and temporal perspective.

The model comprises a macroeconomic block with 19 individual euro area economies<sup>16</sup> and their

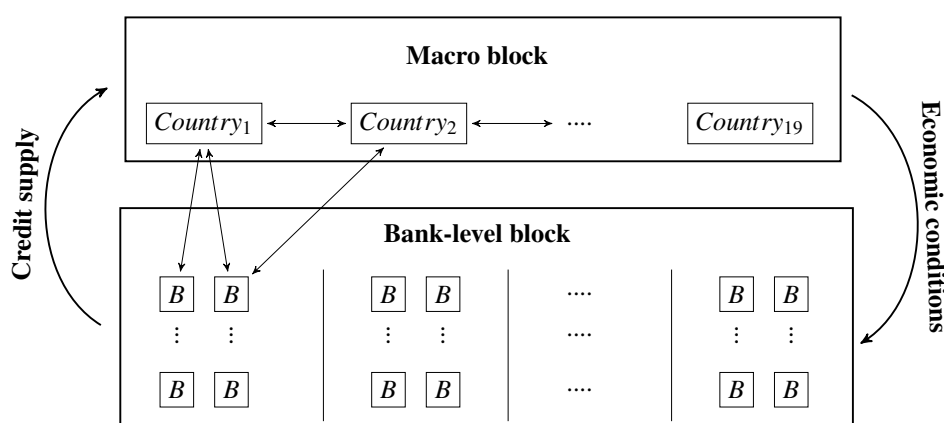
<sup>14</sup>See the discussion of such criteria jointly with examples of application in Budnik et al. [2021c].

<sup>15</sup>The sequence of the application of the two functions on  $\mathcal{Y}$  matters for the result, i.e.  $\mathcal{N}(\mathcal{C}(\mathcal{Y})) = \mathcal{C}(\mathcal{N}(\mathcal{Y}))$ .

<sup>16</sup>As of 2022, the euro area consists of Belgium, Germany, Ireland, Spain, France, Italy, Luxembourg, the Netherlands, Austria, Portugal, Finland, Greece, Slovenia, Cyprus, Malta, Slovakia, Estonia, Latvia, and Lithuania.

18 main non-euro area financial partners.<sup>17</sup> The euro area economies are bound by a common monetary policy and are interconnected by cross-border trade spillovers (see Figure 1).

The banking block consists of 89 of the largest euro-area banks represented on a consolidated level. Economic conditions influence banks in countries with exposure or access to funding. Economic conditions affect bank activity by impacting the quality of bank assets, credit demand, funding costs and availability. Consequently, the model incorporates cross-border financial spillovers commensurate with the international activities of the largest euro-area banks.



Notes:  $Country_1 - Country_{19}$  represent individual euro area economies,  $B$  represents an individual bank. Straight arrows connecting countries highlight the presence of cross-border trade spillovers. Banks headquartered in a country are positioned below the country label. The straight arrows connecting countries and individual banks indicate the two-way interactions between banks and economies, where these banks have exposures or source funding. The two curved arrows on the sides of the figure represent the direction of the two components of the interactions between banks and economies. Source: Budnik et al. [2023].

**Figure 1: Basic model structure**

Bank lending decisions, when aggregated at the country level, influence the macroeconomic outlook of that country. All model equations are solved jointly, which permits acknowledging the contemporaneous nature of feedback loops.

Two sources of uncertainty can enter model simulations. The scenario uncertainty links to residuals in the empirical equations representing the dynamics of the euro area and non-euro area countries. The parameter uncertainty exploits the information on the joint distributions of estimated model parameters concerning country-level and bank-level behavioural equations. Appendix A provides more details about the main mechanisms and properties of the model, while Budnik et al. [2023] includes the comprehensive model description.

<sup>17</sup>These are economies that have the most vital financial links to the euro area. It includes all European Union economies outside the euro area as of 2020 (Bulgaria, Czech Republic, Denmark, Croatia, Hungary, Poland, Romania, Sweden), and two European Free Trade Association (EFTA) economies, Switzerland and Norway. Furthermore, other regions included in the model are Brazil, China, Japan, Mexico, Russia, Sweden, Turkey, the United Kingdom, and the United States.

## 4 Simulation setup

Model simulations start in 2022 Q4 and expand over the 5-year horizon until 2027 Q4. Scenario and parameter uncertainty explore Monte Carlo (MC) simulations. Scenario uncertainty involves bootstrapping reduced-form shocks for the euro area, the rest of the world's economies, and energy prices by applying the wild block algorithm.<sup>18</sup>

This setup looks at three alternative scenario horizons: one year, three, and five years ahead. We focus on a subset of variables describing the macro-financial environment and the banking system. For macro-financial developments, these are GDP, HICP inflation, unemployment rate, equity, house prices, 3-month EURIBOR and 10-year bond yields. They correspond to the set of variables that enter the design of the EBA/SSM scenario. In most instances, and if not mentioned otherwise, the presented macro-financial variables are the euro area averages weighted by the nominal GDP, while the unemployment rate is the average weighted by the country-level labour force.

For the banking sector, we look at the transitional CET1 ratio as a measure of solvency, the change in the transitional CET1 capital, return on assets (ROA) as a measure of bank profitability and the non-performing loan (NPL) ratio as a measure of bank asset quality. Loan volumes and lending rates to the non-financial private sector map the evolution of lending to the euro area real economy. Additionally, we track banks' average debt funding costs and their liquidity mismatches measured by the liquidity coverage ratio (LCR). Bank-level variables are weighted according to the definition of a variable; namely, the CET1 ratio is weighted by bank-level Risk-Weighted Assets, lending volumes, and interest rates to the euro area non-financial system by relative country-level volumes of outstanding and new loans, respectively. Last, we also look at the share of banks, in terms of total banking sector assets, for which the CET1 ratio falls below their joint Pillar I and II requirements (excluding buffers).

## 5 Looking at the full distribution of events

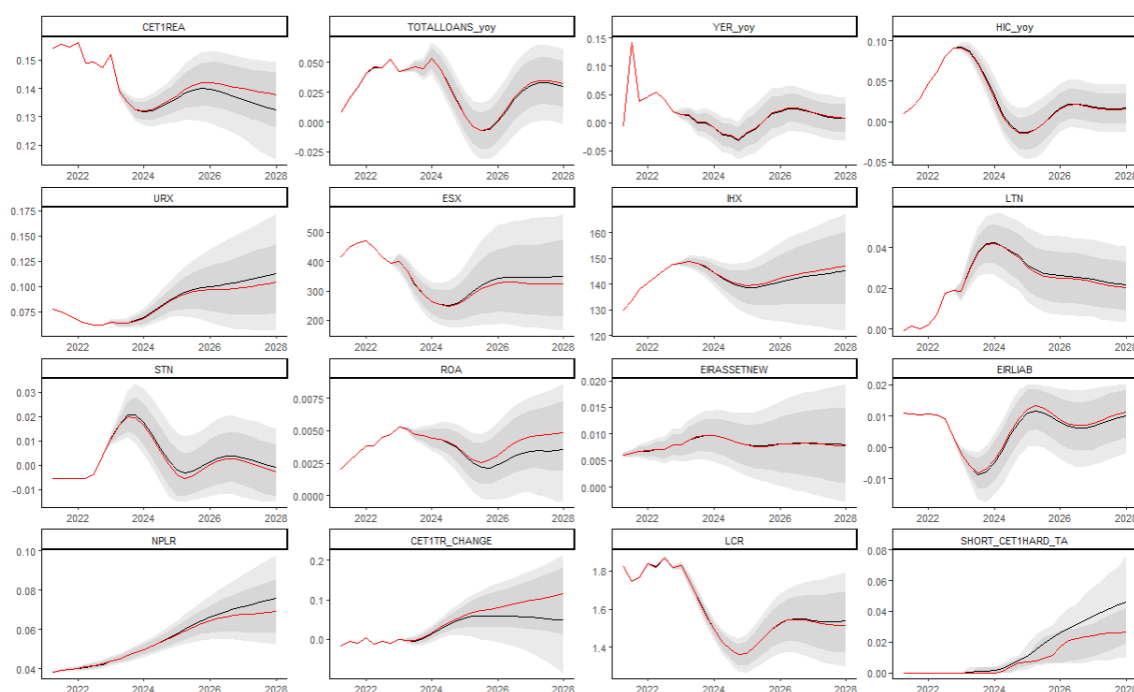
This section first looks at the entire distribution of possible futures. Later, it moves to evaluate solvency risks pertinent to the euro area banking system by discussing the results of the distributional stress test for the overall banking system.

---

<sup>18</sup>We have conducted a detailed robustness check assessing the impact of the sampling scheme on the results by replacing the bootstrapping method with sampling from the estimated parametric distributions of macro-financial reduced-form shocks. With this alternative sampling scheme, the space of future scenarios reflects greater uncertainty. The higher variance of future scenarios translates into more extreme scenarios contained in tails of variable distributions. Otherwise, the sampling scheme does not significantly affect any of the conclusions. The authors can provide the results of the robustness check on request.

## 5.1 Macro-financial scenarios

Figure 2 presents the complete distribution of scenarios. In the medium term, the expected growth rate of GDP and inflation return to their long-term averages of slightly below 2%. The expected 3-month EURIBOR stabilises accordingly. The probability of the 3-month EURIBOR falling below the 0% threshold is around 50%, but that of the 3-month EURIBOR returning to all-time lowest levels observed in the low inflation environment of 2020 – 2022 is significantly below 10%. The expected level of government bond yields remains more than two percentage points above 3-month EURIBOR, reflecting time and risk premium. The euro area stock and house price indices follow an increasing trajectory.



**Figure 2:** Full distribution of plausible scenarios

Notes: CET1REA – CET1 ratio, TOTALLOANS\_yoy – annual bank lending volumes to the non-financial private sector growth rate, YER\_yoy – annual GDP growth rate, HIC\_yoy – HICP inflation, URX – unemployment rate, ESX – equity price index, IHX – house price index, LTN – 10-year bond yields, STN – 3-month EURIBOR, ROA – return on assets, ERASSETNEW - the interest rate on new lending to the non-financial private sector, EIRLIAB - the average debt funding costs, NPLR – NPL ratio, CET1TR\_CHANGE - percentage change in CET1 capital compared to the end 2022, LCR - liquidity coverage ratio, SHORT\_CET1HARD\_TA - the share of banks falling below their CET1 minimum and Pillar II requirements in the total assets of the euro area banking system. Red line: median, black line: mean, dark field 60%, lighter field 80% probability span.

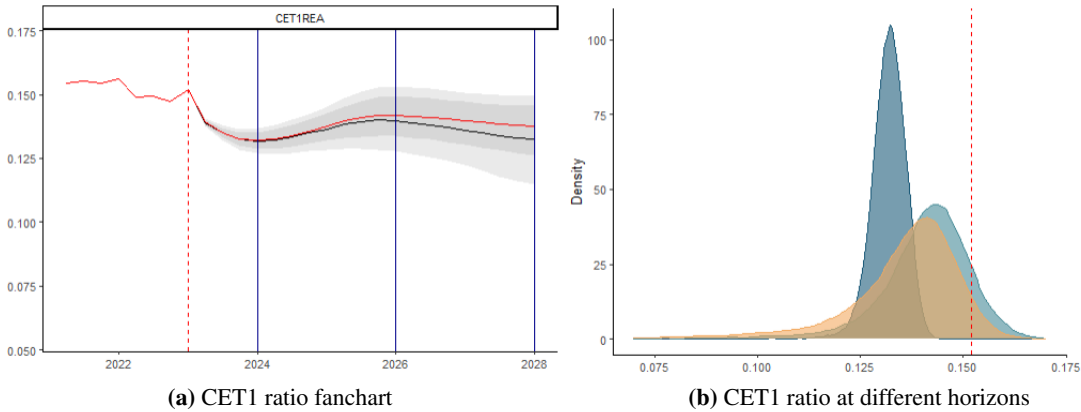
The expected bank profitability stays positive, and the average period ROA is around 0.5% over the simulation horizon. The quality of the bank assets deteriorates slightly, which partially derives from model assumptions such as missing sale-offs of assets and partially from the gradual normalisation of labour markets predicted at the longer end of the horizon.

CET1 capital increases over time in line with the expansion of bank assets, therein loans to the non-financial private sector. At the same time, the ratio of CET1 capital to the risk-weighted amounts decreases somewhat in the longer horizon. This sliding path of the CET1 ratio in the longer horizon reflects conservative assumptions of the model that do not allow banks to access capital markets, exclude bank recapitalisations, entry of new, more efficient banks, adaptations of their business models or restructurings. Lastly, the liquidity coverage ratio stabilises at a level that is safely above the required 100% but below the level in 2022. The latter correction reflects the assumed normalisation of monetary policy and the partial pull-out of central bank liquidity in the medium run.

The estimates of scenario uncertainty are relatively broad and extend over the horizon, especially for index and stock variables such as the CET1 ratio or equity indices.

### 5.2 Tails of bank solvency

The banking system CET1 ratio distribution in Figure 3 comprehensively summarises system-wide solvency risks. The fan chart of the CET1 ratio broadens and becomes more asymmetric with increasing scenario horizon. At the 1-year horizon, the distribution of the CET1 ratio approximately centres around its mean. Over time, the left tail of the CET1 distribution becomes longer and thicker. It reflects an increasing mass of probability falling on scenarios involving amplification mechanisms.



**Figure 3:** CET1 ratio at three horizons: 1-year, 3-years and 5-years

Notes: LHS chart: red line - median, black line - mean, dark field 60%, lighter field 80% probability span. Blue horizontal lines mark, starting from the left hand side, the starting point (end 2022), 1-year, 3-year and 5-year forecast horizon. RHS chart: red horizontal line - the starting point (end 2022) CET1 ratio, Gaussian kernel estimates of the probability density of CET1 ratios at different forecast horizons.

Table 1 selects the mean, median and candidate tail metrics of 10<sup>th</sup>, 5<sup>th</sup> and 1<sup>st</sup> percentiles for the system-wide CET1 ratio. The mean and median CET1 ratio remains around three percentage points

above the regulatory requirements over the 5-year horizon. However, the CET1 ratio in the tail systematically decreases over time, increasing the probability that the banking sector eventually faces significant trouble. The probability of the CET1 ratio dropping 1.5 percentage points below its regulatory thresholds in one year is firmly below 1%. In three years, it is above 1%, but below 5%, and in five years, it is already above 5%. Notwithstanding, these probabilities are reasonably low, speaking of the high capacity of the euro area banking system to contain stress.

