DNB Working Paper

No 860/April 2026

Identifying scenarios of interest under deep uncertainty

Jan Kwakkel, Willem L. Auping and Jan Willem van den End

DeNederlandscheBank

EUROSYSTEEM

Identifying scenarios of interest under deep uncertainty
Jan Kwakkel, Willem L. Auping and Jan Willem van den End*

* Views expressed are those of the author and do not necessarily reflect official positions of De Nederlandsche Bank.

Working Paper No. 860

April 2026

De Nederlandsche Bank NV
P.O. Box 98
1000 AB AMSTERDAM
The Netherlands

Identifying scenarios of interest under deep uncertainty

Jan Kwakkel[a], Willem L. Auping[a], Jan Willem van den End[b,c*]

[a] Delft University of Technology, the Netherlands
[b] De Nederlandsche Bank, the Netherlands
[c] Vrije Universiteit Amsterdam, the Netherlands

Abstract

We introduce Exploratory Modelling and Analysis (EMA) as a framework for macroeconomic scenario analysis under deep uncertainty. EMA traces scenarios of interest from policy-relevant outcomes. These outcomes are forecasted by a model, based on a large set of uncertain conditioning variables. This approach flips the conventional method of scenario analysis, by sampling many scenarios without imposing strong ex-ante priors on the choice, combination, value, or distribution of the conditioning variables. Applying EMA with an interacted VAR model to financial stress scenarios for the euro area shows that scenarios of interest can be uncovered that may be overlooked by common approaches for scenario analysis.

JEL classification: E52, E58, G12
Keywords: Financial shocks, Scenarios, Exploratory modelling, Deep uncertainty

---

* Corresponding author. Email address: w.a.van.den.end@dnb.nl
The authors wish to thank Guido Ascari, Maurice Bun, Maria Sole Pagliari, Sebastiaan Rast and participants of the VU research seminar (March 2026), the DNB research seminar (March 2026) and members of the ESCB export group on scenarios (April 2026) for useful input. Views expressed are those of the authors and do not necessarily reflect official positions of De Nederlandsche Bank or the Eurosystem.

1. Introduction

Scenario analysis is an important instrument in the toolkit of central banks and other policymakers (BIS, 2025). It allows them to assess uncertainties related to alternative hypothetical events or narratives, thereby guiding policy decisions that are robust to uncertain future states of the world. In monetary policy, which relies on macroeconomic forecasts, scenarios are particularly useful because they provide structured ways of thinking about possible futures that deviate from the baseline macroeconomic projection. These deviations may arise from shocks to inflation and output that have implications for the monetary policy stance and transmission.

A main challenge for scenario analysis is that an unlimited set of possible futures is conceivable, while only a subset may be relevant for policy. A systematic approach to scenario selection and design is therefore needed. The ECB (2025) proposes that scenarios should be relevant in terms of their potential impact, likelihood and implications for monetary policy. Scenarios with high potential impact, even if their likelihood is ex-ante perceived to be very low or uncertain, might especially be important, as they carry the risk of significant policy mistakes if ignored. However, future outcomes that are closer to the baseline - and therefore considered more likely - may appear to be more obvious candidates for the risk assessment surrounding the baseline forecast. This tends to exclude extreme tail risks, since such events may be unprecedented and not captured in historical data on which forecasting models are estimated.

Recent developments in central banking highlight the growing importance of scenario analysis. Bernanke (2024) recommended to de-emphasize the baseline forecast and adding alternative scenarios to assess risks arising to unexpected shocks to the economy. This approach helps the public understand policy choices in light of uncertainty and underscores the role of risk management in monetary decision-making (see for instance Laxton et al., 2025). The ECB (2025) highlights the growing importance of scenario analysis in its latest monetary policy strategy review. Over the last years, the ECB has expanded its scenario analysis, incorporating narrative scenarios, sensitivity analyses, and quantitative risk assessments into its macroeconomic projection exercise.

This reflects a broader recognition that traditional forecasting methods, while useful, are insufficient in environments characterized by fundamental uncertainty; situations where the knowledge for assigning probabilities is insufficient. In such contexts, scenario analysis provides a structured yet flexible means of exploring "what if" questions. Typically, scenario analyses begin by selecting a set of topical and plausible scenarios, along with the shock variables that define them. However, this design process inherently restricts the range of uncertain variables and their possible combinations. Such dimensionality reduction can cause surprising developments to be overlooked, leading to potentially relevant scenarios being missed.

This motivates our research question: how can the policy-relevant scenarios be identified in a systematic way, taking into account the fundamental uncertainty about future states of the world? We address this question by applying Exploratory Modelling and Analysis (EMA) as a method to identify scenarios of interest for economic and monetary policy. Our approach fits in the literature on macro-economic scenario analysis developed by central banks. EMA flips the way central banks usually conduct scenario analysis, by sampling a large set of uncertain scenarios without imposing strong ex-ante priors on the choice, combination, value or distribution of the uncertainties that comprise the scenarios. In our application of EMA, the uncertainties are the shock, or conditioning variables, of which many values and combinations are sampled. This creates an uncertainty space with a host of possible future scenarios, without assigning probabilities to them. Based on the sampled uncertainties, a space of outcomes is forecasted by a model. So the model maps the uncertainties to the outcome space. In that space, the outcomes that are relevant for policy are identified. Those outcomes are then linked back to the values and combinations of the conditioning variables in the uncertainty space. These values and combinations of the uncertainties form the scenarios of interest.

We apply EMA to euro area data based on a Interacted Vector Autoregression model with exogenous conditioning variables (I-VARX). Different specifications of the model are used to generate forecasts of inflation and GDP growth, conditional on many samples of financial variables. These are the uncertainties that – in combination - form a large set of scenarios. By linking back the policy-relevant outcomes to the uncertainty space we identify the future states of the world (scenarios) that are of interest for the policymaker.

Our application of EMA yields several general insights relevant for the design and interpretation of scenario analysis. First, we find that the selection of policy-relevant scenarios is highly model-dependent. Larger models with a higher statistical fit tend to generate less dispersed scenario outcomes. This implies that more complex and/or better-fitting models tend to underrepresent tail risks and extreme scenarios ("fallacy of statistical fit"). In addition, non-linearities play an important role in shaping the policy-relevant outcomes. This underscores the importance of using non-linear models to capture scenario dynamics.

Second, the dimensionality of the uncertainty space, in our case the number of uncertain conditioning variables allowed to vary, has important implications for coverage. Designing more specific scenarios by constraining additional variables reduces the explored uncertainty space, increasing the risk that relevant combinations of uncertainties are omitted (the "fallacy of scenario precision").

Third, we find that individual variables that appear influential for the forecasted outcomes when considered in isolation are not necessarily significant when combined with other variables in a multivariate scenario. This finding supports the use of scenarios as combinations of model input beyond sensitivity tests to account for interactions and joint dynamics across variables (Saltelli et al., 2019).

The rest of this paper is structured as follows. In section 2 we put our method in the context of related literature. Section 3 explains the EMA framework. The I-VARX model and the empirical set-up are outlined in sections 4 and 5. Section 6 presents the results. In section 7 some general implications of the results are discussed. Section 8 concludes.

2. Contribution to the literature

The literature distinguishes two main approaches for macroeconomic scenario analysis: the narrative approach and the statistical approach. Both assess risks and uncertainties around the baseline macroeconomic outlook, but differ substantially in methodology.

Narrative scenario analysis relies on structural economic models to simulate the effects of hypothetical shocks. These models capture the transmission channels through which shocks propagate, underpinning internally consistent narratives about alternative economic paths. In practice, central banks frequently use narrative approaches to explore deviations from a baseline projection, informed by expert judgement or policy-relevant hypothetical events (ECB, 2025). Because this approach embeds economic structure, it facilitates the communication of the underlying economic story and transmission channels. Narrative approaches have several limitations though. The forecasted outcomes are typically point estimates, which do not provide information about their likelihood or the distribution. Moreover, the selection of scenarios is to some extent judgmental, and the bounds of the uncertainty set are not well defined.

In contrast, statistical approaches derive risk assessments from the empirical distribution of historical data. Common tools include fan charts based on past forecast errors, sensitivity analyses for individual risk factors, and macro-at-risk frameworks that estimate the probability of adverse tail outcomes. The seminal paper by Adrian et al. (2019) uses quantile regressions to estimate the conditional distribution of GDP growth as a function of financial conditioning variables. Such methods provide a probabilistic assessment of risks (growth-at-risk) by generating density forecasts and outcomes in the tail of the distribution. Furthermore, the use of parsimonious, usually single equation models facilitates that an ensemble of models can be estimated, which mitigates model uncertainty. At the same time, the drawback is that single-equation models do not capture the system dynamics of the economy. This makes statistical methods less suitable for constructing coherent macroeconomic scenarios and narratives.

While narrative and statistical approaches each offer distinct advantages, recent work highlights the benefits of combining them in risk-assessment frameworks. Modern macro-financial risk assessment increasingly relies on such hybrid methods, to capture both the structural coherence of narrative scenarios and the probabilistic features of statistical models. Adrian et al. (2025) propose such a synthetic approach in a framework that merges scenario-based projections with macro-at-risk techniques.

Their approach is based on external inputs (taken from the Federal Reserve) for the percentiles of the baseline and reference distributions for key macroeconomic variables. These percentiles are the outcomes of macro-at-risk models and underpin the scenario mixture distribution, which is a weighted average of different scenarios. The method combines the point forecast (typically the median) of those scenarios with Monte Carlo sampling from the reference distribution to compute scenario density distributions. To ensure consistency with the baseline forecast, the sampled scenarios are adjusted using entropy tilting, which shifts the distribution toward the baseline. This method generates a diverse set of scenarios, expanding beyond the limited number typically produced through expert-driven narrative exercises. Second, it provides a backstop scenario that mitigates the incompleteness of any manually selected scenario set. Third, the resulting mixture distribution yields not only scenario paths but also likelihoods, thereby enhancing the quantitative assessment of risks.

