

Finance and Economics Discussion Series

Federal Reserve Board, Washington, D.C.

ISSN 1936-2854 (Print)

ISSN 2767-3898 (Online)

Risks and Uncertainty in Monetary Policy

Tobias Adrian, Domenico Giannone, Matteo Luciani, and Mike West

2026-061

Please cite this paper as:

Adrian, Tobias, Domenico Giannone, Matteo Luciani, and Mike West (2026).
“Risks and Uncertainty in Monetary Policy,” Finance and Economics Discussion Series 2026-061. Washington: Board of Governors of the Federal Reserve System,
<https://doi.org/10.17016/FEDS.2026.061>.

NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.

Risks and Uncertainty* in Monetary Policy

TOBIAS ADRIAN
International Monetary Fund
tadrian@imf.org

DOMENICO GIANNONE
Johns Hopkins University
domenico.giannone@jhu.edu

MATTEO LUCIANI
Federal Reserve Board
matteo.luciani@frb.gov

MIKE WEST
Duke University
mike.west@duke.edu

June 15, 2026

Abstract

Central banks monitor macroeconomic risk through two traditions: scenario analysis, regularly used since the mid-1990s, and distributional forecasting, practiced since the late 1960s. The two are complementary but separate: scenarios provide narratives without probabilities, while predictive distributions provide probabilities with limited economic interpretation. Treating baseline forecasts and scenarios as conditional predictive densities, and distributional forecasts as reference predictive distributions, places both within a common framework and clarifies their roles. The Scenario Synthesis assigns weights to scenarios consistent with the reference distribution, offering a practical and reproducible tool for risk assessment and policy deliberation under deep uncertainty.

Keywords: Scenarios, fan charts, growth-at-risk, model uncertainty, Bayesian predictive synthesis.

JEL Codes: C1, C11, C53, E32, E37, E58.

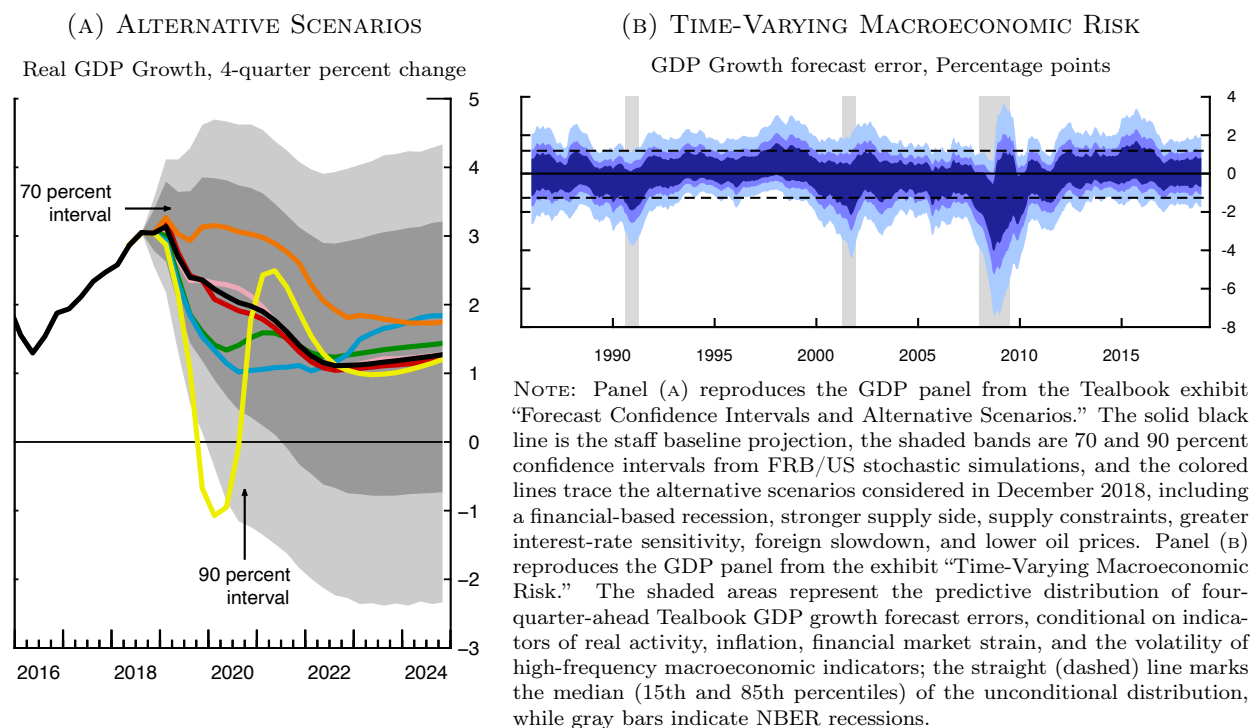
*We are grateful for valuable comments and suggestions from Marta Bańbura, Thomas Carter, Matyas Farkas, Manuel Gonzalez-Astudillo, Edward Herbst, David Hofman, Ruy Lama, Jesper Lindé, Simon Lloyd, Shalva Mkhatrihsvili, Steve Mulema, Pau Rabanal, Yuji Sakurai, Giovanni Sciacovelli, and Dima Solohub.