	Mean	Median	10 perc	5 perc	1 perc
Starting point	15.2%				
1-year	13.2%	13.2%	12.7%	12.5%	11.8%
3-year	13.9%	14.2%	12.7%	11.8%	7.2%
5-year	12.8%	13.8%	11.5%	9.5%	-3.2%

**Table 1:** System-wide CET1 outcomes at different horizons

## 6 Distributional stress testing in a heterogenous banking system

This section illustrates the applications of distributional stress testing to bank-level risks. It shows that the distributional stress testing of individual banks can detect pockets of vulnerabilities inherent to their unique balance sheets. It can also recover the share of systemic and idiosyncratic components of banks' risks and replicate risk indicators such as the systemic expected shortfall (SES) proposed by Acharya et al. [2017].<sup>19</sup> The section also compares bank-level CET1 ratios derived in the distributional stress test with the real-life 2023 EBA/SSM exercise results.

### 6.1 Solvency at risk and systemic stress measures

The lower percentiles of the distribution of their CET1 ratio can most straightforwardly measure the general banks' sensitivity to macro-financial conditions. This measure, built analogously to the measure of system-wide solvency risks and described in the entire space of possible futures, is dubbed the stressed CET1 ratio. The stressed CET1 reflect scenarios that likely differ between individual banks and are specific to their portfolios and business models (Flood and Korenko [2015]). They summarise bank idiosyncratic and systemic vulnerabilities.

Two alternative vulnerability measures examine bank outcomes conditional on system-wide distress. The first metric is related to the SES of Acharya et al. [2017] and amounts to the mean bank-level CET1

<sup>19</sup>The original contributions of Acharya et al. [2017] rely on market data and simple time series models.

conditional on the system-wide CET1 ratio falling in its lower percentile. This measure, which we call the expected systemic CET1 ratio, pins down the relative performance of banks in macro-financial scenarios, which strain the overall system. The second metric measures the probability that a bank is under stress, that is, falls in the lower percentiles of its CET1 ratio, while the banking system being in trouble. This measure, which captures the congruence between bank-specific and system-wide stress, is called the systemic concordance for the CET1 ratios.

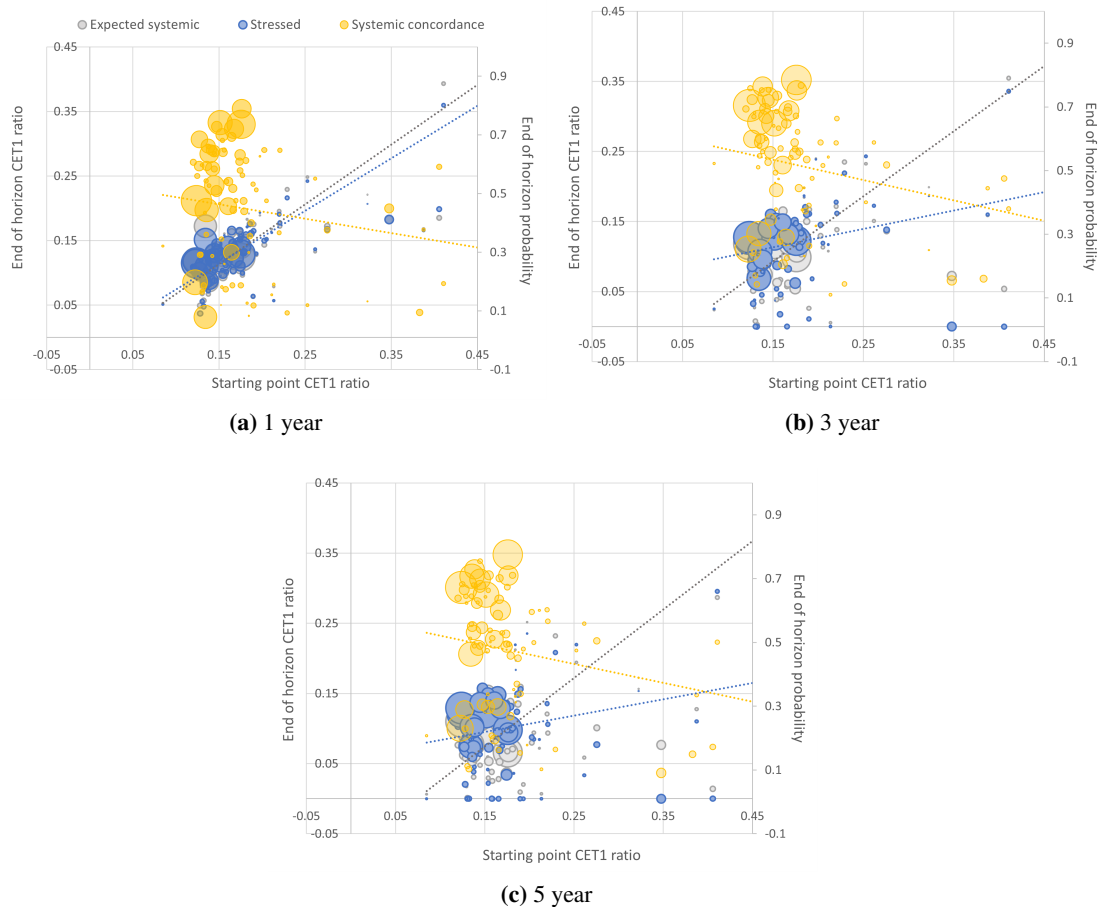
The two measures of systemic vulnerability are complementary. The expected systemic CET1 ratio evaluates a bank's performance in a systemic stress episode. However, whether the event pushes the bank to the limits of its resilience remains uncovered. For example, a deep recession originating in the euro area will likely become a systemic event for the euro area banking system. Nevertheless, a bank with most of its exposures outside the euro area can fare relatively well, at least when compared with a hard-hit crisis in one of its significant exposure countries. The systemic concordance for the CET1 ratios informs about the likelihood that a systemic event is all times worse from the bank's perspective.

Figure looks at these three alternative tail risk measures of the CET1 ratio compared to bank solvency at the horizon's beginning. The blue balls represent the stressed CET1 ratio, the grey balls represent the expected systemic CET1 ratios, and the yellow balls represent the systemic concordance for the CET1 ratios. The measures correspond with the 10<sup>th</sup> percentile of the respective distributions in each case.

A solid positive relationship exists between the initial CET1 ratio and general and systemic resilience. It is vital in the short run and abates moderately with an increasing horizon. This positive dependence on the CET1 ratio from the starting point is stronger and more stable for the systemic expected CET1 ratios over time. The correlation coefficient between the initial and systemic expected bank CET1 ratios is as high as 0.97 in the 1-year and levels off at 0.81 in the 5-year horizon. The correlation coefficient between the initial and stressed CET1 ratios decreases more dramatically, from 0.9 in the 1-year to 0.35 in the 5-year horizon.

The longer the horizon, the more likely the emergence of specific and "unlikely" scenarios that strain banks with specific exposures and a business model without exposing the general banking systems. Intuitively, this happens mostly for smaller banks. The correlation between actual bank solvency and bank propensity to experience significant stress other than a systemic event is weakly positive and approximately constant over time. The corresponding correlation coefficients between the systemic concordance for CET1 ratios and the banks' initial solvency rate are weakly negative and between -0.4 and -0.28 for all horizons.

This set of results concludes that actual bank solvency ratios are a good measure of the bank's ability



**Figure 4:** CET1 threshold of euro area scenarios

Notes: Blue balls - stressed CET1 ratio for 10<sup>th</sup> percentile, gray balls - systemic expected CET1 ratio for 10<sup>th</sup> percentile, yellow balls - systemic concordance for the CET1 ratio for 10<sup>th</sup> percentile. The size of balls corresponds to the size of banks' assets at the end of 2022. The graphs zoom in on the banks with the initial CET1 ratio (at the end of 2022) below 45%. Lines represent the fitted linear regression between (unweighted) variables for the full sample of banks.

to withstand temporary stress and, for the largest banks, also long-lasting systemic stress. They lose their information power the more extended the stress horizon, particularly for smaller banks. Appendix B includes further exploration of these indicators.

## 6.2 Comparison to the EBA/SSM stress test results

The different metrics derived from the distributional stress test can be contrasted with the actual EBA/SSM stress test 2023 results.<sup>20</sup> The left-hand side graph in Figure 5 reports the change in bank-level CET1 ratios from their starting point for the stressed and systemic expected CET1 ratios from the distributional

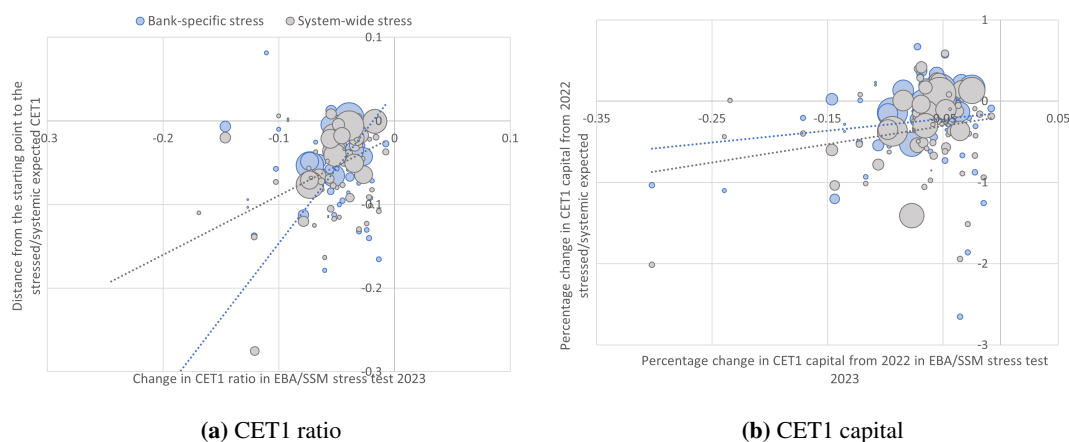
<sup>20</sup>In this, we use that the starting points of our and the EBA/SSM stress test are congruent, although the latter stress test accommodates a more comprehensive update of banks' balance sheet results at the end of 2022.



stress test against the results of the European stress test.

The first observation is that the results of the EBA/SSM stress test are, in broad terms, closer to the systemic expected CET1 ratios than to the stressed CET1 ratios. It is an intuitive result. Relying on a single adverse scenario for all banks is likely to lead to the design of scenarios that reflect system-wide stress. The correlation coefficient between the drop in the CET1 ratio, commensurate with the systemic expected CET1 ratio, and that observed in the actual EBA/SSM results, is very high, amounting to 0.97. Accordingly, the relative performance of banks in the 2023 stress test was very close to that implied by systematically scanning all plausible and unfavourable developments for the euro area banking sector.

The second observation is that, on average, large banks appear to be under sufficient stress in the EBA/SSM exercise, whereas for smaller banks, the picture is more varied. Banks to the right of the two regression lines are those for which the severity of the EU-wide scenario is relatively low, least compared to the distributional stress test metrics derived at the 10<sup>th</sup> percentile threshold. Banks to the left of the regression lines can be assessed as sufficiently stressed. Most of the largest banks are left to the lines, whereas smaller banks are scattered on both sides of the lines.



**Figure 5:** Distributional stress test versus EBA/SSM stress test changes in CET1 ratio and capital over 3-year horizon

Notes: Blue balls - stressed CET1 ratio for 10<sup>th</sup> percentile, gray balls - systemic expected CET1 ratio for 10<sup>th</sup> percentile. The size of balls corresponds with the size of banks' assets at the end of 2022. In the macroprudential stress test capital losses over 100% of the initial CET1 capital are possible, as the stress test incorporates capital transformation triggers for AT1 and T2 instruments.

Another way to compare stress test exercises involves looking at changes in CET1 capital. Such changes measure the actual losses carried out by investors or governments in the case of a bail-in or bailout. Figure 5 shows the stressed and systemic expected capital losses measured in the 10<sup>th</sup> percentile of the respective distributions, compared to the corresponding EBA/SSM stress test results. On average,

the stressed and systemic expected capital losses are larger than the capital losses in the supervisory stress test. However, these differences mainly reflect the constant versus dynamic balance sheet assumption. Capital losses in the dynamic balance sheet tend to be greater (Budnik et al. [2021a], Budnik et al. [2023]) than under the constant balance sheet, although they can even turn negative.<sup>21</sup>

The patterns exhibited by capital losses are generally consistent with those observed for changes in the CET1 ratios. The correlation between the distributional and supervisory stress test results is weaker for capital losses than for the CET1 ratios. However, the largest banks consistently experience relatively high stress in the EBA/SSM against the distributional stress test. The severity of the EBA/SSM compared to the same distributional stress test benchmark differs significantly for smaller banks.

The comparisons illustrate how the distributional stress test can support traditional supervisory stress tests. It can serve as ex-ante or ex-post evaluation metrics for the sufficiency of stress in other bottom-up or top-down exercises, both for the overall banking system and individual banks. Banks identified as "under-stressed" can be subject to additional supervisory scrutiny, for instance, while applying bank-specific capital charges. Then, a distributional stress test can classify banks between those most exposed to systemic risks and those most sensitive to developments in specific markets, risks, or jurisdictions. Finally, the distributional stress test can become a relevant risk assessment tool, possibly more fitted than system-wide stress test exercises, especially for the latter banks.

## 7 Reverse stress testing

This chapter explores the multiple-scenario stress test to delineate plausible scenarios with the desired severity. Reverse stress testing is straightforward to implement when there is a single risk factor and the outcome variable. In this trivial case, reverse stress testing requires an inversion of the mapping function to find risks corresponding to the preferred outcomes. However, such an inversion becomes very complex in real-life stress tests, with multiple macro-financial factors, multiperiodicity, and possibly nonlinear transmission of risks into institutions' balance sheets.

We implement the worst-case and threshold examples of a reverse stress test to illustrate the flexibility of our approach. Furthermore, we discuss an example of selecting plausible and severe scenarios with a desired economic narrative.