However, the synthesis method also has notable limitations. Because it relies on externally provided percentiles and scenario outcomes, it does not incorporate simulations from an internally specified macroeconomic model, limiting its flexibility and interpretability. Moreover, the entropy-tilting procedure, by construction, pulls the distribution toward the baseline, which may dampen extreme tail outcomes and reduce sensitivity to rare but severe events. Finally, the framework does not inherently identify policy-relevant scenarios, as the mixture is driven by statistical properties rather than economic narratives or policy objectives.

Our contribution to the literature is introducing EMA for macroeconomic scenario analysis. It combines elements of the narrative and statistical approaches, by designing multivariate scenarios based on the sampling of conditioning variables. Like Adrian et al. (2019) we use financial variables as conditioning variables. The EMA approach assumes fundamental uncertainty about the underlying distributions of these variables. Moreover, we do not impose strong prior restrictions on the value, distribution and combination of the conditioning variables, by generating a large set of possible futures through sampling, treating each draw as equally likely.

The sampled combinations of conditioning variables form a set of multivariate scenarios, of which the scenarios of interest are identified. This is done by tracing back the values and combinations of conditioning variables that deliver policy-relevant outcomes. Our approach remedies to some extend the pitfall of scenario analysis to overlook surprising developments or structural discontinuities (Kwakkel and Jaxa-Rozen, 2015). This well-known drawback of existing approaches results from limiting the set of uncertain variables to a smaller number of drivers, either based on historical data, a topical narrative or the distance of the scenario outcome to the baseline. Such dimensionality reduction may exclude less obvious extreme scenarios, which however might have a high impact. The open exploration approach by EMA provides a remedy for this.

EMA has some parallels with reverse stress testing, for which macroeconomic scenarios can be used as input (see for instance Aikman et al., 2024). Central banks and supervisors apply reverse stress testing to identify the combination of shocks that would lead to the failure of the financial sector or individual financial institutions. While this shares the idea of exploring adverse conditions, its logic differs from EMA. Reverse stress testing works backwards from a predefined failure threshold, searching for shocks that would cause a breach in solvency or liquidity requirements. It therefore works with a specified and structured set of shocks, assuming that the relevant risk factors, model relationships, and shock ranges are known. By contrast, EMA is a forward, exploratory approach; it samples from deep uncertainties, without assuming that the probability distributions of the uncertain conditioning variables are known. The sampled uncertainties are then analysed to discover patterns in the outcome space, focussing on scenarios that are linked to policy-relevant outcomes. Thus, whereas reverse stress testing identifies a single or narrow set of minimum conditions for failure, EMA identifies subspaces of uncertainties that are associated with policy-relevant outcomes.

3. EMA framework

EMA has been developed as a tool to support decision-making in complex systems characterized by deep uncertainty (Bankes, 1993; Bankes et al., 2013). Its core objective is to identify scenarios that are relevant for policy, without relying on strong assumptions about the underlying distributions of uncertainties. EMA has been applied in a wide range of disciplines, including climate and flood-risk and energy and supply-chain analysis. Wieles et al. (2025) is the first application of EMA to macroeconomics. They apply EMA to a Dynamic Stochastic General Equilibrium (DSGE) model to explore scenarios of interest for the inflation and output objectives of the central bank. In their approach the model parameters are the uncertainties, including parameters that determine the shocks to the economic system.

A stylised overview of the EMA framework is provided in Figure 1. It represents an uncertainty space and outcome space, connected by a model that maps uncertainties into outcomes. Computational experiments are used to sample the uncertainty space, which is spanned by a set of deeply uncertain variables or parameters (u_i for $i = 1,...,n$. In this stylised example $n = 2$). In our application of EMA the uncertainties are the conditioning variables that determine the scenarios. A scenario is a point in the uncertainty space, comprising a combination of values of different uncertainties. The scenarios are randomly generated by Latin Hypercube sampling (McKay et al., 1979). This method divides the uncertainty space into subspaces that have the same probability of being chosen (McKay et al., 2000). The latter implies that a uniform distribution is used to generate the values of the conditioning variables, reflecting that under deep uncertainty the simulated values are equally likely. The values of the uncertainties are bounded by a range, set by the modeller based on the assumed extreme minimum and maximum values of the uncertainties.

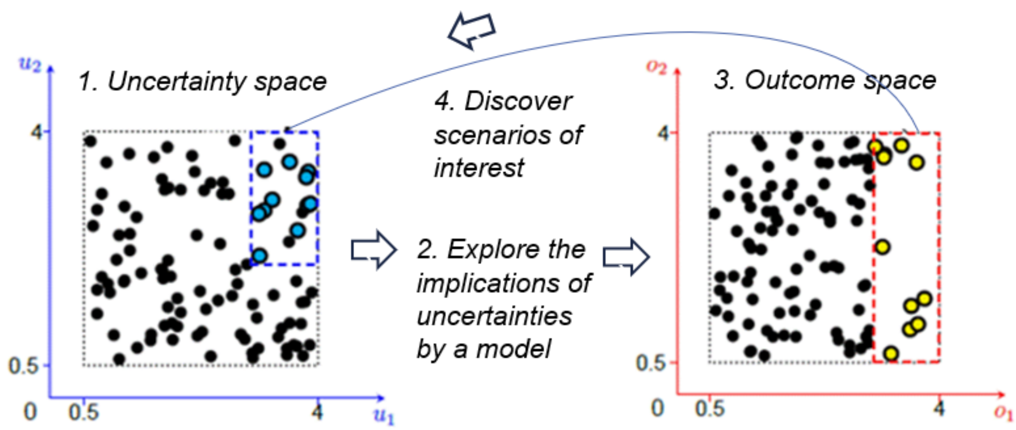

Fig 1. Uncertainty space (1) and outcome space (3) of a model with two uncertainties and two outcomes. For illustration, 100 scenarios (dots in the uncertainty space) and their corresponding model outcomes are visualized. (dots in the outcome space are the outcomes of the model (2)) The region (red, dashed rectangle with the yellow-filled dots)) captures the policy-relevant outcomes of 11 scenarios of interest (blue-filled dots in the blue, dashed area of the uncertainty space that are discovered (4)).

The implications of the sampled uncertainties are systematically explored through a model that maps the uncertainties to the outcome space. The outcomes o_i for $i = 1,...,n$ in the right-hand panel of Figure 1 are the dependent variables of the model, usually the objective variables of the policymaker (in this stylised example $n = 2$). So the model projects the effect of uncertainties into system behaviours that are captured by the outcome space (Bryant and Lempert, 2010; Kwakkel and Jaxa-Rozen, 2016).

Finally, the outcomes that are relevant for policy are linked back to the combinations of uncertainties (the scenarios) that produce them, thereby revealing the scenarios of interest. These are the scenarios in which the outcome variables exceed a certain critical value, for instance a value that conflicts with a policy objective. In Figure 1 these outcomes are marked by the red dashed rectangle in the right-hand panel.

The process of linking back the policy-relevant outcomes to their driving combinations of uncertainties (the scenarios of interest) is called behaviour-based scenario discovery. It identifies the scenarios associated with clusters of the outcome variables that have common characteristics. The intervals of the uncertainties in the uncertainty space that distinguish the outcome dynamics in each cluster are identified by the patient rule induction method (PRIM, see Friedman and Fisher, 1999 and Kwakkel et al., 2016). PRIM searches for orthogonal subspaces (boxes) within the uncertainty space that lead to values of the outcome variables that significantly differ from the average value of the outcomes.

The subspaces of the uncertainty space are spanned by intervals of the uncertainties.[1] The combination of them form a box, which covers scenarios. PRIM identifies boxes that have a high concentration (density) of scenarios of interest. In Figure 1 these scenarios are covered by the blue dashed rectangle in the left-hand panel. The intervals of the various uncertainties that determine the scenarios can be interpreted and communicated in the form of narratives. So the policymaker identifies regions in the uncertainty space that are of interest because they lead to particular outcomes, for instance outcomes that conflict with the policy objective (Bryant and Lempert, 2010; Kwakkel et al., 2013). It implies that policy-relevant outcomes are the basis of scenario discovery.

4. VAR model

Our analysis is based on an I-VARX model. VAR models are more often used for scenario analyses by central banks (BIS, 2025). The I-VARX is a reduced-form non-linear model. It offers several advantages for conducting scenario analysis under deep uncertainty. Capturing the dynamic interactions among multiple variables with minimal structural assumptions, a VAR is an appropriate framework when the underlying relationships are uncertain or unstable, as is likely the case for extreme shocks. Moreover, VAR models are flexible and largely a-

[1] PRIM describes these subspaces in the form of hyperrectangular boxes of the uncertainty space. To identify these subspaces, PRIM uses a lenient or patient, as opposed to greedy, hill climbing optimization procedure (Kwakkel and Jaxa-Rozen, 2016).

theoretical. This makes them less susceptible when structural relations shift, or when nonlinearities play a role. As a result, VARs provide a useful framework for exploring a wide range of plausible scenarios in environments characterized by deep uncertainty.

The I-VARX model is specified as a standard monetary policy VAR with inflation (π), GDP growth (y) and the money market rate (r) as endogenous variables included in vector X. The conditioning variables in vector Z are the uncertainties which determine the scenarios. These are assumed to be exogenous variables in the model. We simulate the impact of financial stress scenarios on the economy and so vector Z includes financial variables. Shocks to financial variables are associated with a negative demand shock (Benguria and Taylor, 2020), as they tighten financial conditions and so lower output growth and inflation. The I-VARX model is specified as,

$$X_t = A_0 + \sum_{p=1}^{p} B_p \, X_{t-p} + \sum_{n=0}^{n} \Gamma_n \, Z_{t-n} + \sum_{s=1}^{s} \Theta_s \, r_{t-s} \, Z_{t-s} + U_t \qquad (1)$$

In the model, the interest rate r responds to inflation and economic growth, like in a Taylor rule. We impose a zero restriction on the coefficients of Z_t in the interest rate equation, implying that the interest rate does not respond directly to the financial variables. This assumption reflects a Taylor rule type policy response in which monetary policy reacts to inflation and output, while financial shocks affect the economy indirectly through their impact on real activity and prices.