Disclaimer: The views expressed in this paper are those of the authors and do not necessarily reflect the views and policies of the Board of Governors, the Federal Reserve System, or the International Monetary Fund, its Management, or its Executive Directors.

1 Introduction

Central banks have monitored macroeconomic risk systematically for roughly three decades. Staff regularly brief policy committees on the risks surrounding the economic outlook. At the Federal Reserve, this practice is evident in the Tealbook, the briefing document on economic and financial conditions that staff prepare for the Federal Open Market Committee (FOMC) ahead of each policy meeting. Since 2010, every Tealbook has contained a chapter titled “Risks and Uncertainty” that assesses the forces that could drive the economy away from the baseline projection. Figure 1 shows the two main tools Federal Reserve Board staff use to describe and communicate risk in the December 2018 Tealbook. Panel (A) presents the *alternative scenarios*, a few fully articulated paths for the economy built on internally consistent assumptions about shocks and their transmission mechanisms. Panel (B) presents a *predictive distribution* for GDP growth centered on the staff’s baseline.

Figure 1: THE TWO APPROACHES TO RISK IN THE DECEMBER 2018 TEALBOOK



SOURCE: Both panels are reproduced from the “Risks and Uncertainty” chapter of the December 2018 Tealbook.

Despite appearing in the same chapter, the two exhibits in Figure 1 are not connected through a formal framework. The gap is conceptual, not merely organizational. The scenarios provide economic narratives but no probabilities: they describe how a financial-based recession might unfold, but not how likely it is. The predictive distribution provides probabilities but no economic interpretation: it shows that downside risk has risen sharply, but

not why. Each tool answers the question the other leaves open. A policymaker who asks “what is the probability of a financial-based recession?” or “what narrative underlies this tail risk?” finds no answer in either exhibit.

This paper makes three contributions. First, it provides a historical account of how and why central banks developed the two approaches—tracing the intellectual origins of scenario thinking and probabilistic forecasting, the institutional pressures that shaped them, and why they developed in parallel rather than in dialogue. Second, it develops a formal framework that bridges them—the Scenario Synthesis—which assigns probabilities to named scenarios consistently with the predictive distribution and recovers economic narratives from probabilistic evidence, allowing the two exhibits to speak to each other. Third, it shows empirically that the framework gives central banks a disciplined way to characterize the balance of risks, assess whether scenario sets span the relevant sources of macroeconomic risk, and attach probabilities to named risks under clear conditions.

The historical survey reveals that these two approaches developed in parallel for distinct institutional and intellectual reasons, and that strong views persist about which one is superior. [Bernanke \(2024\)](#), for example, in a review of the Bank of England’s forecasting framework, criticized its fan charts—a format for communicating predictive distributions—as difficult to interpret and communicate, and recommended greater reliance on narrative scenarios. However, as [Bauer et al. \(2025\)](#) observe in their review for the Fed’s 2025 strategy assessment, “no clear set of best practices for communicating uncertainty and risks to the public has emerged.” We argue that this debate rests on a false dichotomy: scenarios and predictive densities are natural complements, and the challenge is to combine their strengths rather than to choose between them. This complementarity reflects a deeper reality: central banks face “deep uncertainty”—about the model itself, not only about shocks—so the two tools approximate different objects and serve different purposes.