---

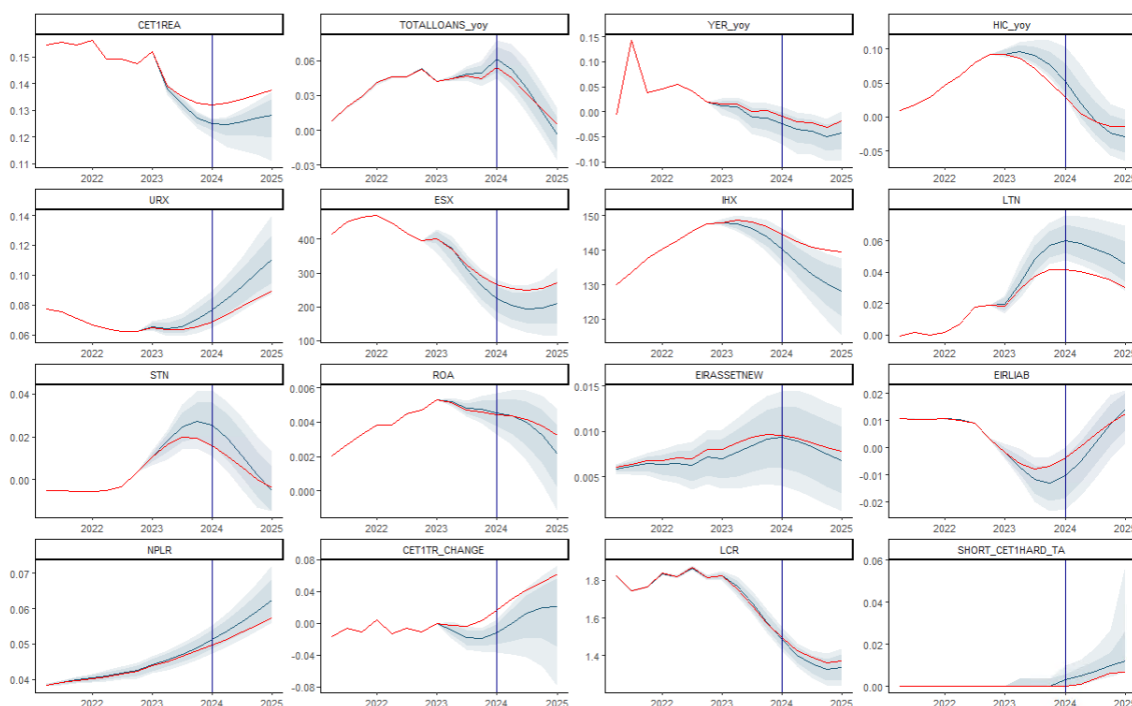
<sup>21</sup>For instance, a bank under stress may choose to reduce its capital charges by reducing higher-risk and increasing lower-risk exposures. Although this portfolio rebalancing is likely to be reflected in cuts in corporate lending and contributes to the bleak economic outlook, individual banks can book positive profits as long as the low-risk exposures are high-yielding. Such adjustments are absent from constant balance sheet stress tests.

## 7.1 Worst-case scenarios

Figure 6 illustrates the 10% worst-case scenarios for the one-year horizon CET1 ratio at the system level. The red lines mark the medians of the variables in the entire distribution (as in Figure 2), while the blue lines and fan charts summarise the medians and distributions of the variables in the worst-case scenarios.

A substantial reduction in bank solvency on a short horizon emerges with low growth and high inflation scenarios. Elevated inflation triggers a monetary policy response, leading to pronounced corrections in asset prices. Falling stock and bond prices translate into bank revaluation losses, shrinking their capital.

High inflation fuels an increase in nominal bank lending volumes. Although loan volumes to the non-financial private sector in real terms are below their median in the whole scenario distribution, their expansion in nominal terms inflates the denominator of CET1.



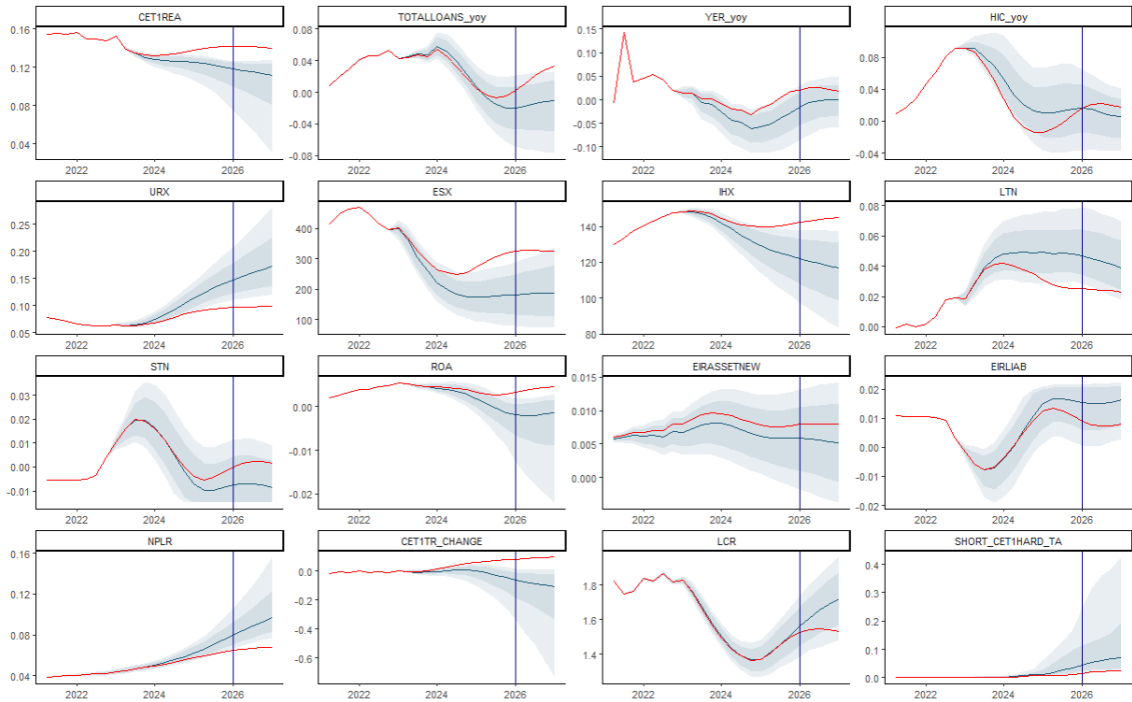
**Figure 6:** Worst-case scenarios for system-wide CET1 ratio at 1-year horizon

Notes: CET1REA – CET1 ratio, TOTALLOANS\_yoy – annual bank lending volumes to the non-financial private sector growth rate, YER\_yoy – annual GDP growth rate, HIC\_yoy – HICP inflation, URX – unemployment rate, ESX – equity price index, IHX – house price index, LTN – 10-year bond yields, STN – 3-month EURIBOR, ROA – return on assets, EIRASSETNEW – the interest rates on new lending to the non-financial private sector, EIRLIAB – the average cost of debt funding, NPLR – NPL ratio, CET1TR\_CHANGE – percentage change in CET1 capital compared to the end 2022, LCR – liquidity coverage ratio, SHORT\_CET1HARD\_TA – the share of banks falling below their CET1 minimum and Pillar II requirements in the total assets of the euro area banking system. Red line: median for the full distribution of events, blue line: median of the relevant worst-case scenarios, blue darker field 60%, blue lighter field 80% probability span of the latter scenarios. The navy blue vertical line marks the end of the 1-year scenario horizon.

Short-term worst-case scenarios for bank solvency resemble the evolution of the euro area economy

at the exit from the COVID-19 pandemic. They combine fragile economic growth, sharply increasing inflation, and normalising monetary policy that increasingly affected financial markets.

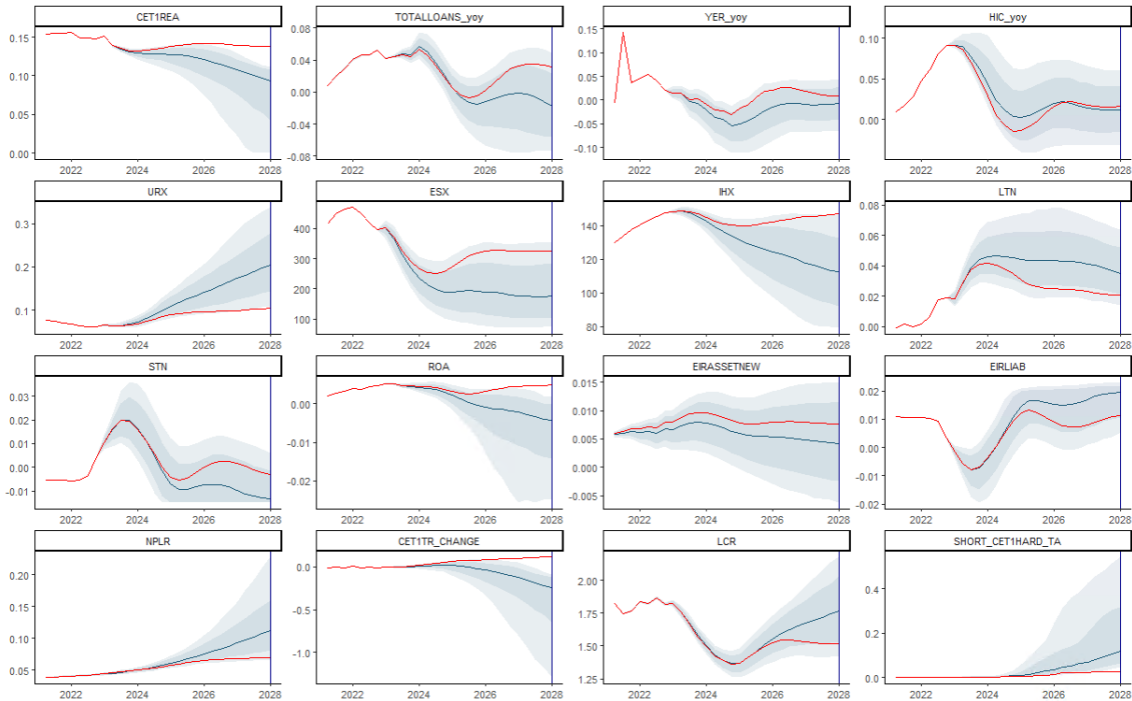
The medium-term risks to the bank’s solvency reflect a classical recession. The 10% worst-case scenarios for the CET1 ratio in the 3-year horizon are illustrated in Figure 7.3. They depict a sustained contraction in output, an increase in unemployment, and low inflation but a distinctive correction in asset prices. Bank funding costs increase on the back of their deteriorating balance sheets and increasing risk premia. However, with weak loan demand, banks struggle to pass through higher funding costs into lending rates. Bank lending volumes contract, their asset quality deteriorates, and profitability decreases.



**Figure 7:** Worst-case scenarios for system-wide CET1 ratio at 3-year horizon

Notes: CET1REA – CET1 ratio, TOTALLOANS\_yoy – annual bank lending volumes to the non-financial private sector growth rate, YER\_yoy – annual GDP growth rate, HIC\_yoy – HICP inflation, URX – unemployment rate, ESX – equity price index, IHX – house price index, LTN – 10-year bond yields, STN – 3-month EURIBOR, ROA – return on assets, EIRASSETNEW – the interest rates on new lending to the non-financial private sector, EIRLIAB - the average cost of debt funding, NPLR – NPL ratio, CET1TR\_CHANGE - percentage change in CET1 capital compared to the end 2022, LCR - liquidity coverage ratio, SHORT\_CET1HARD\_TA - the share of banks falling below their CET1 minimum and Pillar II requirements in the total assets of the euro area banking system. Red line: median for the full distribution of events, blue line: median of the relevant worst-case scenarios, blue darker field 60%, blue lighter field 80% probability span of the latter scenarios. The navy blue vertical line marks the end of the 3-year scenario horizon.

The most prevalent solvency risks in the long run relate to a long-lasting recession. The risks summarised in Figure 8 are qualitatively similar to medium-term risks and reflect a recession lasting not 3 but 5 years. With a high probability, monetary policy hits the lower zero bound. Banks continually run down their capital, leading to very adverse outcomes at the system level.



**Figure 8:** Worst-case scenarios for system-wide CET1 ratio at 5-year horizon

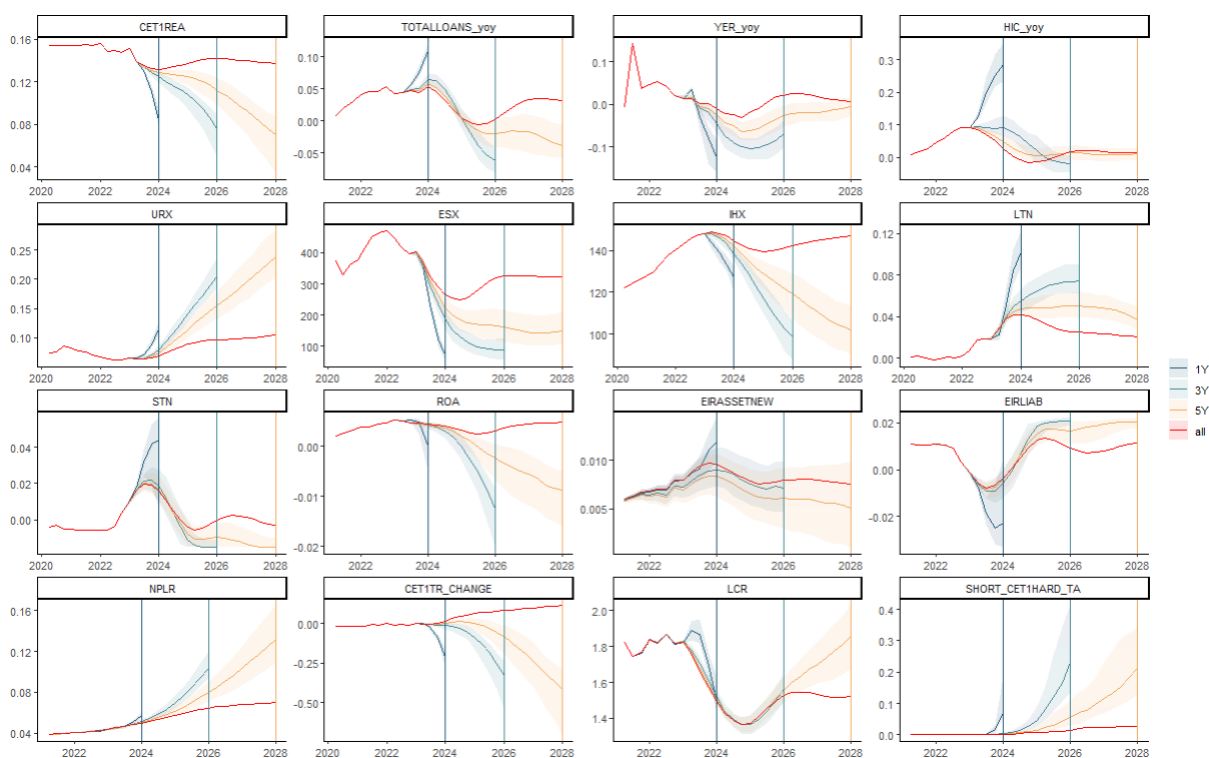
Notes: CET1REA – CET1 ratio, TOTALLOANS\_yoy – annual bank lending volumes to the non-financial private sector growth rate, YER\_yoy – annual GDP growth rate, HIC\_yoy – HICP inflation, URX – unemployment rate, ESX – equity price index, IHX – house price index, LTN – 10-year bond yields, STN – 3-month EURIBOR, ROA – return on assets, ERASSETNEW – the interest rates on new lending to the non-financial private sector, EIRLIAB - the average cost of debt funding, NPLR – NPL ratio, CET1TR\_CHANGE - percentage change in CET1 capital compared to the end 2022, LCR - liquidity coverage ratio, SHORT\_CET1HARD\_TA - the share of banks falling below their CET1 minimum and Pillar II requirements in the total assets of the euro area banking system. Red line: median for the full distribution of events, blue line: median of the relevant worst-case scenarios, blue darker field 60%, blue lighter field 80% probability span of the latter scenarios. The navy blue vertical line marks the end of the 5-year scenario horizon.