The conditioning variables are interacted with interest rate r. This takes into account the possibly nonlinear effects of monetary policy on the economy, conditional on the level of financial stress. This is in line with Aastveit et al. (2017) and Pellegrino (2018), who include variables for financial, economic and policy uncertainty in interaction with the interest rate. In scenario analyses by central banks, the monetary policy response to inflation and growth is sometimes switched off in the model simulations. In section 6.1 we show how this exogenous monetary policy assumption affects the outcomes.

We estimate different specifications of the model, to take into account model uncertainty concerning the parameters and model structure, including the lags of transmission of the shocks. Applications of EMA usually consider uncertainty in the model parameters, to account for model uncertainty. This is also considered by applications of EMA that sample of an ensemble of models (Bankes et al., 2013). While we focus on uncertainty in the conditioning variables, model uncertainty is taken into account by specifying different model specifications: a large, medium and small-scale model and a model excluding interaction effects.

Large model	I-VARX (2,2)	$p = 1, 2$; $n = 1, 2$; $s = 1, 2$
Medium model	I-VARX (2,1)	$p = 1, 2$; $n = 0, 1$; $s = 1$
Small model	I-VARX (2,{1})	$p = 1, 2$; $n = 1$; $s = 1$
No interaction	VARX (2,1)	$p = 1, 2$; $n = 0, 1$

The model specifications differ in terms of the number of lags (p) of the endogenous variables, the lags of the conditioning variables (n) and lags (s) of the interaction term. In the medium-scale model and in the variant excluding interaction effects the conditioning variables have a contemporaneous effect on π, y and r and only a lagged effect in the other two models. The large model includes two lags ($s = 1, 2$) of the interaction term. The small model includes one lag of the conditioning variables and the interaction term. The models are estimated by OLS with data for the euro area over the 2000_Q2 – 2025_Q3 sample period.

If the objective were to predict optimal outcomes for policy purposes, a Bayesian VAR (BVAR) would be the preferred approach. The shrinkage inherent in Bayesian estimation enhances more stable out-of-sample forecasts and is well suited to small sample sizes. Our purpose, however, is methodological rather than predictive. We aim to illustrate how model size, parameter uncertainty and linearity influence the resulting scenarios of interest. For this purpose, OLS offers a more transparent method than Bayesian estimation. Because OLS imposes no priors, the model dynamics are determined by the data. In contrast, BVARs apply shrinkage, meaning that the prior affects the behaviour of interaction terms and the extent to which the model yields extreme outcomes. Consequently, differences in the outcome space partly reflect regularization effects rather than differences related to model size, parameter uncertainty or linearity. To illustrate model behaviour, such as the relationship between fit, size and dispersion of the outcomes, OLS therefore provides a clearer and less confounded representation. The influence of Bayesian estimation is demonstrated in the robustness analysis in Annex A.

5. Empirical set-up

We simulate financial stress scenarios by forecasting the outcome variables π and y with the VAR model, conditional on shocks in the financial variables included in vector Z. These are the uncertainties that determine the uncertainty space. Assuming deep uncertainty about the financial scenario that may hit the euro area economy, we include a range of financial variables in Z which in combination shape a host of multivariate scenarios. The set of conditioning variables includes the euro-dollar exchange rate (FX), the Eurostoxx equity index (EQ), the high-yield corporate bond spread (CSP), the Euro area government bond spread (GSP), the 10 years US Treasury bond yield (US) and the Bitcoin rate (BIT). These variables are included in the model in terms of quarterly changes (see the definitions of the variables in Annex B).

The scenarios are determined by shocks to the financial conditioning variables. The shocks are assumed to occur simultaneously at the start of the scenario horizon (t=0).[2] The one period shocks affect the economy over the full forecast horizon, through the recursive structure of the VAR model. This takes into account that the shocks may have a lagged effect on the economy through various transmission channels. The shock values of the conditioning variables are random draws (samples) from a uniform distribution, implying that each draw is equally likely. A draw is a combination of joint shocks to the conditioning variables and each combination forms a scenario. This way we sample 10,000 scenarios.

The shocks are sampled within a minimum and maximum bound. These restrictions are based on the plausibility of the levels of financial stress. To calibrate the boundaries we use information about the historical lows and highs of the financial variables and supervisory guidance about shocks that banks have to assume for market risks (see Annex A). Figure 2 shows the sampled shocks of the conditioning variables. These are quarter-on-quarter changes, some of which are assumed to have symmetric and others asymmetric upper and lower bounds, based on the information used to calibrate them (see Annex A).

[2] This set-up of the shocks is taken as example for our exercise. The model is flexible to assume the shocks to the conditioning variables occur at a different point in time, also different from each other. The model is also flexible to adjust to forecast horizon.

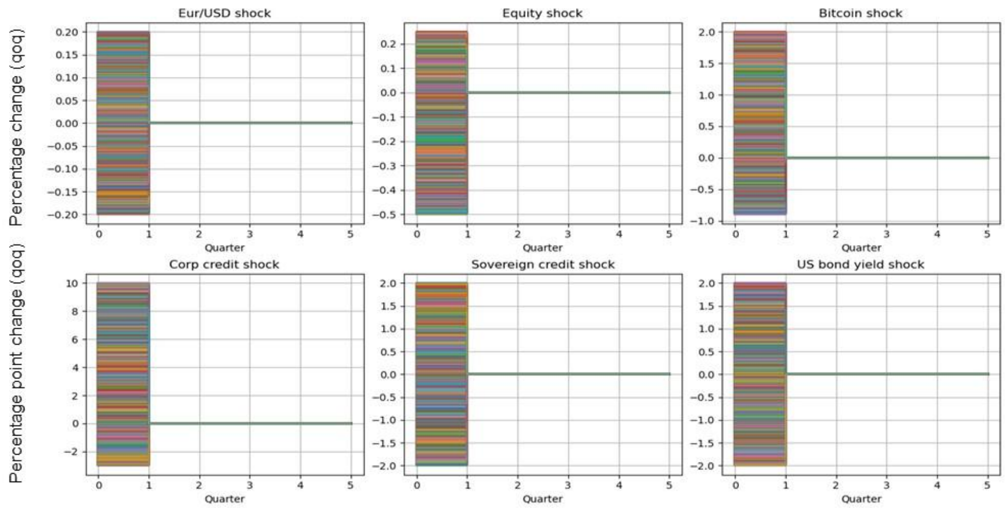

Fig. 2. Sampled shocks in the conditioning variables. Results of 10,000 simultaneously sampled shocks to all conditioning variables at t=0. The shocks are sampled within an upper and lower bound. Shocks to the Euro-dollar, Equity index and Bitcoin rate are percentage changes (upper panels). Shocks to the Corporate credit spread, Sovereign spread and US bond yield are percentage point changes (lower panels).

Following from the shocks, conditional forecasts are made for π and y four quarters ahead. The shocks cause a deviation from the steady state, in which inflation π is assumed to be at the 2% target rate and real GDP growth y to equal potential output growth (assumed to be 1%[3]). The policy-relevant outcomes are states at t=4 in which π and y exceed a certain critical level. We define as critical level for π a deviation of more than 1 percentage point above or below its steady state level and for y a deviation of minus 1 percentage point or more from the steady state growth rate of 1% (such a deviation is associated with a recession). We assume these are the outcomes the central bank wants to avoid (i.c. stagflation and deflationary recession), since these states conflict with its policy objective(s). In the sample period, there are two deflationary recessions; in 2009 following the Global Financial Crisis (GFC) and in 2020 the early stage of the Covid-19 pandemic. Deflationary recessions, i.c. growth slowing down and inflation falling, are typically the outcome of negative demand shocks. The sample does not contain stagflationary periods, which are associated with negative supply shocks.

The stagflation and deflationary recession regimes are highlighted in Figure 3. To avoid those states, the central bank should anticipate the scenarios that lead to the policy-relevant outcomes. For this, it needs to know the scenarios of interest, i.c. the combinations of

[3] The assumed potential growth of 1% is close to the estimates of Bandera et al. (2023) for euro area potential output growth.

financial variables and their shock levels, that are associated with the policy-relevant outcomes. We explore those scenarios by applying the EMA workbench developed by Kwakkel (2017) to the I-VARX model.[4]

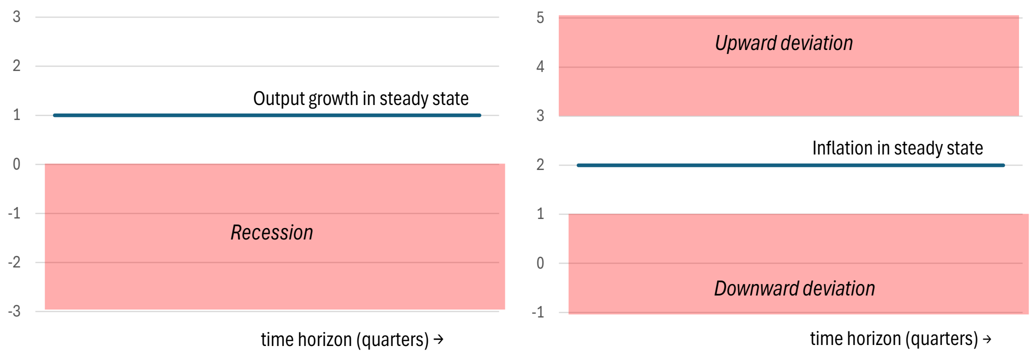

Fig. 3. Stylised picture of policy-relevant outcomes for GDP growth (left panel, annual percentage change on y axis) and inflation (right panel, annual percentage change on y axis).

6. Results

6.1 Outcome space

The outcomes result from the sampled shocks to the conditioning variables. Hence, we do not introduce shocks through the reduced-form residuals, nor do we rely on a structural decomposition of innovations. Instead, the scenarios are generated by shocking the conditioning variables. So the I-VARX is employed to generate conditional forecasts of the outcome variables.

Figure 4 shows the outcome space for each of the four model variants. Each dot is a four quarters ahead conditional forecast for the deviation of π and y from steady state, following joint shocks to all financial variables. The shocks are randomly sampled by 10,000 draws, as explained above.

[4] The documentation and scripts of the EMA workbench are available in EMA Workbench documentation — Exploratory Modeling Workbench.