We formalize this distinction by treating baseline forecasts and scenarios as *conditional* predictive densities: each is a forecast of the economy conditional on a specific set of assumptions and a specific model. The reference predictive distribution, instead, aims at approximating the *unconditional* distribution—the probability of all outcomes, not conditioned on any particular model or set of assumptions. The Scenario Synthesis then asks whether this unconditional distribution can be approximated as a weighted combination of conditional views and, if so, what weights each scenario should receive.

We apply the Scenario Synthesis framework to Tealbook data from two contrasting episodes—the turbulent pre-crisis environment of December 2007 and the more stable conditions of December 2018. In 2007, the scenario set provided only limited coverage of downside risks to GDP growth (the Credit Crunch scenario captured meaningful probability mass, but

the 2008–09 outcome had an even fatter left tail), and alternative reference densities implied sharply different weights. In 2018, by contrast, the scenario set was much closer to spanning the relevant risks and yielded an internally coherent assessment. More broadly, the exercises show how the Synthesis can be used to characterize the balance of risks, assess whether the available scenario set spans the relevant sources of macroeconomic risk, and, when it does not, guide the design of improved scenarios.

Beyond staff forecasting, the framework helps structure disagreement within policy committees. Under deep uncertainty, policymakers may disagree about shocks, transmission mechanisms, or the weighting of risks. By representing alternative views as conditional predictive densities and aggregating them into a committee-level assessment, the Scenario Synthesis makes these sources of disagreement explicit—and thus serves not only to assign probabilities to scenarios but to discipline policy deliberation.

The rest of the paper is organized as follows. Sections 2 and 3 trace the development of risk analysis at the Federal Reserve and other central banks. Section 4 discusses the disconnect between scenarios and predictive densities. Sections 5 and 6 present the framework and the Scenario Synthesis. Section 7 contains the empirical applications to the 2007 and 2018 Tealbooks. Sections 8, 9, and 10 discuss the broader implications for scenario design, alternative economic models, and policy deliberation. Section 11 concludes.

2 Risk Analysis at the Federal Reserve: A Historical Account

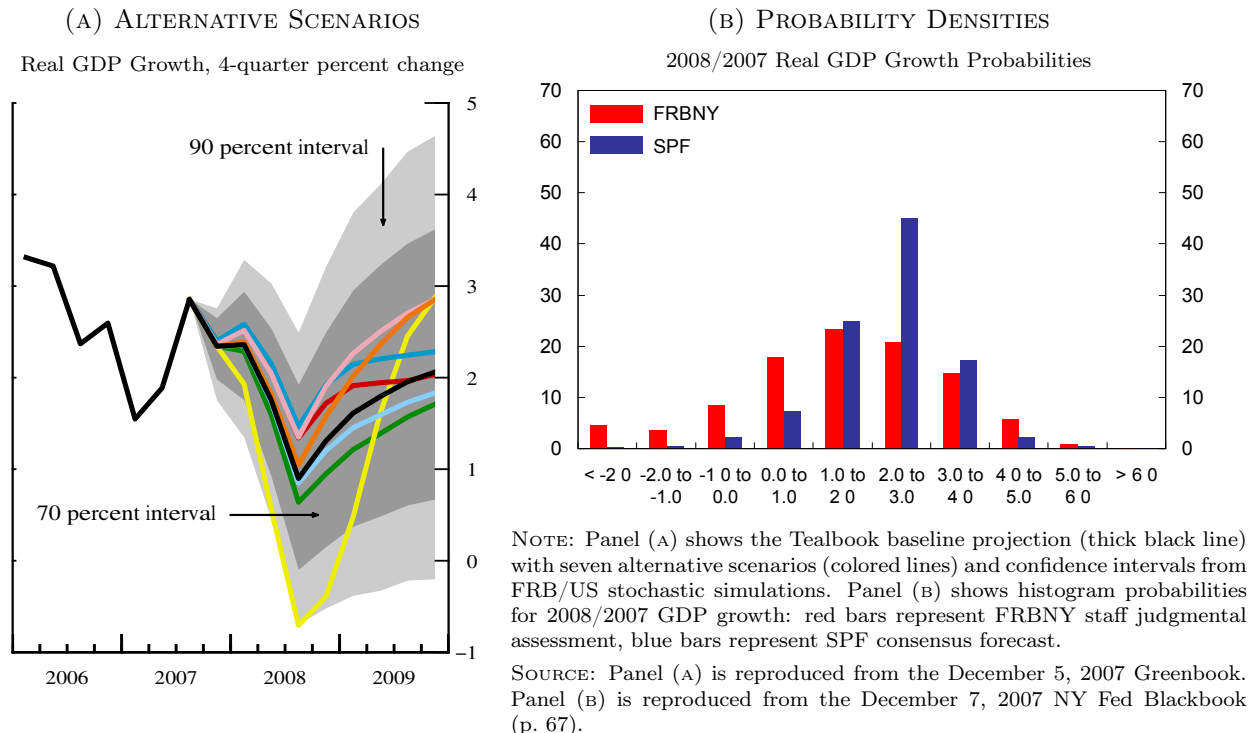
The systematic monitoring of macroeconomic risk by central banks dates back only to the 1990s. The practice became more important after the Global Financial Crisis and has gained further prominence since the COVID-19 pandemic. Yet documentation of these practices is scarce and fragmented. In the spirit of Sims (2002), this section reconstructs the Federal Reserve’s risk-analysis practices and traces their evolution, drawing on publicly available material from the Tealbook, prepared by Federal Reserve Board (FRB) staff for the FOMC,² and the Blackbook, prepared by the Federal Reserve Bank of New York for its President. Both documents are classified when prepared and released with a five-year lag.