## 7.2 Threshold scenarios

The alternative avenue to ensure sufficient severity of a stress test is to focus on scenarios with a postulated maximum level of the CET1 ratio. Figure 9 illustrates the outcome of scenario selection with the system-wide CET1 ratio not higher than 10% on three alternative stress test horizons.

The short- to long-term risks to bank solvency identified with this approach are qualitatively similar to those uncovered by the corresponding worst-case scenarios. Short-term risks to bank solvency relate to a sudden acceleration of inflation and a decisive reaction of monetary policy. Medium- to long-term risks emerge in a recessionary environment. The recession that brings the bank solvency below 10% in the 3-year horizon is deeper and commensurate with sharper corrections in asset prices than the recession resulting in a below 10% solvency rate in the 5-year horizon.

There are different trade-offs in designing worst-case and threshold scenarios that surface mostly in short-term stress tests. The main advantage of worst-case selection is controlling for the scenario's plau-



**Figure 9:** Macro-financial scenarios with CET1 ratio  $\leq 10\%$  for 1-year, 3-year and 5-year horizon

Notes: CET1REA – CET1 ratio, TOTALLOANS\_yoy – annual bank lending volumes to the non-financial private sector growth rate, YER\_yoy – annual GDP growth rate, HIC\_yoy – HICP inflation, URX – unemployment rate, ESX – equity price index, IHX – house price index, LTN – 10-year bond yields, STN – 3-month EURIBOR, ROA – return on assets, ERASSETNEW – the interest rates on new lending to the non-financial private sector, EIRLIAB - the average cost of debt funding, NPLR – NPL ratio, CET1TR\_CHANGE - percentage change in CET1 capital compared to the end 2022, LCR - liquidity coverage ratio, SHORT\_CET1HARD\_TA - the share of banks falling below their CET1 minimum and Pillar II requirements in the total assets of the euro area banking system. Red line: median for the full distribution of events, dark blue line: median for threshold scenarios with the  $\leq 10\%$  CET1 ratio in the 1-year horizon, lighter blue line: median for threshold scenarios with the  $\leq 10\%$  CET1 ratio in the 3-year horizon, orange line: median for threshold scenarios with the  $\leq 10\%$  CET1 ratio in the 5-year horizon. Fancharts span 40% probability span of the corresponding threshold scenarios. Vertical lines mark the end of the scenario horizons.

sibility and ensuring its meaningful quantitative and qualitative interpretation. At the same time, the level of the outcome variable must remain consistent with policymakers' expectations. Threshold scenarios provide more flexibility in choosing the minimum stress on the outcome variable but can compromise scenario plausibility. The probability of the CET1 ratio falling below 10% in the 5-year horizon is 6%, 3-year, 2%, and 1-year a minuscule 0.1%. Such low plausibility can harm the interpretation of the stress test results.

Another observation is that for both types of reverse stress testing, a significant nonlinearity emerges between the severity of the outcome variable and the adversity of macro-financial developments. An additional reduction of 1 percentage point in the system-wide CET1 ratio is related to a more considerable deterioration in the economic outlook in the permissive stress test and more minor in the already

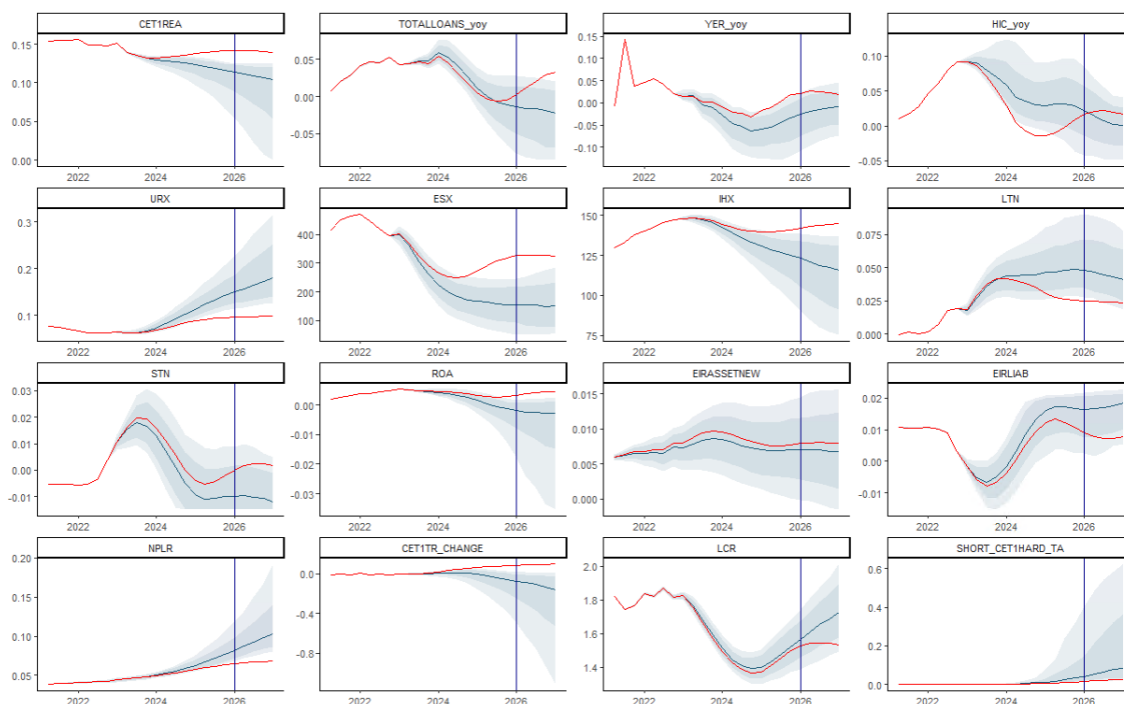
conservative ones. This nonlinearity partially reflects the emergence of amplification mechanisms in the tails of the CET1 distribution. The Appendix C presents a detailed discussion of these aspects.

These properties of reverse stress test scenarios contain relevant information for designing any stress test. For any preferred scenario horizon, the distribution of the outcome variable should inform the ambition concerning stress test severity. Informed choices of scenario severity can prevent the construction and communication of highly implausible scenarios. Another aspect is that scaling risk factors does not result in a proportionate increase in the severity of the stress test. Even moderate changes in the adversity of macro-financial scenarios, such as in the EBA/SSM process, can sometimes lead to substantial changes in the resulting CET1 ratios or the overall stress test plausibility. The reverse holds also. Large changes in macro-financial factors must not cause significant shifts in CET1 ratios, potentially "disappointing" stress test designers.

### **7.3 Selecting a desired narrative**

Here, we consider a simple example in which a scenario designer, having set stress test plausibility and severity, wishes to emphasise a narrative related to monetary policy stance. Figure 10 and 11 contrast the worst-case scenarios with different monetary policy stances. The two families of scenarios emerge from cutting the space of worst-case scenarios for 10% system-wide CET1 ratio (as in Figure 7.3) into two equal slices based on the level of the three-year average of the real interest rate. The relatively loose monetary policy stance in Figure 10 is synonymous with the real interest rate being below its median in the original worst-case scenario, and the relatively tight stance amounts to the real interest rate being above its median.

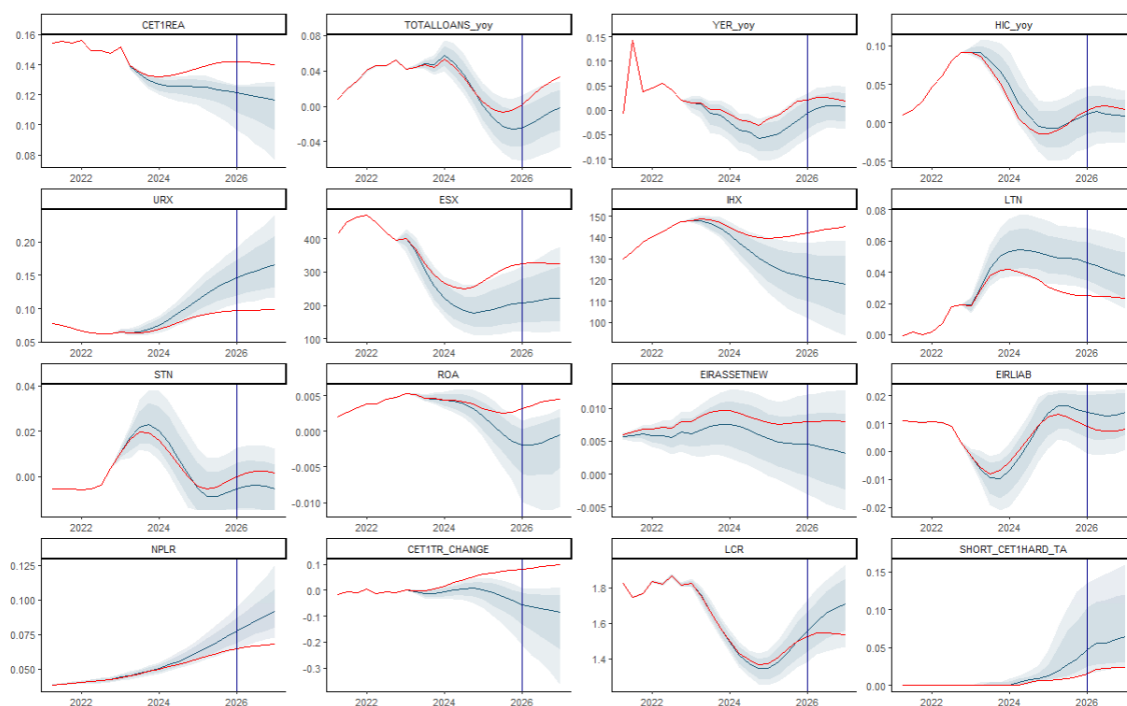
The two scenarios have similar severity and the same probability of realisation of 5%, while they differ in the behaviour of some macro-financial variables. Relatively loose monetary policy emerges in stagflation scenarios with high inflation and low GDP growth, while relatively tight policy emerges in scenarios with a more positive economic outlook. Scenarios with looser monetary policy and higher inflation appear to put more pressure on the real economy. There, unemployment is higher, and asset prices and real lending to the non-financial private sector are lower than in scenarios with higher real interest rates. The results suggest that at the beginning of 2023, a future monetary policy that was too loose rather than too tight constituted a pronounced risk to the euro area's macro-financial system.



**Figure 10:** Worst-case scenarios for system-wide CET1 ratio at 3-year horizon with looser monetary policy

Notes: CET1REA – CET1 ratio, TOTALLOANS\_yoy – annual bank lending volumes to the non-financial private sector growth rate, YER\_yoy – annual GDP growth rate, HIC\_yoy – HICP inflation, URX – unemployment rate, ESX – equity price index, IHX – house price index, LTN – 10-year bond yields, STN – 3-month EURIBOR, ROA – return on assets, ERASSETNEW – the interest rates on new lending to the non-financial private sector, EIRLIAB - the average cost of debt funding, NPLR – NPL ratio, CET1TR\_CHANGE - percentage change in CET1 capital compared to the end 2022, LCR - liquidity coverage ratio, SHORT\_CET1HARD\_TA - the share of banks falling below their CET1 minimum and Pillar II requirements in the total assets of the euro area banking system. Looser monetary policy corresponds to the period average real interest rate (EURIBOR minus HICP inflation rate) below its median for worst-case scenarios with 10% plausibility. Red line: median for the full distribution of events, blue line: median of relevant worst-case scenarios with looser monetary policy, blue darker field 60%, blue lighter field 80% probability span of the latter scenarios.





**Figure 11:** Worst-case scenarios for system-wide CET1 ratio at 3-year horizon with tighter monetary policy

Notes: CET1REA – CET1 ratio, TOTALLOANS\_yoy – annual bank lending volumes to the non-financial private sector growth rate, YER\_yoy – annual GDP growth rate, HIC\_yoy – HICP inflation, URX – unemployment rate, ESX – equity price index, IHX – house price index, LTN – 10-year bond yields, STN – 3-month EURIBOR, ROA – return on assets, EIRASSETNEW – the interest rates on new lending to the non-financial private sector, EIRLIAB - the average cost of debt funding, NPLR – NPL ratio, CET1TR\_CHANGE - percentage change in CET1 capital compared to the end 2022, LCR - liquidity coverage ratio, SHORT\_CET1HARD\_TA - the share of banks falling below their CET1 minimum and Pillar II requirements in the total assets of the euro area banking system. Tighter monetary policy corresponds to the period average real interest rate (EURIBOR minus HICP inflation rate) above its median for worst-case scenarios with 10% plausibility. Red line: median for the full distribution of events, blue line: median of relevant worst-case scenarios with tighter monetary policy, blue darker field 60%, blue lighter field 80% probability span of the latter scenarios.