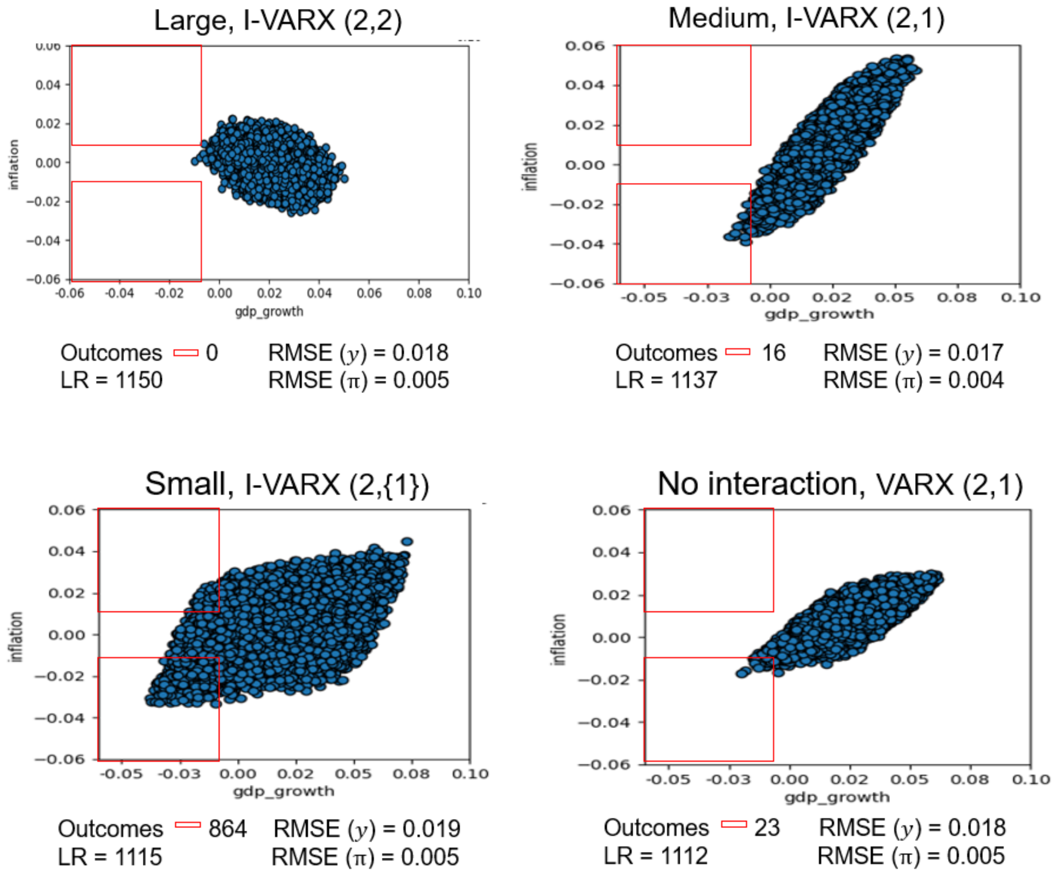

Fig. 4. Outcome spaces filled with conditional forecasts of the outcome variables inflation and GDP growth (π and y as deviation from their steady state). The forecasts result from a large, medium and small-scale I-VARX model and the model excluding interaction effects (VARX). The dots in the outcome spaces are the results of the sampled shocks to the conditioning variables in the model. LR is likelihood ratio as measure of statistical fit. RMSE is root mean square error of the static in-sample forecast of π and y.

In the outcome space, the policy-relevant outcomes can be detected. These are marked by the red boxes in Figure 4. They present states of stagflation and deflationary recession, as determined by the critical levels for π and y specified above. It shows that no such outcomes are forecasted by the large model. The medium-scale model and the model excluding interaction effects forecast several deflationary recessions (captured by the lower left red boxes). The small model forecasts both stagflations (in the upper left red boxes) and deflationary recessions. Overall, the models forecast more cases of deflationary recession than stagflation following the financial shocks. This is in line with the common finding in the literature that a financial crisis presents a negative demand shock, characterised by reduced growth and lower inflation (Forni et al., 2024).

The results show that the small model produces more dispersed outcomes than the medium and large models; its outcome space has the widest coverage. In contrast, the forecasted outcomes of the large model are most concentrated and least dispersed. This suggests that larger models tend to forecast less extreme outcomes and thereby less likely identify policy-relevant scenarios, which are often associated with tail outcomes. A possible explanation for this is the difference in statistical fit across models. Larger models with a higher in-sample fit less likely generate extreme outcomes that are not included in the sample distribution. The small and medium-scale models have lower likelihood ratios (LRs) than the large model, indicating a worse in-sample fit. A poorer model fit can be associated with higher forecast variance, which in turn increases the probability of more extreme forecast outcomes (we elaborate on this in the discussion section).

Comparing the outcome space of the small model with that of the model excluding interaction effects (VARX) shows that the former generates much more policy-relevant outcomes than the latter, while both models have a comparable statistical fit (LR). This suggests that interactions effects – included in the small I-VARX model – are an important driver for extreme outcomes, which are inherently policy-relevant.

Figure 5 illustrates how monetary policy affects the outcome space of the medium model. When monetary policy is exogenised (by setting the corresponding coefficients in the interest rate equation of the VAR to zero, analogous to eliminating the Taylor-rule feedback), the number of scenarios of interest increases substantially (compare the left and right panels of Figure 5). This highlights the stabilising role of monetary policy in the model. By adjusting the policy rate whenever inflation or GDP growth deviates from the policy objective, the central bank counteracts the emergence of adverse macroeconomic outcomes. In our case, active monetary policy dampens the downward effects on GDP growth and inflation resulting from financial stress scenarios. These are associated with negative demand shocks, to which the central bank most obviously would respond by monetary easing. This mitigates the risk of a deflationary recession, as shown by the lower number of policy-relevant outcomes in the left panel of Figure 5 (including monetary policy response) compared to the right panel (no monetary policy response).

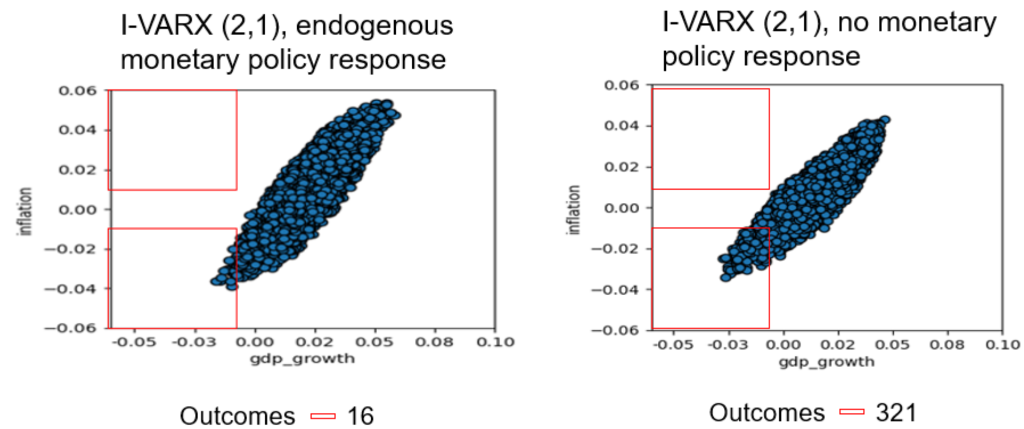

Fig. 5. Outcome spaces with conditional forecasts of inflation and GDP growth (π and y as deviation from their steady state), based on the medium VAR model I-VARX(2,1) including (left panel) and excluding (right panel) the monetary policy response.

6.2 Uncertainty space

By behaviour-based scenario discovery the policy-relevant outcomes are linked to the combinations of uncertainties that meet the critical levels of the outcomes. This reveals the scenarios of interest. Those scenario can be detected in the uncertainty space, which spans all combinations of the sampled values of uncertainties. Figures 6 and 7 show the uncertainty space of the small and medium-scale model (the large model does not have policy-relevant outcomes). The red boxes represent subspaces which cover (most of) the scenarios of interest. These scenarios are denoted by the orange dots and represent the combinations and levels of the sampled conditioning variables which are related to the policy-relevant outcomes.

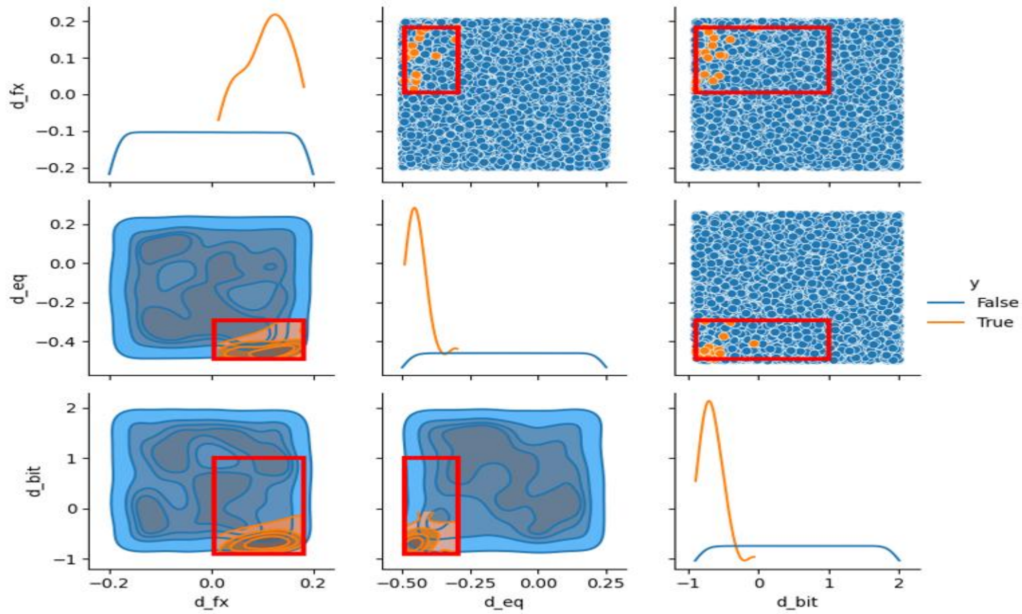

Fig. 6. Uncertainty space of medium-scale model. The dots in each panel are combinations of different shock values of two conditioning variables. Orange dots (surfaces) in the red boxes are combinations (scenarios) of interest. Blue dots (surfaces) are scenarios that do not meet the outcome criterium for policy relevance. The diagonal panels show the density of the shock values for each conditioning variable.