Two broad approaches have emerged. The first identifies salient economic risks and uses models to produce alternative scenarios—conditional forecasts describing what happens if those risks materialize. The second constructs predictive distributions over possible future outcomes using surveys, judgment, or statistical models. This distinction is central to the paper. The Introduction showed scenarios and predictive densities side by side for December

²The Tealbook replaced the Greenbook, which served the same function from the 1960s through 2009. To simplify exposition, we use “Tealbook” throughout, while preserving “Greenbook” in archival titles, source notes, and citations.

2018 (Figure 1), a useful “normal-times” benchmark. Here we add an earlier, more revealing episode: December 2007, one year before the Global Financial Crisis. Figure 2 displays the Tealbook scenarios alongside the predictive densities from the NY Fed Blackbook and the Survey of Professional Forecasters. The contrast between these assessments lies at the heart of this paper. We begin with scenarios, turn to predictive densities, and then discuss the gap between them.

Figure 2: SCENARIOS AND PREDICTIVE DENSITIES FOR DECEMBER 2007



2.1 The Baseline

The natural starting point is the Tealbook baseline, because both scenarios and predictive densities measure risk around it. The baseline is the FRB staff’s judgmental projection under a particular set of conditioning assumptions, and it anchors the Fed’s risk analysis. The Tealbook baseline is not the forecast of any single model. It combines sectoral expertise, incoming data, conditioning assumptions, econometric models, financial-market information, outside forecasts, and judgmental add-factors (Fischer, 2017).³ When, in Section 5, we associate the baseline with a “baseline model” M_0 , we mean the baseline process that anchors the staff projection, not the mechanical output of FRB/US.

³Public Tealbooks explicitly contrast the staff forecast with model forecasts. In the December 2018 “Alternative Model Forecasts” exhibit, for example, the staff projection for 2019 real GDP growth is 2.4 percent versus 1.1 percent for FRB/US (Federal Reserve Board, 2018, p. 91).

Uncertainty around the baseline is summarized in several ways. The confidence intervals in Figures 1(A) and 2(A) come from stochastic simulations of FRB/US, the Federal Reserve Board’s workhorse semi-structural model (Brayton and Tinsley, 1996; Brayton et al., 1997, 2014). FRB/US, a large-scale model in the Cowles Foundation tradition, contains hundreds of equations covering consumption, investment, housing, labor markets, prices, fiscal and monetary policy, and international linkages, combining micro-founded and reduced-form elements; its main strengths are comprehensiveness and interpretability. Since its introduction in 1996, it has been used extensively to communicate ideas to the Board and the FOMC and to produce domestic alternative scenarios (Tetlow and Ironside, 2007). The Tealbook also reports intervals derived from historical forecast errors, which provide an alternative way to summarize uncertainty around the baseline projection.⁴

Both approaches have important limitations. Because FRB/US is most often used in a (conditionally) linearized form, its simulated uncertainty via stochastic simulations is close to symmetric.⁵ Historical forecast-error intervals can be more asymmetric, but both approaches are time-invariant: they adjust only after shocks appear in the data, and cannot respond to contemporaneous indicators of risk such as current financial conditions (Bauer et al., 2025). For this reason, the Fed’s broader assessment of risk complements baseline intervals with alternative scenarios and time-varying predictive densities, discussed next.