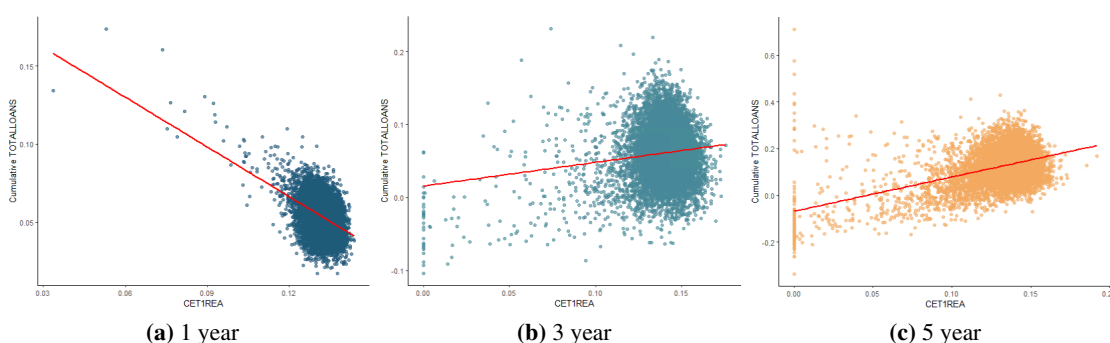
## 8 Looking beyond system-level solvency

This section elaborates on applying multiple-scenario stress testing in cases where, due to circumstantial or institutional reasons, we aim to look at the financial system’s resilience in a more holistic manner. The first example concerns a stress test that simultaneously focuses on bank solvency and lending. The second example looks at a stress test emphasising the emergence of endogenous financial stability risks. Additionally, Appendix D discusses the design of a joint solvency liquidity stress test.

### 8.1 Macroprudential trade-offs: solvency and lending

So far, all examples of applying our methodology have referred to solvency criteria. However, from a macroprudential viewpoint, bank lending is an equivalently relevant outcome of a stress test. A macroprudential policymaker would like to know whether, in adverse circumstances, banks will be in a position to continue providing lending to the economy.

Figure 12 points out that the system-wide solvency and lending outcomes are correlated. Interestingly, this correlation changes over the stress test horizon. In the short term, it is negative (the correlation coefficient is close to 0.2, with a p-value of zero). It reflects the predominance of the denominator effect, where a sudden increase in lending volumes leads to the expansion of risk-weighted amounts and at least a temporary drop in the CET1 ratio. The longer the horizon, the more precise the positive relationship between bank solvency and their ability to provide lending to the economy. In the five-year horizon, the correlation coefficient is close to 0.8, with a p-value of zero. The medium- to long-run positive correlation emerges because banks adapt their assets to preserve stable solvency rates and consistently meet capital requirements and buffers.

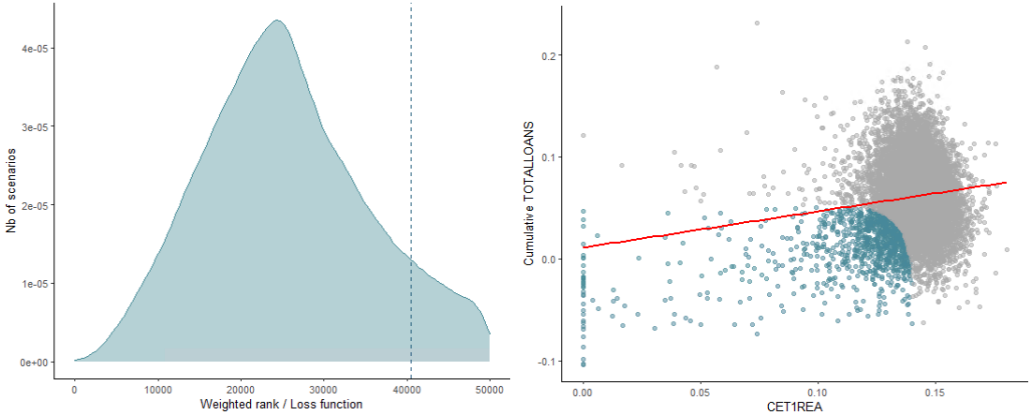


**Figure 12:** Correlation of system-wide CET1 and euro area lending in different futures

Notes: Observations in the graphical representation are filtered using Cook’s distance for better readability.

A macroprudential policymaker can proceed with designing a stress test with a sufficient level of severity for solvency and lending. The first step involves describing the relative relevance of the two variables. We take an ambivalent stance in our example and assign system-wide solvency and lending outcomes equal weights. However, a conservative policymaker in benign times can aim at a higher weight assigned to solvency versus lending. In turn, a crisis may ask for higher weight on lending. We should also focus on the 3-year horizon and worst-case scenarios with a 10% severity level.

The left panel in Figure 13 shows the distribution of scores for scenarios sorted along with the level of CET1 ratio and the cumulative change in lending to the non-financial private sector at the end of the horizon. The right panel of Figure 13 illustrates the selection of the two-dimensional space of the CET1 ratio and cumulative lending at the end of the horizon.



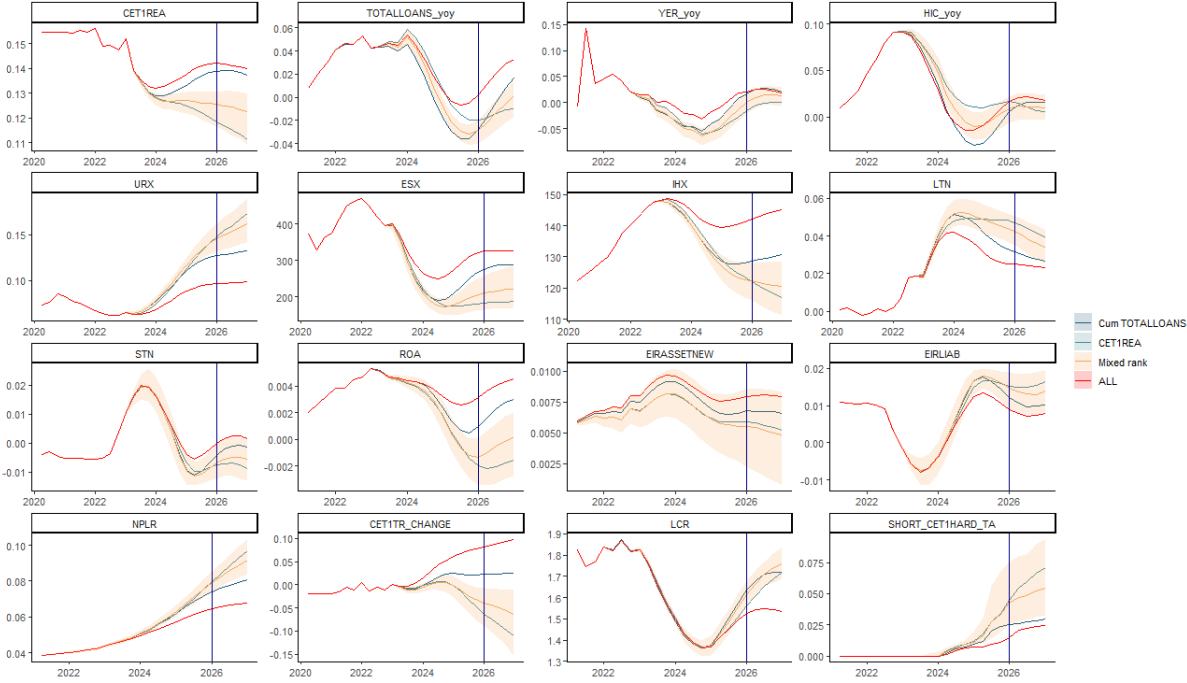
(a) Loss function at 3-year horizon (equal weights) (b) Worst-case selection for CET1 ratio and cumulative lending change

**Figure 13:** Loss function and scenario selection for the joint CET1 ratio and lending growth stress test

The results of such a stress test are illustrated in Figure 14 and contrasted with two reference stress tests, which aim at 10<sup>th</sup> percentile of bank solvency and bank lending separately. The scenarios stressing macroprudential resilience are somewhere in between those stressing the CET1 ratio and lending individually. The system-level CET1 ratio is the lowest in the worst-case scenarios, focusing only on solvency and more benign in the joint stress test. Analogously, the growth of lending to the non-financial private sector is lower in the worst-case scenarios with lending and less negative in the joint scenarios.

Risks to bank solvency and lending in the medium term are related to an economic recession. Inflation initially overshoots as compared to the median of the entire scenario distribution but goes down toward the horizon’s end. When comparing these three sets of results, the interesting realisation is that unemployment and asset prices emerge as risk factors for bank solvency, but to a lesser extent for bank

lending. Lending, in turn, appears to be more sensitive to monetary policy stance, therein to the level of real interest rates and the expansion of the ECB assets.<sup>22</sup>



**Figure 14:** Worst-case scenarios for bank solvency and lending

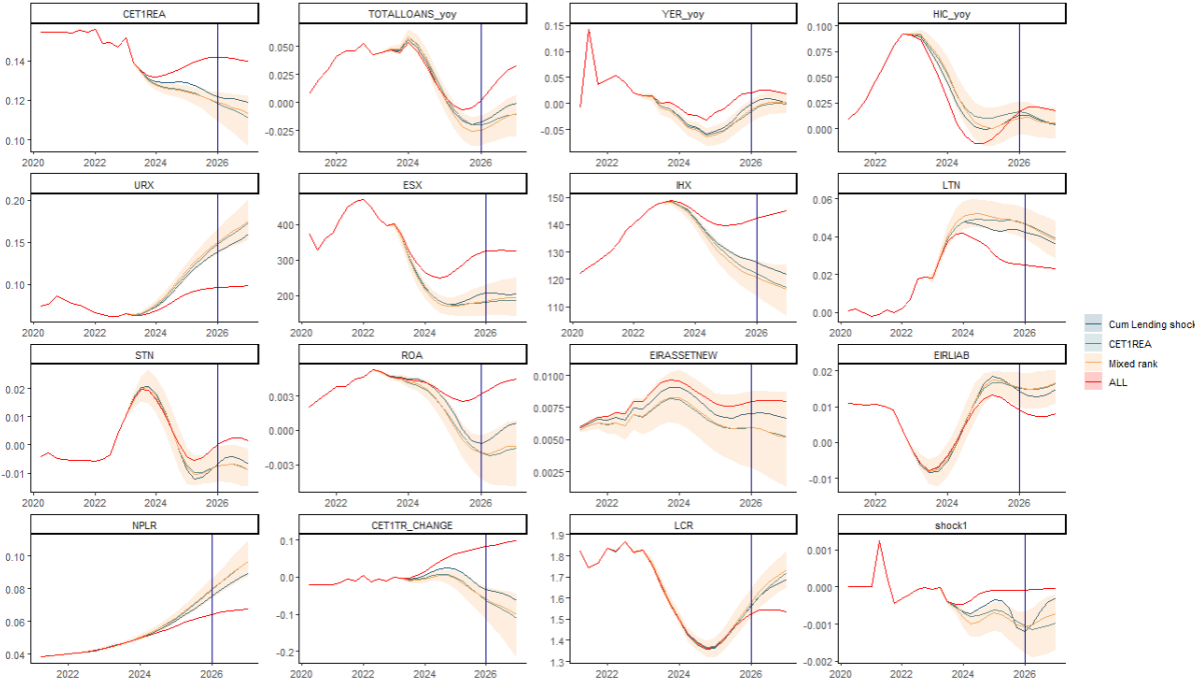
Notes: CET1REA – CET1 ratio, TOTALLOANS\_yoy – annual bank lending volumes to the non-financial private sector growth rate, YER\_yoy – annual GDP growth rate, HIC\_yoy – HICP inflation, URX – unemployment rate, ESX – equity price index, IHX – house price index, LTN – 10-year bond yields, STN – 3-month EURIBOR, ROA – return on assets, EIRASSETNEW – the interest rates on new lending to the non-financial private sector, EIRLIAB - the average cost of debt funding, NPLR – NPL ratio, CET1TR\_CHANGE - the cumulative change in CET1 capital, LCR - liquidity coverage ratio, SHORT\_CET1HARD\_TA - the share of banks falling below their CET1 minimum and Pillar II requirements in the total assets of the euro area banking system. Red line: median for the full distribution of events, dark blue line: median of worst-case scenarios for lending growth, light blue line: median of worst-case scenarios for CET1 ratio, yellow: median of worst-case scenarios for CET1 and lending growth jointly. Fancharts mark 40% probability span for the worst-case scenarios for CET1 and lending growth jointly.

## 8.2 Systemic risks

An alternative approach to thinking of a macroprudential stress test is to put to the front the ability of the banking system to absorb rather than amplify macro-financial shocks. One may wish for scenarios that reveal coordination failures and trespassing the capacity of the banking system to sponge up distress. Such scenarios can uncover hidden vulnerabilities with potentially pronounced effects on the economy that go under the radar of stress tests focused on the adversity of particular banking sector metrics. Moreover, it can be used to cross-check the outcomes of more straightforward scenario designs or, when

<sup>22</sup>It should be noted that, in general, this approach will provide a different assessment of risks than the sequential application of the two criteria in the spirit described in Section 7.3. The latter prioritises the severity described along with the first criteria, such as bank solvency. And will use the second criterion to strengthen the narrative of the scenario.

monitored over time, to deliver information about the evolving resilience of the banking sector.



**Figure 15: Worst-case scenarios for system-wide amplification**

Notes: CET1REA – CET1 ratio, TOTALLOANS\_yoy – annual bank lending volumes to the non-financial private sector growth rate, YER\_yoy – annual GDP growth rate, HIC\_yoy – HICP inflation, URX – unemployment rate, ESX – equity price index, IHX – house price index, LTN – 10-year bond yields, STN – 3-month EURIBOR, ROA – return on assets, ERASSETNEW – the interest rates on new lending to the non-financial private sector, EIRLIAB - the average cost of debt funding, NPLR – NPL ratio, CET1TR\_CHANGE - the cumulative change in CET1 capital, LCR – liquidity coverage ratio, Shock1 - loan supply shock resulting from the banking sector developments. Red line: median for the full distribution of events, dark blue line: median of worst-case scenarios for amplification, light blue line: median of worst-case scenarios for CET1 ratio, yellow: median of worst-case scenarios for CET1 and lending growth jointly. Fancharts mark 40% probability for the worst-case scenarios for CET1 and lending growth jointly.