The medium model results in 16 scenarios of interest (out of 10,000 sampled scenarios, see Figure 6). These scenarios are described in the boxes by combinations of three conditioning variables: the euro-dollar rate, the equity index and bitcoin rate. Those combinations of values of the conditioning variables predict outcomes for GDP growth and inflation that present deflationary recessions. In case of the medium model, all scenarios of interest fall inside the red boxes, meaning that coverage = 1 (coverage defined as the number of scenarios of interest within the box over the total number of scenarios of interest). The identification of 16 scenarios of interest out of 10,000 sampled scenarios implies a likelihood of 0.2% for a scenario of interest, assuming each sample draw is equally likely.

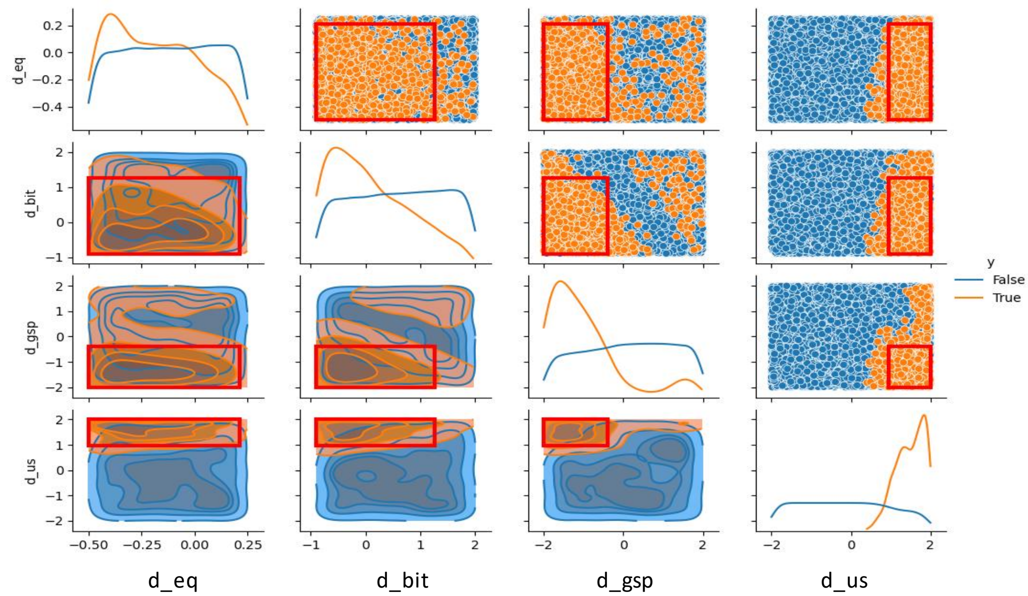

Fig. 7. Uncertainty space of small model. The dots are in each panel are combinations of different shock values of two conditioning variables. Orange dots (surfaces) in the red boxes are combinations (scenarios) of interest. Blue dots (surfaces) are scenarios that do not meet the outcome criterium for policy relevance. The diagonal panels show the density of the shock values for each conditioning variable.

The boxes in the uncertainty space of the small model capture 586 scenarios of interest, out of a total of 864 scenarios of interest (see Figure 7). The scenarios in the boxes are described by combinations of four conditioning variables: the US bond yield, the equity index, bitcoin rate and the Euro area government bond spread. Hence, not only the number of scenarios of interest produced by the small model differs from the medium model, but also the uncertainties (conditioning variables) that determine them. It reflects the model dependency of the identified scenarios, related to the different set of policy-relevant outcomes forecasted by the small model. These outcomes reflect both states of stagflation and deflationary recessions, while the medium model only forecasts the latter. The more diverse behaviour of outcomes in the small model implies that different uncertainties describe the outcomes.

Moreover, in contrast to the medium model, not all scenarios of interest are covered by the red boxes. For example, the bifurcation in the relationship between the government bond spread and both the equity index and the bitcoin rate is not captured (see the third column of Figure 7). This bifurcation implies that both low and high values of the spread are associated with policy-relevant outcomes. Hence, the number of scenarios of interest in the boxes of the uncertainty space in Figure 7 is less than the total number of policy-relevant outcomes in Figure 4, implying coverage < 1. Higher coverage can be obtained by sacrificing density, defined as the number of scenarios of interest within the box over the total number of scenarios within the box. Thereby, density reflects the probability that the scenarios described by the box are of interest, or policy-relevant.

The PRIM algorithm searches the optimal trade-off between coverage and density (Kwakkel and Jaxa-Rozen, 2015). The optimal trade-off depends on the specific application and the EMA workbench allows the analyst to make his own trade-off. The trade-off curve in Figure 8 shows the series of boxes – denoted by the dots - resulting from the PRIM algorithm. All dots on the trade-off curve represent boxes in the uncertainty space that meet the criterium for the policy-relevant outcomes. The right-hand vertical axes of Figure 8 show the number of conditioning variables (uncertainties) that defines the scenarios in the box. The colour code indicates how many uncertainties are used for this. These uncertainties are restricted by the bounds within they are sampled.

The trade-off curve is delivered by the PRIM algorithm that peels (removes) parts of the uncertainty space, by trimming the ranges of the uncertain variables. The restricted dimensions count how many uncertainties have any range removed and so are being restricted. In the end, the scenarios are defined by the restricted uncertainties. Variables whose ranges are not reduced by PRIM are irrelevant for the scenario and do not define the scenario box. Thereby a movement along the trade-off curve upper left is a movement from general-to-specifc, since each peel restricts the box in the uncertainty space more, by removing ranges of the uncertainties. Putting more restrictions (constraints) on the scenario box means that its density increases and the scenarios described by the box become more specific.

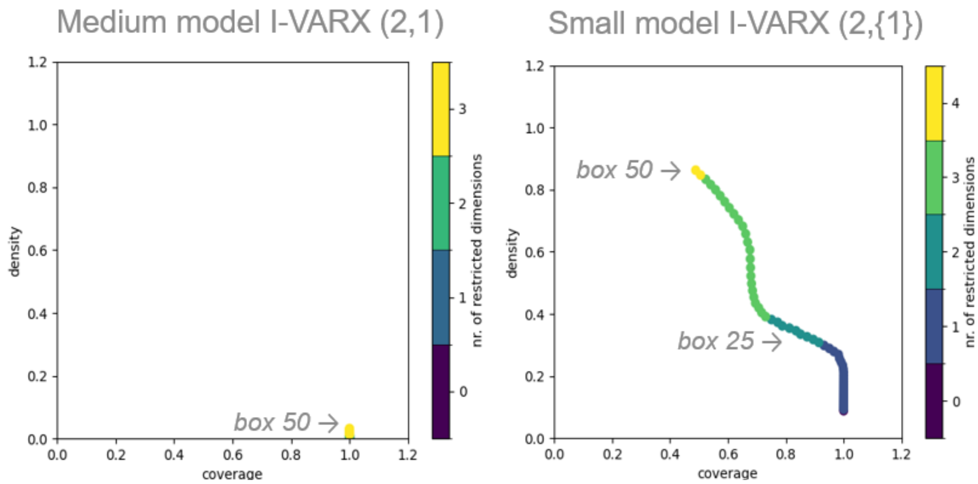

Fig. 8. Trade-off between coverage and density. Each dot is a box in the uncertainty space with a different trade-off. All dots are associated with policy-relevant outcomes. The right-hand vertical axes show the number of conditioning variables (uncertainties) that defines the scenarios in the box. The colour code indicates how many uncertainties are used for this.

The trade-off curve of the medium model (left-hand panel of Figure 8) indicates that the behaviour of the outcomes is only described by combinations of three uncertainties. It implies there are no trade-offs between coverage and density. This reflects the homogenous character of the policy-relevant outcomes, which all reflect a deflationary recession. This state is described by a limited and fixed set of three uncertainties, as described above. There is no further peel that improves the density. This is explained by the very sparse number of policy-relevant outcomes and related scenarios of interest (0.2% of the samples).

The trade-off curve of the small model (right-hand panel of Figure 8) shows that density increases, but coverage decreases if more restrictions are used to define the scenarios of interest. This suggests that more complex and specific scenarios (made up by a higher number of conditioning variables) increases the likelihood that some scenarios of interest are overlooked, by not being covered in the box. They would only emerge as dimension of the scenarios of interest when the conditioning variables vary freely, by their values being not or less restricted by bounds. Another disadvantage of a high number of restrictions used to describe the scenarios of interest is that the scenarios become less easy to interpret. Every added restricted dimension means that an extra variable is used to describe the scenarios of interest. This complicates the narrative building.

6.3 Uncertainties

The PRIM algorithm provides the bounds of the uncertainties that define the identified boxes. Thereby the box gives a more precise description of the scenarios of interest. For the medium model these are defined by a fall of the equity index by 30 to 50% and an appreciating of the euro up to 20% (Figure 9). The upper bound of the change in the equity index and the lower bound of the euro appreciation are statistically significant, as indicated by the quasi p-values. This underscores that both variables are important for distinguishing the box.

The bitcoin is also a dimension of the scenarios of interest. However, the restricted changes in the bitcoin rate range through both negative and positive values (with the quasi p-value being less discriminative, only being significant at a 10% level). This implies that the bitcoin is not an informative variable for the scenarios of interest. Hence, the policy-relevant outcomes of the medium model, i.c. deflationary recessions, are associated with scenarios in which the equity index drops sharply and the euro appreciates. A narrative that could describe this scenario is the situation following the Liberation day event in April 2025, when a flight to safety caused the euro to appreciate and stock prices to fall.