2.2 Monitoring Risk Using Scenarios

A scenario is a forecast produced under alternative assumptions and, sometimes, an alternative model relative to the baseline. The model can be a formal econometric model or a “mental model”—a qualitative narrative translated into a quantitative path through expert judgment. Each scenario is typically accompanied by a narrative explaining the economic mechanisms: which transmission channels are activated, why the assumptions lead to different outcomes, and how the scenario unfolds.

SCENARIOS AT THE FEDERAL RESERVE BOARD. At the Federal Reserve Board, combining baseline projections with alternative scenarios has been standard since the mid-1990s. Until 2009 these scenarios appeared in the U.S. outlook section of the Greenbook; since 2010 they have appeared in the Tealbook’s “Risks and Uncertainty” section. The scenario set changes

⁴See the exhibits “Prediction Intervals Derived from Historical Tealbook Forecast Errors” (since 2017) and “Selected Tealbook Projections and 70 Percent Confidence Intervals Derived from Historical Tealbook Forecast Errors and FRB/US Simulations” (since 2004).

⁵FRB/US is not itself a linear model, and even its linearized version respects the effective lower bound. The Board’s stochastic-simulation procedure has been refined to generate more realistic recession depth, duration, and left-tail behavior by González-Astudillo and Vilán (2019), who introduced a Markov-switching residual procedure. The resulting asymmetries are modest for GDP growth and inflation but more pronounced for the fed funds rate and unemployment.

from round to round, depending on which risks staff consider most salient.

Early on, the Board constructed virtually all domestic scenarios using FRB/US (Herbst et al., 2026). It later expanded the model set to include SIGMA (Erceg et al., 2006), a multicountry DSGE model for international scenarios; EDO (Edge et al., 2008), an estimated DSGE model of the U.S. economy; and models with richer financial sectors, such as Gertler–Karadi-type credit-accelerator frameworks (Gertler and Karadi, 2011).

SCENARIOS AT THE NY FED. Scenarios have appeared in the Blackbook since at least 2005, historically built from expert judgment. By the December 2017 Blackbook—the last publicly available one—the process had become more model-based: scenarios were built using a medium-scale DSGE model (Del Negro and Giannoni, 2017) combined with a large BVAR (Crump et al., 2025) estimated on the 31 variables in the Tealbook judged most important. The DSGE supplies a structural interpretation, while the BVAR, though reduced-form, is closely tied to the structural models and serves as a cross-check and a flexible tool for conditional forecasting.

Unlike the FRB staff, FRBNY staff also assigned judgmental weights to each scenario and combined them into a full probability distribution. This anticipates a key element of the Scenario Synthesis—the explicit weighting of scenarios—except that the Blackbook weights were judgmental, whereas the Synthesis chooses them statistically via a concordance criterion.

LIMITATIONS OF SCENARIOS. Scenarios provide narrative structure—economic mechanisms, transmission channels, and causal stories policymakers can reason about—but they offer no probability assessment and no formal metric to evaluate them as a set. Two questions therefore go unanswered: *What is the probability of each scenario?* and *How well does the scenario set capture current macroeconomic risks?* These are not merely conceptual concerns: when presented with an adverse scenario, policymakers naturally ask whether it warrants a response or a contingency plan, and the answer depends on information the scenario itself does not provide.

A second limitation is scenario selection. From round to round there is often little continuity: new scenarios appear, old ones disappear, and most persist only briefly. There is typically no formal criterion for which risks are explored, nor a systematic record of why a scenario is introduced or removed, making it hard to compare assessments over time or to tell whether changes reflect the economy or judgment. This critique is not unique to scenarios: statistical risk tools are themselves evolving research products rather than fixed reference objects.⁶

⁶For example, in April 2019 the staff introduced a new two-year macro-risk exhibit because the one-year “Time-Varying Macroeconomic Risk” exhibit could show balanced risks while staff perceived larger downside

A third limitation is that scenarios are typically reported as point paths, with no predictive distribution; and even when models characterize uncertainty around them, those models are usually linear or log-linear, so that uncertainty within each scenario is typically symmetric and time-invariant, with little scope for fat tails, skewness, or state-dependent amplification. In practice, the risk assessment comes primarily from the choice of scenarios and the assumptions imposed on them, not the models’ internal dynamic.