The selection of scenarios that emphasise systemic amplification of risks focuses on measuring the feedback from the banking sector to the real economy. This measure amounts to the appropriately weighted sum of non-linear adjustments in lending to the non-financial private sector by individual banks.<sup>23</sup> As in previous cases, we select 10% worst-case scenarios, yet this time with the strongest amplification mechanism.<sup>24</sup>

Figure 15 further contrasts the worst-case 10% amplification scenarios with the worst-case scenarios only for solvency and the solvency and lending together. The amplification measure is placed in the

<sup>23</sup>Both selection design (the choice of the measure of negative lending supply feedback from the banking sector to the real economy as relevant selection metrics) and its results are necessarily specific to the model. However, we still aim to emphasise the most universal aspect of such a selection. By focusing on the magnitude of the amplification mechanism (in any model), macroprudential policymakers can receive scenarios at the core of their mandate.

<sup>24</sup>The highest period average level of the measurement variable.

lower left corner of Figure 15 and is lower in all worst-case scenarios than the median in the entire distribution of scenarios. Amplification mechanisms emerge in scenarios with weak economic activity and when banks experience low profitability and capital losses.

Compared to the benchmark joint solvency and lending worst-case scenarios, the stress test that zooms in on the amplification risks hangs stronger on the loan supply channels. Generally, scenarios emphasising solvency and lending and amplification are similar. However, the former instils a higher harmful loan demand component, reflected in a relatively muscular contraction in loan volumes and interest rates.

## 9 Conclusions

Real-life stress test scenarios ordinarily rely on complex macro-financial scenarios based on agreed economic narratives. However, their methodology may not guarantee scenarios simultaneously plausible, severe, and capable of capturing the most critical risks at any given moment (Breuer and Summer [2018]).

The challenges are numerous. First, the concept of "sufficient severity" is often blurred. Although stress tests usually aim to assess financial outcomes, such as banking system solvency, the focus often shifts away from this ultimate goal, and severity becomes a metric attributed to macro-financial outcomes. Although this focus on the severity of risk factors may be justified in linear environments, where their impact on the financial system is always proportional, this is hardly the case in genuine economic environments. Furthermore, adverse scenario narratives, while comprehensive, are susceptible to human biases, including overlooking plausible realities or overestimating the likelihood of improbable events. Consequently, there are often trade-offs between designing statistically plausible scenarios and following the risk narrative.

This paper introduces a conceptually neat approach that addresses many of the challenges faced in regulatory stress tests. The essential step is initiating the stress test design by examining the distribution of potential future scenarios. This distribution should possess three crucial attributes.

First, it should include realistic scenarios incorporating clear-cut economic narratives. Reduced-form models and purely data-driven approaches are unlikely to offer such scenarios. We propose using models with structural elements, such as semi-structural models that combine statistical analysis with structural identification. Semi-structural setups are flexible and can accommodate additional indicators or policy beliefs, enhancing the interpretability of produced scenarios.

Then, the scenario space should encompass a wide range of scenario risks, including model uncer-

tainty. While our application addresses parameter uncertainty in generating scenario spaces, one can also consider combining results from different models, each emphasizing different frictions or utilizing different datasets.

Finally, the scenario space should encompass all risk factors and outcome variables of the stress test, such as banks' CET1 ratios. This completeness allows for an accurate description of the severity of the desired scenario and, consequently, the stress test itself.

Our framework facilitates the conduct of distributional stress tests, reverse stress tests, and the design of stress test scenarios for various other exercises. We provide examples illustrating how multiple scenario distributional stress testing can yield various at-risk measures, differentiate between idiosyncratic institution-specific and system-wide systemic risks, and comprehensively depict the evolution of risks within the banking system. Our reverse stress test bypasses the need to invert complex models of banks and economies. The examples we provide demonstrate how the search and selection of sufficiently adverse scenarios within a scenario space can yield scenarios tailored to meet supervisory, macroprudential, or broader policy objectives.

Although our primary contribution is methodological, we also contrast some of our results with those of the EBA/SSM stress test in 2023. We confirm certain intuitions about the exercise. The communication of the EU-wide stress test usually underlies testing individual banks' solvency. However, its design leads to outcomes better interpreted as testing individual banks' solvency in the face of system-wide solvency strains. The flip side of this finding is that the exercise may lack the ability to assess bank idiosyncratic risks or risks specific to smaller banks. Additionally, while the EBA/SSM stress test aims to challenge bank solvency, its macro-financial scenarios appear well crafted for assessing the banking system's ability to provide lending to the real economy, akin to the ECB macroprudential stress test.

We add new arguments to discussing the optimal design of stress-testing frameworks. We contribute to the literature advocating supplementing conventional single-scenario stress testing with multiple-scenario approaches. Multiple scenario designs can help validate and produce single-scenario results and accurately identify the distribution of risks within the banking system between individual institutions. Importantly, they have become increasingly achievable with the pronounced evolution of stress test infrastructures in recent years.

## References

- V. V. Acharya, L. H. Pedersen, T. Philippon, and M. Richardson. Measuring Systemic Risk. *The Review of Financial Studies*, 30(1):2–47, 10 2017.
- T. Adrian and M. K. Brunnermeier. Covar. *American Economic Review*, 106(7):1705–41, July 2016.
- T. Adrian, N. Boyarchenko, and D. Giannone. Vulnerable growth. *American Economic Review*, 109(4): 1263–89, April 2019.
- P. Baudino, R. Goetschmann, J. Henry, K. Taniguchi, and W. Zhu. Stress-testing banks – a comparative analysis. FSI Insights on policy implementation No 12, Bank for International Settlements, 2018.
- C. Borio, M. Drehmann, and K. Tsatsaronis. Stress-testing macro stress testing: Does it live up to expectations? *Journal of Financial Stability*, 12(C):3–15, 2014.
- T. Breuer and I. Csiszár. Systematic stress tests with entropic plausibility constraints. *Journal of Banking Finance*, 37(5):1552–1559, 2013.
- T. Breuer and I. Csiszár. Measuring distribution model risk. *Mathematical Finance*, 26(2):395–411, 2016.
- T. Breuer and M. Summer. Systematic Systemic Stress Tests. Working Papers 225, Oesterreichische Nationalbank (Austrian Central Bank), Dec. 2018.
- T. Breuer, M. Jandacka, K. Rheinberger, and M. Summer. How to Find Plausible, Severe and Useful Stress Scenarios. *International Journal of Central Banking*, 5(3):205–224, September 2009.
- K. Budnik. A bird’s-eye view of the resilience of the European banking system: results from the new macroprudential stress test framework. *Macroprudential Bulletin*, 7, 2019. URL 1.
- K. Budnik, M. Balatti, I. Dimitrov, J. Groß, I. Hansen, M. Kleemann, F. Sanna, A. Sarychev, N. Sinienko, M. Volk, G. Covi, and di Iasi. Macroprudential stress test of the euro area banking system. Occasional Paper Series 226, European Central Bank, July 2019.
- K. Budnik, L. Boucherie, M. Borsuk, I. Dimitrov, G. Giraldo, J. Groß, M. Jančoková, J. Karmelavičius, M. Lampe, G. Vagliano, and M. Volk. Macroprudential stress test of the euro area banking system amid the coronavirus (COVID-19) pandemic. Report, European Central Bank, October 2021a.



- K. Budnik, I. Dimitrov, C. Giglio, J. Groß, M. Lampe, A. Sarychev, M. Tarbé, G. Vagliano, and M. Volk. The growth-at-risk perspective on the system-wide impact of Basel III finalisation in the euro area. Occasional Paper Series 258, European Central Bank, July 2021b.
- K. Budnik, I. Dimitrov, J. Groß, M. Jančoková, M. Lampe, B. Sorvillo, A. Stular, and M. Volk. Policies in support of lending following the coronavirus (COVID 19) pandemic. Occasional Paper Series 257, European Central Bank, May 2021c.
- K. Budnik, I. Dimitrov, J. Groß, M. Lampe, and M. Volk. Macroeconomic impact of Basel III finalisation on the euro area. *Macprudential Bulletin*, 14, 2021d.
- K. Budnik, L. Boucherie, and J. Panos. Looking at the evolution of macroprudential policy stance: A growth-at-risk experiment with a semi-structural model. Occasional Paper Series 301, European Central Bank, Sept. 2022.
- K. Budnik, J. Gross, G. Vagliano, I. Dimitrov, M. Lampe, J. Panos, S. Velasco, L. Boucherie, and M. Jančokova. BEAST: A model for the assessment of system-wide risks and macroprudential policies. Working Paper Series 2855, European Central Bank, October 2023.
- M. Cihak. Stress Testing: A Review of Key Concepts. Research and Policy Notes 2004/02, Czech National Bank, Apr. 2004.
- R. Cont, A. Kotlicki, and L. Valderrama. Liquidity at risk: Joint stress testing of solvency and liquidity. *Journal of Banking Finance*, 118:105871, 2020. ISSN 0378-4266.
- EBA. 2021 EU-wide stress test methodological note. Technical report, EBA, January 2021.
- ESRB. Macro-financial scenario for the 2021 eu-wide banking sector stress test, January 2021.
- M. D. Flood and G. G. Korenko. Systematic scenario selection: stress testing and the nature of uncertainty. *Quantitative Finance*, 15(1):43–59, January 2015.
- P. Glasserman, C. Kang, and W. Kang. Stress scenario selection by empirical likelihood. *Quantitative Finance*, 15(1):25–41, January 2015.
- P. Grundke and K. Pliszka. A macroeconomic reverse stress test. *Review of Quantitative Finance and Accounting*, 50(4):1093–1130, 2018.

- J. Henry. Reflections on macroprudential reverse stress testing. In *Reverse Stress Testing in Banking*, chapter Chapter 8. De Gruyter, May 2021.
- X. Huang, H. Zhou, and H. Zhu. A framework for assessing the systemic risk of major financial institutions. *Journal of Banking Finance*, 33(11):2036–2049, 2009.
- P. Kapinos and O. A. Mitnik. A Top-down Approach to Stress-testing Banks. *Journal of Financial Services Research*, 49(2):229–264, 2016.
- Y. Kopeliovich, A. Novosyolov, D. Satchkov, and B. Schachter. Robust risk estimation and hedging: A reverse stress testing approach. *The Journal of Derivatives*, 22(4):717–728, May 2015.
- A. J. McNeil and A. D. Smith. Multivariate stress scenarios and solvency. *Insurance: Mathematics and Economics*, 50(3):299–308, 2012.
- K. Pliszka. System-wide and banks’ internal stress tests: Regulatory requirements and literature review. Discussion Paper No 19, Deutsche Bundesbank, 2021.
- M. Quagliariello, editor. *Stress-testing the Banking System Methodologies and Applications*. Cambridge University Press, 2009.
- A. Sarychev. Scenario generation vs. forecasting: predictive performance criteria and the role of vague priors, March 2014.
- T. Schuermann. Stress testing banks. *International Journal of Forecasting*, 30(3):717–728, 2014.
- G. Studer. Market risk computation for nonlinear portfolios. *Journal of Risk*, 1(4):33–53, 1999.
- P. Traccucci, L. Dumontier, G. Garchery, and B. Jacot. A Triptych Approach for Reverse Stress Testing of Complex Portfolios. Papers 1906.11186, arXiv.org, June 2019.

## A Appendix: BEAST high-level description

The complete description of the model structure and estimates can be found in Budnik et al. [2023]. This appendix briefly discusses the main mechanisms of the model and its elements, which play the most crucial role in stochastic simulations.

### A.1 Macroeconomic block

The dynamics of individual euro area economies are mapped by a set of equations derived from the country-level structural vector autoregressive model (SVAR). Each country, denoted as  $C$ , is characterised by the following set of equations:

$$\begin{aligned}
 Y_t^C &= a_Y^C + \sum_L A_{Y,L}^{Y,C} Y_{t-L}^C + \sum_L B_L^{Y,C} M_{t-L}^{EA} + \sum_L E_L^{Y,C} X_{t-L}^C + \sum_L F_L^{Y,C} Z_{t-L}^C + v_t^{Y,C,+} \\
 M_t^C &= a_M^C + \sum_L A_{M,L}^{Y,C} Y_{t-L}^C + \sum_L B_L^{M,C} M_{t-L}^{EA} + \sum_L E_L^{M,C} X_{t-L}^C + \sum_L F_L^{M,C} Z_{t-L}^C + v_t^{M,C,+} \\
 M_t^{EA} &= \left( \sum_{C \in EA} w^C \times M_t^C > M^* \right) . + M^*
 \end{aligned} \tag{18}$$

where  $Y_t^C$  is a vector of 10 country-specific variables, including real GDP, HICP, unemployment rate, the spread between 10-year government bond yield and 3-month EURIBOR, import volume, export price, residential property price, bank loan volumes and lending rate for the non-financial private sector, and equity price index. Vectors  $M_t^C$  and  $M_t^{EA}$  each include the EURIBOR 3-month rate and the assets of the Eurosystem as a measure of unconventional monetary policy. Vector  $X_t^C$  represents country-specific measures of foreign demand and competitors' export prices, and  $Z_t^C$  includes an additional set of variables, including dummies related to the episode of the COVID-19 pandemic and the index of energy prices.  $M_t^*$  imposes a floor level of -1.5% on the 3-month EURIBOR, and  $w^C$  represents the nominal GDP share of each country in the nominal GDP of the euro area (in 2021). Vector  $v_t^{C,+} = [v_t^{Y,C,+}, v_t^{M,C,+}]'$  relates to the vector of reduced-form residuals from the estimation of country-level SVARs,  $v_t^C$ . The latter are assumed to be independent and identically distributed with a mean of zero and a constant covariance matrix. The number of lags,  $L$ , is set to 2, and  $t$  represents the period. Vectors  $a$  and matrices  $A$ ,  $B$ ,  $E$ , and  $F$  contain the model's estimated coefficients.

Equations of a reduced-form VAR represent individual countries in the rest of the world segment:

$$\tilde{Y}_t^C = \tilde{a}_{\tilde{Y}}^C + \sum_L \tilde{A}_L^C \tilde{Y}_{t-L}^C + \tilde{v}_t^C \quad (19)$$

with  $\tilde{v}^C$  representing a  $K^C$ -dimensional white noise process characterised by a time-invariant positive definite covariance matrix. Vector  $\tilde{Y}$  comprises  $K^C$  variables, including real GDP, import volumes, and export prices for each country.