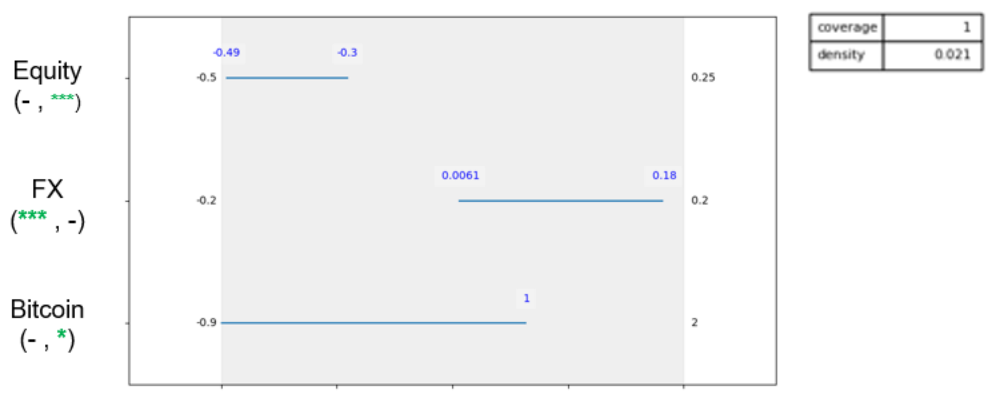

Fig. 9. Box with uncertainties restricted in the medium model, including quasi-p values for each restriction. Upper and lower bounds are denoted on the right and left vertical axes. Coverage and density at the top right side.

For the small model, the scenarios of interest are described by an increase of the US bond yield by 100 to 200 basis points and a drop in the euro area government bond spread by 40 to 200 basis points (Figure 10). The statistical significance of the bounds of both variables underscores that they are important for distinguishing the box. The equity index and bitcoin are also dimensions of the scenarios of interest, but the restricted changes of both variables range through both negative and positive values. This implies that the equity index and bitcoin are not really informative for the scenarios of interest.

Hence, the policy-relevant outcomes forecasted by the small model, i.c. stagflation and deflationary recessions, are associated with scenarios in which the US bond yield rises strongly and the euro government bond spread drops. A narrative that could describe such a scenario is a global increase of risk-free bond yields (if the German bund yield increases, the euro government bond spread drops). Rising yields can reflect upward shift in inflation expectations as in stagflationary regimes. Rising bond yields also tighten financial conditions and this could be a harbinger of deflationary recessions.

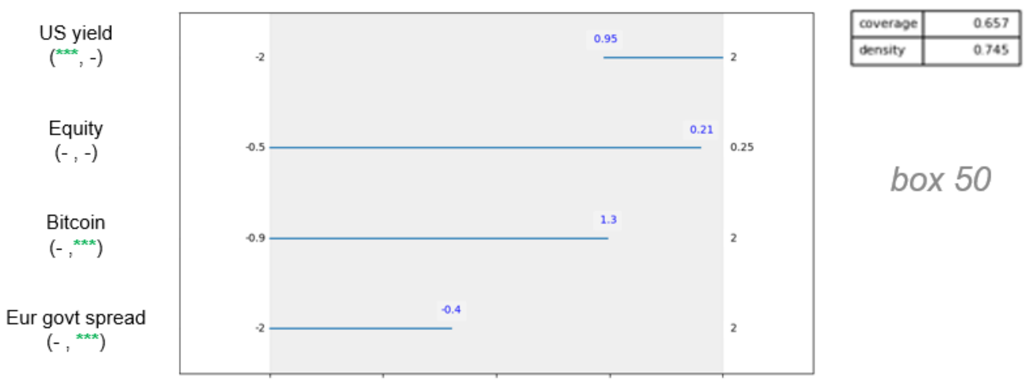

Fig. 10. Box 50 with uncertainties restricted in the small model, including quasi-p values for each restriction. Upper and lower bounds are denoted on the right and left vertical axes. Coverage and density at the top right side.

The larger number of conditioning variables used to distinguish the box in case of the small model complicates the interpretation of the scenarios. By shifting down along the PRIM trade-off curve we can reduce the number of restricted variables. This enlarges the coverage of the box at the expense of a lower density. Figure 11 shows that shifting from box 50 to box 25 (along the curve in the right-hand panel of Figure 8) results in only two conditioning variables distinguishing the box. Using less conditioning variables as restrictions means that more scenarios of interest (774) are covered by box 25 than by box 50 (568). This underscores that designing scenarios by fewer conditioning variables reduces the likelihood that scenarios of interest are overlooked. The identification of 774 scenarios of interest out of 10,000 sampled scenarios implies a likelihood of 7.8% for a scenario of interest captured by the box, assuming each sample draw is equally likely.

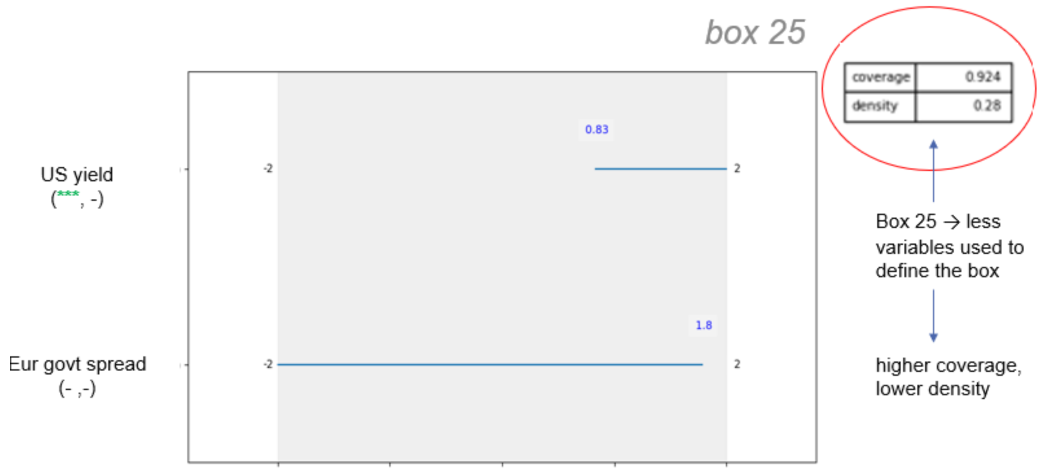

Fig. 11. Alternative box (25) with uncertainties in the small model, including quasi-p values for each restriction. Upper and lower bounds are denoted on the right and left vertical axes. Coverage and density at the top right side.

6.4 Driving features

Feature scores can be used to assess the influence of the various conditioning variables and model parameters on the policy-relevant outcomes. Feature scoring is a machine learning technique for identifying the relative importance of factors for particular model outcomes (Kwakkel, 2017). The feature scores could be used for the selection of conditioning variables in the model and so guide the dimensionality reduction in scenario discovery.

The policy-relevant outcomes of the medium model are most influenced by changes in the bitcoin rate, equity index and the euro area government bond spread (Figure 12, left panel). This indicates that the single conditioning variables that influence the policy-relevant outcomes most are not necessarily similar to the variables that comprise the scenarios of interest. In the small model, this is more likely the case, given the relative strong influence of the US bond yield and – to a lesser extent – the euro area government bond spread on the policy-relevant outcomes (Figure 12, right panel). This indicates that these conditioning variables are consistently important (robust drivers) for the policy-relevant outcomes.

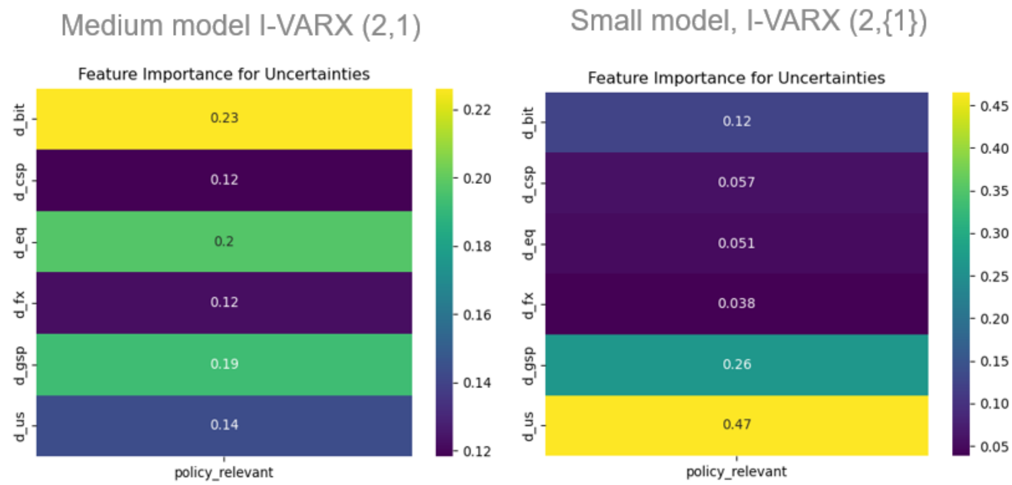

Fig. 12. Feature scores for conditioning variables in the medium-scale model (left panel) and small-scale model (right panel).

To assess the relative contribution of the interaction effects on the policy-relevant outcomes, we compute the feature scores for the model parameters of the conditioning variables that determine the scenarios of interest in the medium model: the equity index and exchange rate.[5] It concerns the parameters in model equation 1 included in vector Γ (coefficients of EQ and FX) and vector Θ (coefficients of the interaction between the interest rate and EQ and FX respectively).

The feature scores in Figure 13 show that the interaction effects are relative important drivers of the policy-relevant outcomes. In both the equations for output growth and inflation the parameters of the interaction effects have a relative high feature score. This indicates that the non-linear dimension of the model is an important driver of more extreme model outcomes, which inherently determine the policy-relevant model outcomes.

[5] To compute the feature scores of the model parameters, we assume that they are also uncertainties in the sampling exercise (with a very narrow bound between the upper and lower parameter values, to not outweigh the influence of the conditioning variables, taking into account that a wider bound increases the feature score).