In sum, scenarios provide rich narratives but remain deeply judgmental—in which risks are explored, the assumptions defining them, and the models quantifying them. What is missing is a systematic way to evaluate whether the scenario set as a whole captures the relevant risks. The Scenario Synthesis of [Adrian et al. \(2025\)](#) fills this gap by leveraging information from predictive densities.

2.3 Monitoring Risk Using Predictive Densities

A predictive density assigns probabilities to all possible future realizations of a variable, describing the full range of outcomes—including particularly favorable or adverse ones—rather than delivering only a point forecast. A point forecast compresses the distribution into a single number that is often asked to carry two objects at once: what is most likely and what matters for decision-making. Econometricians define a point forecast as the minimizer of expected loss under a given loss function ([Granger and Machina, 2006](#); [Elliott and Timmermann, 2008](#); [Gneiting, 2011](#)).⁷ Predictive densities separate the two: the density describes the risks, and the loss function weights them.

THE SURVEY OF PROFESSIONAL FORECASTERS. The Survey of Professional Forecasters (SPF) is the longest-running continuous source of predictive densities for the U.S. economy. Launched in 1968 by the American Statistical Association and the National Bureau of Economic Research, and administered by the Philadelphia Fed since 1990, it asks forecasters for point forecasts and for probabilities assigned to bins covering possible GDP growth realizations. Averaging these probabilities across respondents yields a consensus density summarizing the community’s collective assessment. The blue bars in [Figure 2b](#) show this consensus density for December 2007.

PREDICTIVE DENSITIES AT THE NY FED. The New York Fed has produced predictive densities in its internal Blackbook since at least 2005, though the Blackbook’s existence and methodology have been sparsely documented even within the Federal Reserve System. From

risks further out ([Federal Reserve Board, 2019a](#), pp. 75, 78). During COVID-19, the staff suspended the two-year exhibit, judging its distributions unreliable.

⁷Expected loss under a predictive density is the statistical concept of risk. In this paper we use a broader notion of risk: the probabilistic assessment of possible future events.

May 2007 to March 2010, the New York Fed Research Group presented density forecasts using histograms in a format similar to the SPF, plotted side by side with SPF distributions. Figure 2(B) shows an example from the December 2007 Blackbook: FRBNY staff assigned non-negligible probability mass to negative GDP growth outcomes, while the SPF remained centered on positive growth—illustrating how internal density forecasts can highlight downside tail risks that may be less visible in surveys. The next subsection discusses why the NY Fed’s assessment differed so sharply from the SPF.

This exercise later became more formalized. In March 2017, the Blackbook featured SEPIA (Summary of Economic Projections with Individual Assessment of Uncertainty), an internal survey of about 13 Research Group economists. Each economist assigned probabilities to outcome bins, and averaging these probabilities produced a consensus density analogous to the SPF. SEPIA thus provided an internal, survey-based assessment of macroeconomic risks that complemented the Blackbook’s other risk tools.

WHY THE NY FED’S RISK ASSESSMENT DIFFERED FROM THE SPF. The contrast in Figure 2(B) is striking. One year before the Great Recession, the NY Fed assigned roughly 15% probability to negative GDP growth, while the SPF consensus remained near-symmetric with limited downside risk—a difference reflecting the NY Fed’s distinct informational environment.

The NY Fed houses the markets desk, which executes open-market operations and maintains daily contact with primary dealers, and it supervises the largest bank holding companies, giving it direct visibility into balance sheets, leverage, and funding conditions. This proximity to Wall Street gave it real-time intelligence about emerging stress in money markets, mortgage-backed securities, and interbank lending that was not captured by the macroeconomic data available to the SPF.