Cross-country trade spillovers are introduced by representing countries' foreign demand and price variables with weighted averages of import volumes and export prices of their trading partners.

$$X_t^C = \left[ \sum_{T \in \{EA \setminus C, RoW\}} w_{FDR}^T \times MTR_t^T, \sum_{T \in \{EA \setminus C, RoW\}} w_{CXD}^T \times XTD_t^T \right]' \quad (20)$$

where  $w_{FDR}^T$  represents the share of exports from country  $C$  to country  $T$  in the total exports of country  $C$  and  $MTR$  are country import volumes, and  $w_{CXD}^T$  represents the share of imports from country  $T$  in the total imports of country  $C$  and  $XTD$  are country export prices.

## A.2 Bank balance sheet and profit and loss accounts

Bank assets consist of banking and trading books holdings, with interest-bearing securities that can be classified in either book (see Table 2). In the banking book, the model tracks loan exposures to the non-financial corporate sector (*NFC*), household loans backed by real estate (*HHHP*), and household credit for consumption purposes (*HHCC*) that can exhibit significantly different dynamics depending on the geographical location. It also considers exposures to sovereigns (*SOV*), the financial sector (*FIN*), and central banks (*CB*), where banks can adjust the amounts of asset holdings without changing their geographical composition.

On the liability side, a bank's balance sheet comprises equity, sight, and term deposits from corporates (*NFC*) and households (*HH*), secured funding through repos, issued collateralised debt securities, and unsecured wholesale funding, including inter-bank liabilities and debt securities. Banks can adjust private sector deposits separately for different geographical regions while the geographical composition of other liabilities remains constant. Finally, banks can adjust their capitalisation through profit retention, although they cannot be recapitalised or issue new shares.

Bank net profits consider impairments resulting from credit risk, net interest income, asset reval-

Assets	Liabilities
Loans NFC	Capital
Loans HHP	
Loans HHCC	Sight deposits HH
Loans FIN	Sight deposits NFC
Loans CB	Term deposits
Loans OTHER	Deposits CB
	Deposits SOV
Equity exposures	Repo
Securitized portfolio	Debt securities (secured)
	Debt securities (unsecured)
Securities SOV	
Securities NFC	Wholesale funding (unsecured)
Securities FIN	
Trading assets	

**Table 2:** Schematic illustration of bank’s balance sheet

uation, and net trading and fee-commission income. Within the model, the flows between the three IFRS9 asset impairment stages – performing, with increased credit risk since initial recognition, and credit impaired – are monitored for each distinct banking book portfolio. Changes in asset quality affect the corresponding loan loss provisions, which, when aggregated, are included in the profit and loss statement.

Each banking book portfolio has its assigned credit risk weight based on an internal model-based approach (IRB) or a standardised approach (STA). Total risk-weighted amounts combine credit risk charges on exposures in the banking book with capital charges associated with market and operational risk. These risk-weighted amounts are the denominator for calculating the Common Equity Tier 1 (CET1) capital ratio.

### **A.3 Bank behavior**

Banks in the model operate in monopolistic competition in lending markets while acting as price-takers in funding markets. Consequently, they can discriminate between different lending markets in which they operate and face a downward-sloping demand curve in each. Loan demand primarily depends on macro-financial variables, including the business cycle, GDP, unemployment, inflation dynamics, and market interest rates. On the other hand, the supply of loans reflects the specific circumstances of each bank, such as its solvency, leverage, profitability, asset quality, and funding costs.

The structure of banks’ debt funding depends on the pecking order principle. In the first line, banks

finance newly issued assets or replace maturing liabilities with retail, sovereign, and central bank funding, which are relatively low cost but in limited supply. Banks can access the generally higher-cost wholesale market if these funding sources are insufficient. There, a bank can secure funding at close to risk-free interest rates by posting collateral or issuing unsecured debt, which carries an additional credit spread.

Banks' use of wholesale funding introduces the funding-solvency feedback loop. Deteriorating bank solvency leads to a higher credit spread on unsecured wholesale funding and an increased reliance on wholesale funding overall. The resulting higher average cost of funding reduces bank profitability and can further deplete its existing capital. A solvency-funding costs feedback loop is particularly prone to emerge during adverse macroeconomic conditions and with already elevated risk margins in wholesale markets.

Profit retention follows a straightforward rule: a bank distributes profits as long as it can maintain its internal target capital ratio. This internal target combines regulatory requirements, buffers, and an additional bank management buffer reflecting the bank's business model and balance sheet characteristics.

Banks are subject to capital requirements, buffers, and liquidity regulation. By adjusting their lending and dividend payout policies, they control the distance between actual CET1 and leverage ratios and their regulatory thresholds. There is an essential non-linearity in banks' responses to a capital shortfall versus surplus compared to their capital requirements and buffers. A bank facing a CET1 capital shortfall reduces its lending to the non-financial private sector by a higher absolute amount than a bank with the same magnitude surplus increases. The deviation from the liquidity coverage ratio (LCR) requirements and the net stable funding ratio (NSFR) affects the composition of bank wholesale funding.

#### **A.4 Closing the model**

In closing the model, bank lending decisions impact the real economy. The feedback loop takes the form of an additional credit supply shock derived by aggregating at the country level a measure of bank excessive deleveraging ( $LoansupplyInnov_t^C$ ). The measure of excessive deleveraging equals the non-linear part of bank lending response resulting from financial and regulatory strains, therein from the capital shortfall to regulatory thresholds.<sup>25 26</sup>

Then:

---

<sup>25</sup>Another element of banks' non-linear response can be triggered by their earlier high risk-taking in market segments that turned out to be characterised by high default rates.

<sup>26</sup>The model can operate with two alternative feedback loops. The first feedback loop is used throughout this paper. The second feedback loop supersedes the dynamics of country-level lending volumes and interest rates with their aggregated bank-level counterparts. Comparative exercises carried out using the two feedback loops have shown that the choice of one of them does not significantly affect the presented results.

$$v_t^{C,+} = v_t^C + e_1 \circ \frac{1}{d^C} \times LoansupplyInnov_t^C \quad (21)$$

where  $d^C$  is the first element of the matrix  $D^C$  that provides the mapping between the reduced-form residuals and orthogonal structural shocks  $\varepsilon_t^C \sim \mathcal{N}(0, 1)$  ( $v_t^C = D^C \varepsilon_t^C$ ) and where the credit supply shock is the structural shock ordered first.<sup>27</sup>  $e_i$  is a vector with all zero elements except unit  $i$ -th element.  $LoansupplyInnov_t^C$  is defined in a way that  $\frac{1}{d^C} \times LoansupplyInnov_t^C$  implies an instantaneous change in loan volumes 1%.

Finally, the model equations are stacked in one system and solved simultaneously and sequentially for each forecast horizon period.

## A.5 Scenario and parameter uncertainty

The VAR equations describing the evolution of euro area economies are estimated in a Bayesian panel setup. The estimates provide the set of historical residuals  $v_t$  and posterior estimates of the model parameters in the matrices  $a$ ,  $A$ ,  $B$ ,  $E$  and  $F$  (see equation 18). Analogous information is available for VARs representing non-euro area economies with the recognition that those are estimated individually with a multivariate least squares estimator. Additionally, three energy price indices entering euro area VARs, representing oil, coal and gas prices, are estimated in a multivariate GARCH, delivering the corresponding residuals and model parameters estimates.<sup>28</sup>

Bank behavioural equations related to their lending, funding choices, and the setting of management buffer are usually estimated in a fixed-effects panel framework with frequentist methods. These estimates provide information on the distributions of the corresponding regression coefficients.

Scenario uncertainty is mapped by Monte Carlo (MC) stochastic simulations of  $v$ ,  $\tilde{v}$  and energy GARCH residuals. Two alternative sampling schemes can be applied: one employs the distributional assumptions from the estimation stage, and the other relies on bootstrapping methods. The former sampling scheme employs the estimates of shock variance-covariance matrices and imposes zero correlation assumption between shocks in the euro area, non-euro area countries and energy prices. It also imposes zero correlation between reduced-form shocks between any two non-euro area countries. In contrast, the

<sup>27</sup>We do not discuss the structural representation of shocks to country-level SVARs because it does not directly imprint the results presented in this paper. The model identifies and constrains nine structural shocks through a combination of sign and zero restrictions.

<sup>28</sup>The sample for macroeconomic equations reaches back to 1991. Accordingly, the estimated residuals should capture at least one episode of a significant financial crisis for any country.

wild block bootstrap recognises the complete cross-correlation structure between the shocks drawn for all economies and energy prices.

Parameter uncertainty is captured by repetitively drawing parameters from the distributions of estimated model parameters. For all equations representing euro area country dynamics, the realisations of parameters are drawn from their joint posterior distributions. For non-euro area, energy prices and bank-level equations, it is assumed that the Gauss-Markov theorem holds, and the realisations of the corresponding parameters are drawn from a joint normal distribution with appropriately set means and variance-covariances. The realisations of parameters are drawn independently for any two equations that have been estimated separately.

A small fraction of simulations that lead to the explosive dynamics of the model is discarded in the process of pruning.

## **B Appendix: Bank vulnerability indicators**

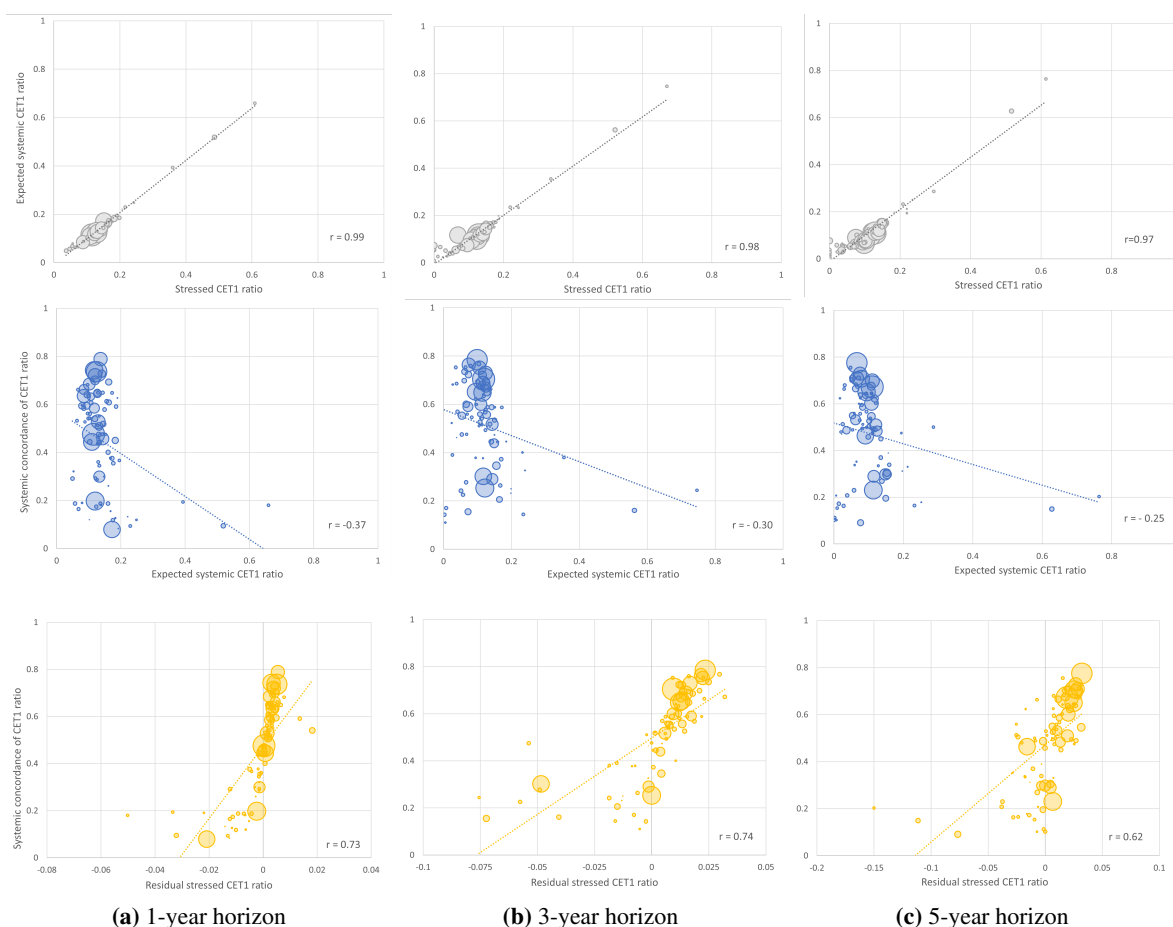
This appendix documents cross-correlations between different bank-level vulnerability indicators from the distributional stress test presented in Section 6.

The first row of Figure 16 shows a strong positive correlation between the stressed and systemic expected CET1 ratios. It partly relates to the specificity of simulations that exclude bank-specific exogenous shocks. In the simulations, the dynamic differences between the banks arise due to the heterogeneous structure of their balance sheets. The presence of additional idiosyncratic shocks could weaken the observed correlation.

The second row of Figure 16 documents the negative correlation between the systemic expected CET1 ratios and the systemic concordance for the CET1 ratios. The attractive aspect is the large dispersion in the systemic concordance, holding both for smaller and most significant banks, for similar levels of banks' systemic expected CET1 ratio. It demonstrates significant variation in the relative relevance of systemic events for individual banks. Banks with a low level of systemic concordance are likely to fare comparatively well in the case of systemic events (or in the EU-wide stress test exercises). However, it is only falsely reassuring because other plausible future scenarios can still materially shake their resilience.

The last row of Figure 16 combines the information from all three indicators, mapping the relationship between the difference between the stressed versus systemic expected CET1 ratios and the systemic concordance of the CET1 ratios. It delivers an attractive interpretation of the systemic concordance for stress test applications. The systemic concordance can indicate benefits from testing banks with addi-





**Figure 16:** The relationship between different bank vulnerability measures at three horizons: 1-year, 3-years and 5-years

Notes: Stressed, systemic expected CET1 ratios and systemic concordance of CET1 ratios for  $10^{th}$  percentiles of respective distributions. The residual stressed CET1 ratio is the difference between a bank's stressed and systemic expected CET1 ratio. The size of balls corresponds to the size of banks' assets at the end of 2022. Lines correspond to the fitted linear regression, and  $r$  is the correlation coefficient between the (unweighted) variables.

tional bank-specific stress scenarios, for instance, via the lenses of distributional stress test, on the top and beyond scenarios reflecting system-wide strains. The benefits from additional scrutiny would be highest, especially for banks to the left of the regression line.