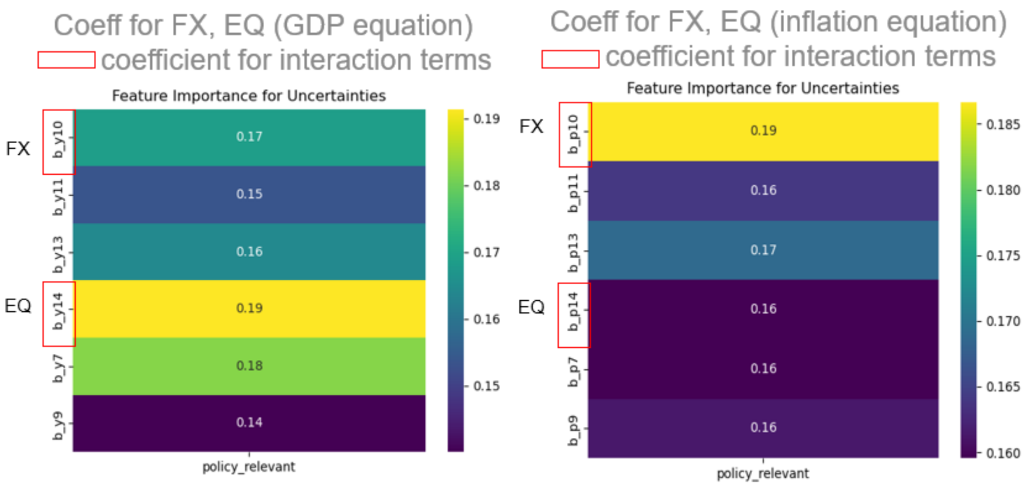

Fig. 13. Feature scores for parameters of the exchange rate (FX) and equity index (EQ) in the medium model for the GDP growth equation (left panel) and inflation equation (right panel). Parameters b_y7 and b_y9 (b_y11 and b_y13) and b_p7 and b_p9 (b_p11 and b_p13) are the contemporaneous and one quarter lagged coefficients for FX (EQ) in the GDP growth and inflation equation. Parameters b_y10 and b_y14 (b_p10 and b_p14) are the coefficients of the interaction term with FX and EQ in the GDP growth (inflation) equation.

7 Discussion

The results of the EMA exercise provide several insights about scenario analysis. First, the identification of scenarios that are associated with policy-relevant outcomes is model dependent. Even variants of the same model type (the VAR model in this case) identify different conditioning variables as dimensions of scenarios of interest. This underscores the use of a suite of models in scenario analyses. This can be complemented by narratives that inform the design of the model structure, including the relevant economic dynamics (Moallemi et al., 2017).

Second, narratives are also useful to build and communicate a coherent story based on the variables identified by EMA as drivers of scenarios of interest, as demonstrated in the previous section. The scenarios uncovered by EMA may involve combinations of variables that appear non-intuitive or less likely, because they have not occurred in the data. Consequently, EMA compels the analyst and policymaker to reflect systematically on potential developments that deviate from historical experience. In doing so, EMA supports the building of previously unseen narratives, as a means of stimulating imagination about unknown futures (Bankes et al., 2013; Kwakkel and Pruyt, 2013).

Third, larger models with a higher in-sample fit less likely generate tail outcomes that are (inherently) policy-relevant. Models that fit the historical data well will less likely forecast future states of the world that are unprecedented and not included in the sample distribution. Relatedly, the probability distributions that underlie the estimation of structural models are not informative under deep uncertainty, when the knowledge for assigning probabilities is insufficient. This is another argument to use a suite of models, including simple models, for scenario analyses.

Fourth, the results suggests that scenarios should not be too specific, meaning that the number of conditioning variables and the constraints put on their ranges should be limited. Putting more restrictions on the variables (to identify scenarios of interest) increases the risk that scenarios of interest are overlooked. Another argument is that restricting more variables in the scenario design make scenarios less easy to interpret. This complicates the narrative building.

Fifth, our results suggest that the non-linear dimension of the model is an important driver of policy-relevant model outcomes, which are inherently associated with tail outcomes. In a survey among central banks, non-linearities are acknowledged as an important feature to address in scenario analyses (BIS, 2025). Another insight from the feature scores is that single conditioning variables that mostly influence the policy-relevant outcomes are not necessarily similar to the variables that comprise the scenarios of interest. This underscores the use of both sensitivity tests (based on shocks to single variables) and multivariate scenarios for exploring uncertain future states of the world. In particular when models feature nonlinearities and interactions, it is useful to consider the set of all possible combinations of conditioning variables as in an uncertainty space, beyond sensitivity tests of single model inputs (Saltelli et al., 2019).

8 Conclusion

EMA is a useful complementary tool for macroeconomic scenario design under deep uncertainty. By sampling a large set of multivariate financial shocks and forecasting their economic impact through an interacted VARX model, we identify combinations of conditioning variables that lead to policy-relevant outcomes. From this we trace the scenarios of interest that are associated with outcomes of stagflation and deflationary recessions.

The EMA framework is applied to different variants of the I-VARX model. The results show that the identification of scenarios of interest is model-dependent. In the medium-scale VAR model the policy-relevant outcomes are associated with scenarios in which the equity index drops sharply and the euro appreciates. In the small model the forecasted policy-relevant outcomes are associated with scenarios in which the US bond yield rises strongly and the euro government bond spread drops. The EMA approach can be combined with the narrative approach to build and communicate a coherent story, based on the variables identified as drivers of scenarios of interest.

Our findings demonstrate that EMA offers a valuable complement to existing scenario analysis frameworks used by central banks. The EMA approach can uncover scenarios of interest that may be overlooked by narrative or statistical approaches, because these conventional methods tend to restrict the range of scenarios by selecting topical scenarios that are deemed plausible because they do not diverge too much from the baseline. This design process inherently limits the range of uncertain variables and their possible combinations. Such dimensionality reduction can result in surprising developments being overlooked. Moreover, the use of large models that fit the historical data well can further reduce the likelihood of identifying extreme or unprecedented scenarios.

As monetary policy increasingly has to navigate environments characterized by extreme shocks and deep uncertainty, the use of EMA can enhance the robustness of scenario analyses and so support more resilient policy design. Future research could extend the EMA framework by exploring the use of alternative model classes, or integrating EMA with conventional scenario frameworks.

Literature

Aastveit, K.A., G. J. Natvik and S. Sola (2017), Economic uncertainty and the influence of monetary policy, Journal of International Money and Finance, 76(C), 50-67.

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

Adrian, T., D. Giannone, M. Luciani and M. West (2025), Scenario Synthesis and Macroeconomic Risk, IMF Working paper, 105.

Aikman, D., R. Angotti and K. Budnik (2024), Stress testing with multiple scenarios: a tale on tails and reverse stress scenarios, ECB Working paper, 2941.

Bandera, K., J. Bodnár, B. Le Roux and B. Szörfi (2023), The impact of the COVID-19 shock on the euro area potential output – a sectoral approach, SUERF presentation.

Bank for International Settlements, BIS (2025), Monetary policy decision-making and communication under high uncertainty, BIS papers, 163.

Bankes, S. C. (1993), Exploratory Modeling for Policy Analysis, Operations Research, 41(3), 435-449.

Bankes, S. C., Walker, W. E. and J. H. Kwakkel (2013), Exploratory Modeling and Analysis. In G. S. and M. C. Fu (Eds.), Encyclopedia of Operations Research and Management Science (3rd ed.), Berlin, Germany: Springer.

Basel Committee on Banking Supervision BCBS (2019), Minimum capital requirements for market risk (January 2019, revised 2021).

Basel Committee on Banking Supervision BCBS (2023), Calculation of RWA for market risk.

Basel Committee on Banking Supervision BCBS (2024), Recalibration of shocks in the interest rate risk in the banking book standard.

Benguria, F. and A. M. Taylor (2020), After the Panic: Are Financial Crises Demand or Supply Shocks? Evidence from International Trade, American Economic Review: Insights, American Economic Association, 2(4), 509-526.

Bernanke, B. (2024), Forecasting for monetary policy making and communication at the Bank of England: a review, Bank of England Independent Evaluation Office.

Bryant, B. P. and R. J. Lempert (2010), Thinking Inside the Box: a participatory computer-assisted approach to scenario discovery, Technological Forecasting and Social Change, 77(1), 34-49.

ECB (2025), A strategic view on the economic and inflation environment in the euro area, ECB Occasional paper, 371.

Friedman, J. H. and N.I. Fisher (1999), Bump hunting in high-dimensional data, Statistics and computing, 9(2), 123–143.

Forni, M., Gambetti, L., Maffei-Faccioli, N. and L. Sala (2024), Nonlinear Transmission of Financial Shocks: Some New Evidence. Journal of Money, Credit and Banking, 56, 5-33.

Kwakkel, J. H., W.L. Auping and E. Pruyt (2013), Dynamic scenario discovery under deep uncertainty: The future of copper, Technological Forecasting and Social Change, 80(4), 789–800.

Kwakkel, J. H., M. Haasnoot and W. E. Walker (2016), Comparing robust decision-making and dynamic adaptive policy pathways for model-based decision support under deep uncertainty. Environmental Modelling & Software, 86, 168-183.

Kwakkel, J. H. and M. Jaxa-Rozen (2016), Improving scenario discovery for handling heterogeneous uncertainties and multinomial classified outcomes, Environmental Modelling & Software, 79, 311–321.

Laxton, D., H. Igityan and S. Mkhatrishvili (2025), Monetary policy credibility, avoiding dark corners, and risk management: a response to Ben Bernanke's review of monetary policy-making at the Bank of England, Oxford Review of Economic Policy, 41(2), 452–483.

McKay, M. D., R.J. Beckman and W.J. Conover (1979), A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code, Technometrics, 21(2), 239-245.

Moallemi, E. A., F. de Haan, J. Kwakkel and A. Lu (2017), Narrative-informed exploratory analysis of energy transition pathways: A case study of India's electricity sector, Energy Policy, 110(C), 271-287.

Pellegrino, G. (2018), Uncertainty and the real effects of monetary policy shocks in the Euro area, Economics Letters, 162(C), 177-181.

Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S. and Q. Wu (2019), Why so many published sensitivity analyses are false: a systematic review of sensitivity analysis practices, Environmental modelling and software, 114, 29-39.