Remarks by NY Fed president Timothy Geithner at the August 7, 2007 FOMC meeting illustrate this advantage. He warned that the turmoil had “the potential to cause substantial damage through the effects on asset prices, market liquidity, and credit; through the potential failure of more-consequential financial institutions; and through a general erosion of confidence,” and noted that “a lot of that risk has gone to leveraged funds that have much less capacity to absorb this kind of shock” ([Federal Open Market Committee, 2007](#), p. 60). Cautioning more inflation-focused colleagues, he added that waiting for clearer evidence on aggregate demand would make them “inevitably be too late” ([Federal Open Market Committee, 2007](#), p. 62). The December 2007 Blackbook confirms this picture: FRBNY staff wrote that “downside risks to growth have increased significantly,” with credit spreads at post-2001 highs and interbank markets deteriorating beyond their August peaks ([Federal Reserve Bank of New York, 2007](#), p. 20). The NBER later dated the recession to that month.

In short, the NY Fed used financial conditions—credit spreads, leverage, and funding markets—as a signal of downside risk that the SPF, relying primarily on traditional macroeconomic data, did not capture.

FROM JUDGMENT TO ECONOMETRICS: THE GROWTH-AT-RISK FRAMEWORK. A major development in the analysis of macroeconomic risk is the Growth-at-Risk (GaR) framework of [Adrian et al. \(2019\)](#). Developed at the New York Fed to update forecast distributions at the roughly six- to seven-week FOMC frequency, it formalizes the link between financial conditions and macroeconomic risk through the entire conditional distribution of future GDP growth, not just the mean. Tighter financial conditions sharply worsen downside risk while leaving the upper part of the distribution comparatively stable: lower GDP-growth quantiles respond strongly to credit spreads, equity volatility, and leverage, upper quantiles much less. Financial conditions thus matter mainly for the tails.⁸ This asymmetry is consistent with nonlinear financial-accelerator mechanisms, discussed in [Section 9](#), in which deteriorating balance sheets, widening risk premia, and credit contraction amplify downside risk.

At the Federal Reserve Board, this line of work entered the Tealbook in 2017 through the Time-Varying Macroeconomic Risk (TVMR) exhibit ([Figure 1b](#)). TVMR reports a predictive distribution for Tealbook forecast errors, with dispersion and asymmetry driven by indicators of real activity, inflation, economic uncertainty, and financial conditions (see [Engstrom and Gonzalez-Astudillo, 2017](#)). In stress episodes, it implies wider downside risks; in more tranquil periods, such as late 2018, the distribution is closer to symmetric and centered around the baseline.

At the New York Fed, the same line of work is reflected in Outlook-at-Risk (OaR), published since 2023. OaR provides one-year-ahead predictive distributions for key macroeconomic variables and builds on extensions of the Growth-at-Risk framework by [Adams et al. \(2021\)](#) and [Boyarchenko et al. \(2026\)](#), who construct risk distributions around consensus forecasts for GDP growth, unemployment, and inflation.⁹

Taken together, TVMR and OaR illustrate how central banks have increasingly moved from judgment-based assessments of tail risk toward transparent and replicable econometric measurement.

⁸The original framework ([Adrian et al., 2019](#)), which used financial conditions as the main driver, was later extended to economic uncertainty ([Engstrom and Gonzalez-Astudillo, 2017](#); [Keijsers and van Dijk, 2025](#)) and geopolitical uncertainty ([Anobile et al., 2025](#)).

⁹Related predictive-density methods are now an active area of research throughout the Federal Reserve System, with recent applications to unemployment, inflation, corporate spreads, tail risks in BVARs, financial shocks, and investment-at-risk ([Kiley, 2022](#); [López-Salido and Loria, 2024](#); [Caldara et al., 2021](#); [Carriero et al., 2024](#); [Scotti, 2023](#); [Amburgey and McCracken, 2025](#)).

SUMMARY OF FEDERAL RESERVE PRACTICE. Risk has been monitored at the Fed through two main quantitative approaches. Scenarios have appeared in the Tealbook since 1995. Predictive densities have been elicited through the SPF (since 1968), built by the NY Fed Blackbook by weighting scenarios judgmentally (since about 2000), surveyed via SEPIA (2017), and produced statistically by the Board’s TVMR exhibit using financial conditions (2017). Crucially, the NY Fed’s approach of building densities *from* scenarios is the essential intuition the Scenario Synthesis formalizes.