## C Appendix: Trade-offs in reverse stress testing

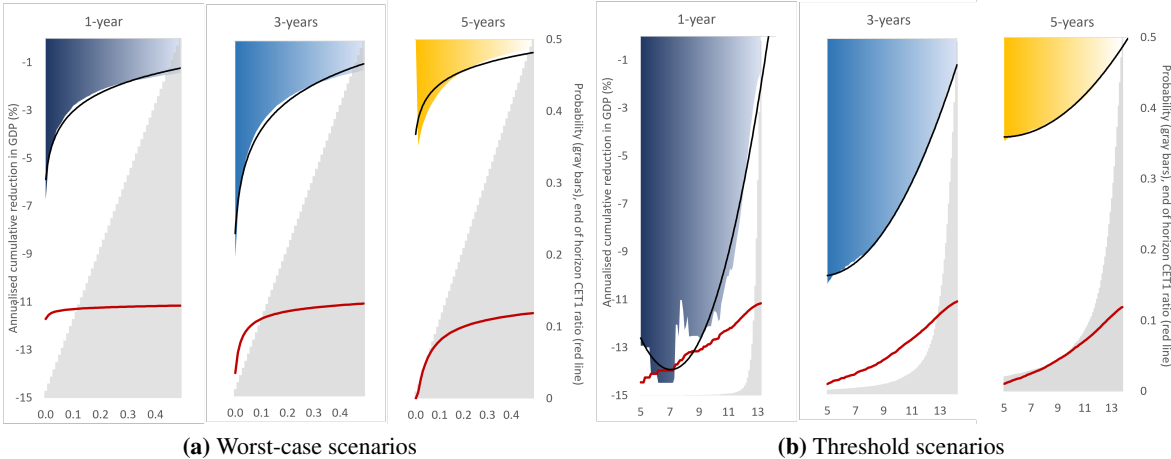
This appendix compares the implementation challenges of the worst-case scenarios with those of threshold scenarios. Furthermore, it discusses the non-linearities in the pass-through of macro-financial scenarios into outcome variables.

Figure 17 summarises the result of progressively identifying severe scenarios by applying the worst-

case and threshold approach. For each reverse stress test scenario, we plot its probability of realisation (grey bars and the right axis), the resulting median system-wide CET1 ratio (red line and the right axis) and the cumulative drop in the euro area GDP compared to the last quarter of 2022 on annualised terms (shaded fields). For worst-case scenarios, we consider realisation probabilities from 0 to 50%, while for threshold scenarios, CET1 ratios from 5% to around 14%.<sup>29</sup>

The threshold approach offers greater flexibility than worst-case scenarios to arrive at very adverse outcomes, especially for short-term horizons. For both approaches, the plausibility of arriving at very adverse scenarios increases with the horizon of the scenario, along with the increasing uncertainty of the CET1 ratio. The flip side of this statistical fact is that scenarios with a sharp drop in the CET1 ratio in a short horizon are relatively unlikely. Even though the threshold approach appears handy in such situations, excessively conservative thresholds compromise scenario plausibility, potentially harming its meaningful interpretation.

A conclusion is that threshold scenarios, or stress tests designed to achieve very low CET1 ratios, should be handled carefully. On the other hand, multiple scenario stress tests that offer the possibility of inspecting the distribution of the outcome variable along the spectrum of plausible futures can become helpful in anchoring policymakers' expectations.



**Figure 17:** Trade-offs between plausibility, severity and horizon for scenarios targeting CET1 ratio

Notes: Primary OY axis shows the cumulative change in the euro area GDP at the end of the scenario horizon compared to the end of 2022 in the reverse stress test. The secondary OX axis shows the probability of the realisation of the corresponding scenario, the level of the median CET1 ratio at the end of the horizon. LHS: The OX-axis marks increasing probability levels. Black lines are the fitted logarithmic regression lines. The logarithmic specification explains 96% of GDP variation in the 1-year horizon (versus 69% for linear specification), 98% in the 3-year horizon (versus 71%), and 58% in the 5-year horizon (versus 68%). RHS: OX-axis marks increasing CET1 ratio thresholds. Black lines are the fitted second-degree polynomial regression lines. The polynomial specification explains 93% of GDP variation in the 1-year horizon (versus 73% for linear specification), and over 99% in the 3-year and 5-year horizons (versus 94%).

<sup>29</sup>The most lenient threshold for each horizon is set so that the probability of the resulting scenario is 50%.

Another conclusion from inspecting the results of reverse stress testing is that the pass-through of macro-financial variables, or risk factors, into bank variables must not be linear. In discussing scenario severity, the salient rule of thumb is that to increase the severity of a stress test, one must provide more adverse macroeconomic conditions. Especially in the approach in which macro-financial scenario design is decoupled from the assessment of the target variable, it is often the only yardstick available. Its reflections are present in, e.g., Henry [2021], who scales macro-financial risk factors with a set of scalars to arrive at a grid of scenarios with different severity.

The marginal impact of macro-financial outlook on the banking sector is low for permissive solvency stress tests. In the range for the CET1 ratio between around 11% and 13% in the 3-year horizon, marginal changes in the CET1 ratio demand measurably worse GDP results. This can be seen in the left panel of Figure 17 when comparing the relative flatness of the CET1 ratio line in this range, with the steepness of the fitted curve for the drop in the GDP level. Analogously, within this rate, the CET1 ratio line is also less steep than the GDP level curve.

For more conservative stress tests, changes in the adversity of the macro-financial outlook have substantial impact both on scenario severity and plausibility. From the level of the CET1 ratio below 11%, marginal changes in CET1 can be generated in moderately worse macro-financial developments. Following this break-even point even slightly worse macro-financial scenario triggers substantial further deterioration in bank solvency and scenario plausibility. This partially reflects the appearance of amplification mechanisms in this part of the distributions.

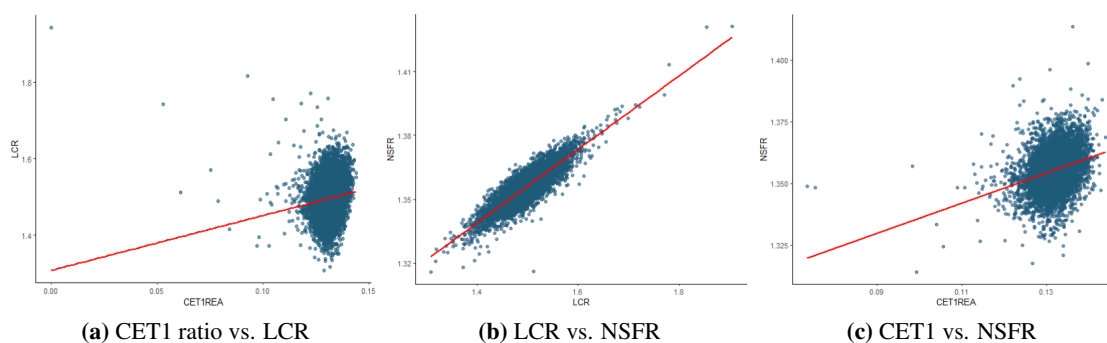
Overall, designing macro-financial scenario around the break-even point can lead to surprises. The non-linearity in the pass-through means that moderate adjustments of the macro-financial scenarios can trigger substantially worse bank outcomes, or substantial deterioration of macro-financial conditions can have disappointing effect on bank results.

## **D Appendix: Joint solvency and liquidity stress testing**

This appendix expands on the ideas discussed in Chapter 8 by illustrating the case of joint liquidity and solvency stress testing. An important disclaimer is that although the model we use includes a comprehensive representation of bank liquidity, the quarterly frequency of the model and relatively stronger focus on the real economy than on financial markets make it less suitable to study liquidity as compared to solvency risks.

## D.1 Bank liquidity and solvency trade-offs

In the one-year horizon, which is most significant for the stress testing LCR, the correlation between the system-wide CET1 ratio and the LCR is positive. As illustrated in Figure 18. There is also a positive correlation between LCR and NSFR.<sup>30</sup>



**Figure 18:** Correlation of system-wide CET1 ratio, LCR and NSFR at 1 year horizon

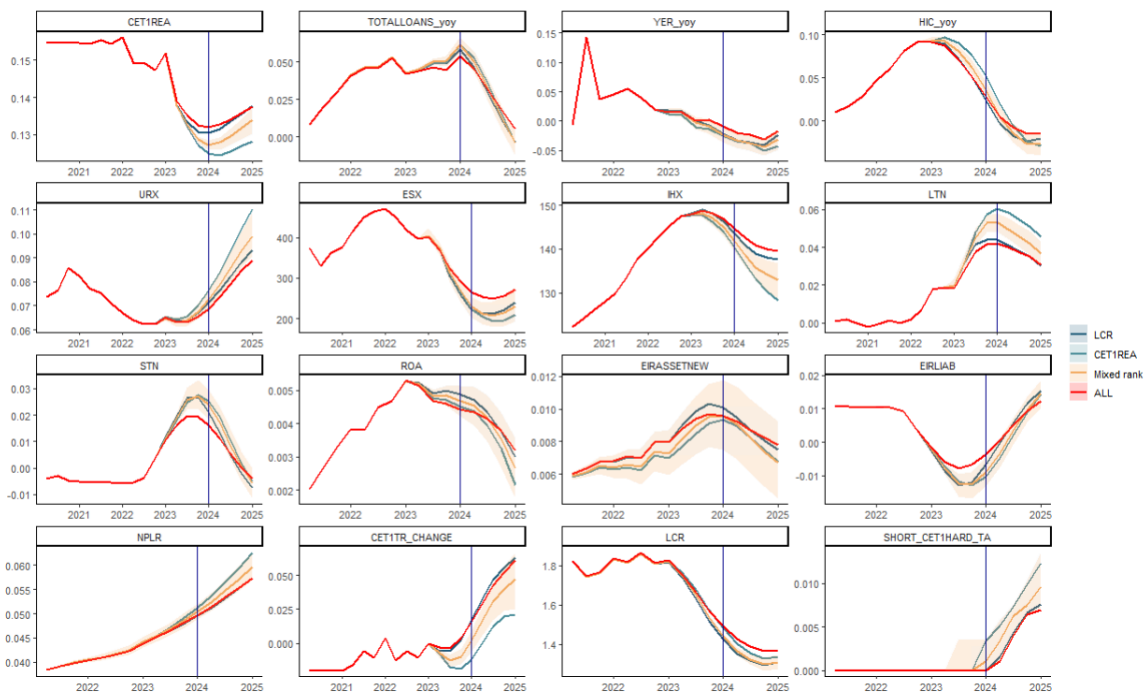
Notes: Observations in the graphical representation are filtered using Cook's distance for better readability.

## D.2 One-year ahead worst-case joint solvency and liquidity scenarios

Figure 19 plots stress scenarios for the CET1 ratio and LCR considered jointly and with equal weights. It contrasts them with scenarios with the same probability 10%, but focusing on bank solvency or liquidity at a time.

Liquidity tensions can emerge even in a solvent banking system along with a modest economic slowdown, as long as they are accompanied by a relatively high level of real interest rates. In comparative terms, solvency stress involves a more significant deterioration in asset prices and unemployment. The combined stress test hits the middle field, placing similar strain on bank maturity mismatches as the solo liquidity stress test, but involving lower capital losses than the stress test designed to evaluate banks' CET1 ratios.

<sup>30</sup>In the longer horizons, the correlation between the CET1 ratio and LCR gradually turns negative, while remaining positive for the CET1 ratio and NSFR. This is an intuitive result closely related to the definitional similarities between the CET1 ratio and the NSFR.



**Figure 19:** Worst-case scenarios (10<sup>th</sup> percentile) for bank solvency and liquidity in 1-year horizon

Notes: CET1REA – CET1 ratio, TOTALLOANS\_yoy – annual bank lending volumes to the non-financial private sector growth rate, YER\_yoy – annual GDP growth rate, HIC\_yoy – HICP inflation, URX – unemployment rate, ESX – equity price index, IHX – house price index, LTN – 10-year bond yields, STN – 3-month EURIBOR, ROA – return on assets, ERASSETNEW – interest rates on new lending to the non-financial private sector, EIRLIAB - the average cost of debt funding, NPLR – NPL ratio, CET1TR\_CHANGE - the cumulative change in CET1 capital, LCR - liquidity coverage ratio, NSFR - net stable funding ratio. Red line: median for the full distribution of events, dark blue line: median of worst-case scenarios for LCR, light blue line: median of worst-case scenarios for CET1 ratio, yellow: median of worst-case scenarios for CET1 and LCR jointly. Fancharts mark 40% probability bands for the worst-case scenarios for CET1 and LCR jointly.

## Acknowledgements

The authors would like to thank Ivan Dimitrov for his continuous advice on the various steps of the project.

### David Aikman

King's College London, London, United Kingdom; email: [david.aikman@kcl.ac.uk](mailto:david.aikman@kcl.ac.uk)

### Romain Angotti

Sciences Po, Paris, France; European Central Bank, Frankfurt am Main, Germany; email: [romain.angotti@ecb.europa.eu](mailto:romain.angotti@ecb.europa.eu)

### Katarzyna Budnik (Corresponding author)

European Central Bank, Frankfurt am Main, Germany; email: [Katarzyna.Budnik@ecb.europa.eu](mailto:Katarzyna.Budnik@ecb.europa.eu)

## © European Central Bank, 2024

Postal address 60640 Frankfurt am Main, Germany

Telephone +49 69 1344 0

Website [www.ecb.europa.eu](http://www.ecb.europa.eu)

All rights reserved. Any reproduction, publication and reprint in the form of a different publication, whether printed or produced electronically, in whole or in part, is permitted only with the explicit written authorisation of the ECB or the authors.

This paper can be downloaded without charge from [www.ecb.europa.eu](http://www.ecb.europa.eu), from the [Social Science Research Network electronic library](#) or from [RePEc: Research Papers in Economics](#). Information on all of the papers published in the ECB Working Paper Series can be found on the [ECB's website](#).

PDF

ISBN 978-92-899-6689-4

ISSN 1725-2806

doi:10.2866/641657

QB-AR-24-058-EN-N