Wieles, C., J. Kwakkel, W. L. Auping and J. W. van den End (2025), Scenario discovery to address deep uncertainty in monetary policy, Journal of Economic Dynamics and Control, 179, 1-16.

Annex A. Outcomes of Bayesian VAR (BVAR) models

To illustrate the influence of Bayesian estimation, we forecast the outcome space of the four model specifications with Bayesian VAR (BVAR) models. BVARs apply shrinkage, meaning that the prior affects the behaviour of interaction terms and the extent to which the model yields policy-relevant outcomes. Bayesian estimation incorporates prior beliefs about the parameters, for instance through a Minnesota prior in VAR applications, which pulls coefficients toward benchmark dynamics, such as a random walk or zero. This shrinkage stabilizes the model by reducing parameter variance, preventing overfitting, and dampening implausibly large responses. The resulting BVAR forecasts therefore reflect a combination of information from the data and the prior, rather than relying solely on the data as in OLS. The influence of Bayesian estimation is illustrated in Figure A.1 across the four model specifications.

In the large BVAR specification, I-BVARX(2,2), the forecast distribution spans a substantial number of policy-relevant outcomes, predominantly associated with stagflationary scenarios (highlighted by the red box in the upper-left panel of Figure A.1). By contrast, the corresponding large OLS-estimated VAR does not generate outcomes in these regions (see Figure 4 in section 6.1). This divergence reflects the interaction between Bayesian shrinkage and model dimensionality. In the large OLS VAR, the proliferation of parameters leads to imprecise coefficient estimates, which attenuate dynamic propagation and keep the forecasts close to the historical sample distribution. As a result, the implied outcome space remains relatively concentrated. In contrast, the Minnesota prior in the BVAR regularizes the high-dimensional coefficient space by shrinking poorly identified parameters toward parsimonious structures that favour persistence and own-lag dynamics. This allows shocks to propagate over longer horizons, expanding the range of attainable outcomes and increasing the likelihood of policy-relevant scenarios.

The outcome spaces of the medium-scale VAR and BVAR models also display a different dispersion and coverage of policy-relevant scenarios. The OLS-estimated I-VARX(2,1) forecasts more dispersed outcomes overall, as reflected in the broader spread of the outcome cloud (see Figure 4). However, despite this larger dispersion, the OLS VAR covers relatively few policy-relevant outcomes. In contrast, the outcomes generated by the Bayesian specification, I-BVARX(2,1), are more concentrated, reflecting the regularising effect of the prior. While this results in a narrower overall dispersion of outcomes, the BVAR covers a substantially larger number of policy-relevant outcomes. The shrinkage imposed by the prior stabilises the estimated dynamics and leads to a denser clustering of outcomes in policy-relevant regions, including stagflationary regimes (see Figure A.1). In this application, Bayesian shrinkage reduces overall dispersion but leads to more policy-relevant outcomes.

The small BVAR model (I-BVARX(2,{1})) generates fewer policy-relevant outcomes than its OLS counterpart (see the lower-left panel of Figure A.1). This difference also relates to Bayesian shrinkage. In the OLS specification, the coefficients are estimated without prior restrictions, allowing the dynamics (particularly those associated with the interaction terms) to operate freely. As a consequence, shocks propagate more strongly and produce a wider dispersion of forecasted outcomes for inflation and GDP growth. In the Bayesian version of the same model, shrinkage compresses the interaction effects and pulls coefficients toward the prior, thereby dampening the dynamic responses and yielding a narrower outcome space with fewer extreme observations.

The opposite pattern arises when the model excludes interaction terms. Although the BVAR model without interactions (BVAR(2,1)) has the similar number of parameters as its OLS counterpart, it is less prone to over-parameterisation, which means that the prior binds less tightly. With weaker shrinkage, the BVAR without interactions allows coefficients to fluctuate more strongly and to extrapolate more from the data, producing a wider outcome space and thicker tails (as shown in the lower right panel of Figure A.1). As a result, linear models estimated with a BVAR prior can paradoxically produce more volatile outcomes than their interaction-augmented counterparts. The outcome distribution can thus become more elongated and tilted, with more outcomes in the extreme regions, in this case regions with low inflation with negative GDP growth (deflationary recessions).

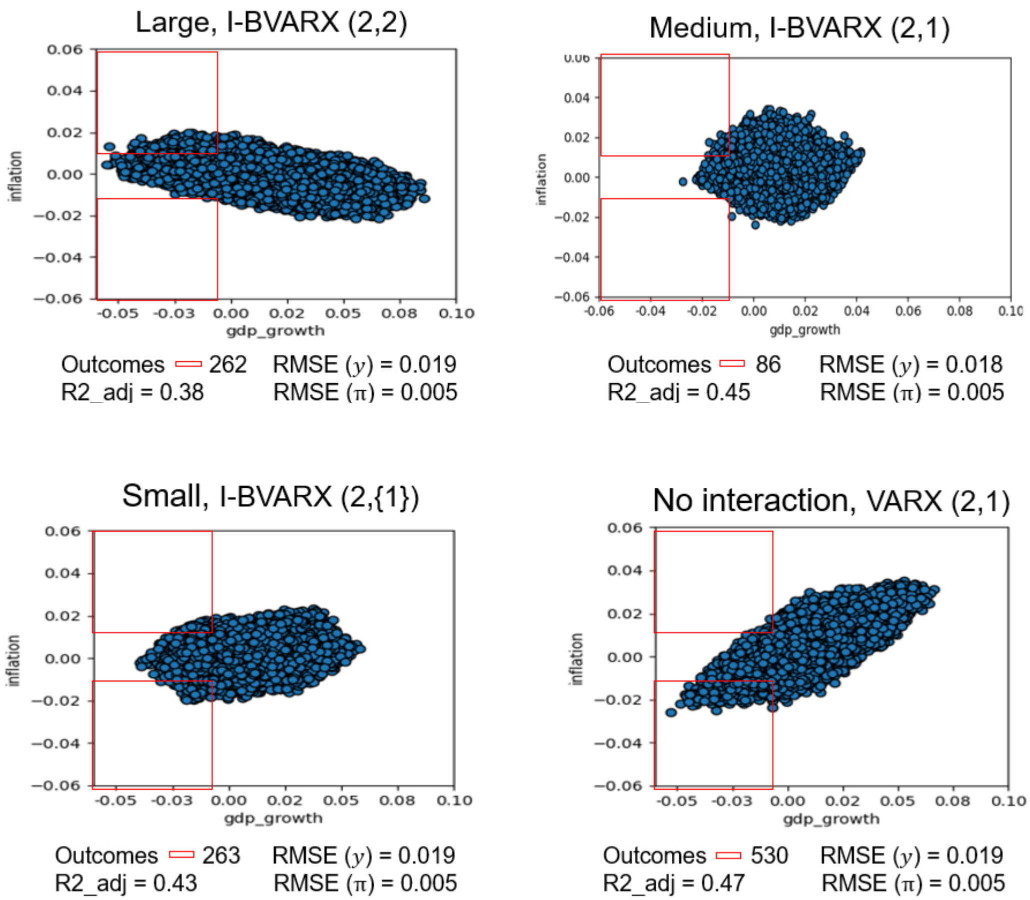

Fig. A.1 Outcome spaces with conditional forecasts of the outcome variables inflation and GDP growth (π and y as deviation from their steady state). The forecasts result from a large, medium and small-scale I-BVARX model and the model excluding interaction effects (BVARX). Bayesian estimation is conducted using a Minnesota prior with a tightness parameter of $\lambda = 0.1$. The dots in the outcome spaces are the results of the sampled shocks to the conditioning variables in the model. R2_adj is the adjusted R2 ratio as measure of statistical fit. RMSE is root mean square error of the static in-sample forecast of π and y.

Annex B. Definition of model variables

The endogenous (outcome) variables in the I-VARX model are inflation (π), GDP growth (y) and the money market rate (r) for the euro area. Variable r is the 3 month Euribor interbank rate, which is used as proxy for the ECB policy rate. Variables π and y are deviations from their steady state levels and are included in terms of annual percentage point changes. Variable r is included in terms of levels. Source of these variables is the ECB data portal.

The conditioning variables are included in terms of quarterly percentage change (Euro-dollar (FX), Eurostoxx equity index (EQ), Bitcoin rate (BIT)) or quarterly percentage point change (high-yield corporate bond spread (CSP, defined as the Euro area corporate bond yield minus the German government bond rate), the euro area government bond spread (GSP, defined as the Euro area average government bond yield minus the German government bond rate) and the 10 years US Treasury bond yield (US).

The transformed endogenous and conditioning variables included in the model are all stationary. Based on the augmented Dickey–Fuller test, the null hypothesis of a unit root is rejected for all variables at the 1% significance level, except for the money market rate, for which the null hypothesis is rejected at the 5% significance level.

Table A.1 provides the sources of the conditioning variables and the boundaries of the sampled shocks. The boundaries are calibrated, based on the observed lows and highs of the conditioning variables and on supervisory guidance of the BCBS for market risk shocks to banks' portfolio holdings.

Table A.1 Boundaries and sources of the conditioning variables (quarterly changes, either percentage or percentage point change)

	FX	EQ	BIT	US	CSP	GSP
EMA boundaries						
minimum	-20%	-50%	-90%	-2	-3	-2
maximum	20%	25%	200%	2	10	2
Historical min,max						
minimum	-12%	-24%	-35%	-0.8	-5.5	-0.7
maximum	8%	12%	172%	1.0	10.7	0.6
Stdev	4%	7%	32%	0.4	1.6	0.2
BCBS	-8%	-50%	-	2	-20%	-
Source	ECB	ECB	FRED	FRED	FRED	FRED

EMA boundaries are the minimum and maximum bounds of the sampled shocks applied in the EMA workbench. Historical min, max are the minimum and maximum levels observed in the 2000_Q2 -2025_Q3 period. Stdev is standard deviation in that period. BCBS is the Basle supervisory guidance for shocks to market risk on banks' portfolio holdings (BCBS, 2019, 2023, 2024). ECB is ECB data portal. FRED is St. Louis Fed database.

EUROSYSTEEM

De Nederlandsche Bank N.V.
Postbus 98, 1000 AB Amsterdam
020 524 91 11
dnb.nl