There is also a third channel, not reviewed separately above because it is not a quantitative tool in the same sense: the staff’s judgmental *Assessment of Risks*. This assessment provides a qualitative characterization of the balance of risks and uncertainty around the baseline and often carries substantial weight in policy deliberations. It may draw on scenarios, predictive densities, and other statistical exhibits, but it is not reducible to any one of them. In July 2019, for example, staff judged risks to be tilted to the downside even though the one-year-ahead time-varying-risk estimates were “not unusually wide or skewed” (Federal Reserve Board, 2019b, p. 69). Scenario Synthesis can be viewed as a structured way to make this judgmental assessment explicit, quantitative, and reproducible.

This historical reconstruction shows that the Federal Reserve has developed a rich but fragmented toolkit for assessing macroeconomic risk. We next place these practices in a broader international context.

3 Risk Analysis at Other Central Banks and Institutions

The Federal Reserve’s experience is part of a broader global effort. Other central banks, international institutions, and private firms have developed their own approaches—sometimes independently, sometimes influenced by the Fed. We organize the discussion around the three building blocks introduced above: baseline projections, scenarios, and predictive densities.

3.1 Baselines Across Institutions

Across central banks, baseline projections are typically produced with a combination of structural and reduced-form models. We briefly illustrate this with two prominent examples.

EUROPEAN CENTRAL BANK. The ECB’s baseline projections are staff projections supported by a suite of models. Two central components are ECB-Base (Angelini et al., 2019), a large semi-structural model for the euro area, and ECB-MC (Angelini et al., 2026), a multi-country semi-structural model covering euro-area member countries; both are designed in the same broad tradition as FRB/US. The third central component is the New Area-Wide Model II (NAWM II), a micro-founded open-economy DSGE model estimated on euro-area data using Bayesian methods (Christoffel et al., 2008; Coenen et al., 2018). The ECB also

maintains a range of satellite models—including small DSGE models, small semi-structural models, and time-series models—that provide cross-checks on the baseline projection (see Ciccarelli et al., 2024, for a description of satellite models).

BANK OF ENGLAND. The Bank of England’s baseline model is COMPASS, a medium-scale open-economy New Keynesian DSGE model estimated on UK data (Burgess et al., 2013). Adopted in 2011, it was designed as a central organizing model within a forecasting platform that also includes more than forty supplementary models. A medium-scale Bayesian VAR serves as its reduced-form companion (Domit et al., 2019), matching the DSGE’s coverage while being more robust to structural misspecification. In practice, however, the Bank’s forecast infrastructure relies heavily on judgment and supplementary models, so the published baseline is best understood—as at the Federal Reserve—as a judgmental staff product rather than the output of a single model.¹⁰

3.2 Scenarios Across Institutions

The Federal Reserve Board has the longest continuous tradition of systematic scenario analysis, producing scenarios in the Tealbook since at least 1995. Elsewhere, scenario analysis was historically reserved for exceptional circumstances or internal stress tests rather than regular communication. That changed after the Global Financial Crisis, which exposed the limits of single-baseline forecasting under deep uncertainty. Since then, scenarios have spread widely, though formalization, publication frequency, and integration with probabilistic tools vary considerably.

EUROPEAN CENTRAL BANK. Before the pandemic, the ECB used scenarios only occasionally. COVID-19 was a turning point: in June 2020, staff first built three pandemic scenarios—mild, severe, and very severe—and only then chose which would become the published baseline, inverting the usual logic and treating the baseline as one element of a broader scenario space. Scenarios have since become regular features of ECB projections and communication, used extensively during the Russia-Ukraine energy shock and, more recently, for U.S. tariff risks and Middle East geopolitical risks. In each case, they structured deliberations around defined contingencies and communicated uncertainty to the public.

BANK OF ENGLAND. The Bank historically communicated uncertainty through fan charts: probability distributions around the central projection, calibrated from forecast errors and judgment, that became a widely imitated template. Following the Bernanke (2024) review,

¹⁰In his independent review, Bernanke (2024) recommended replacing or substantially revamping COMPASS, citing limitations in its monetary-transmission channels. Consistent with this modernization agenda, the Bank of England released an updated medium-scale DSGE model for the UK economy in 2025 (Albuquerque et al., 2025).

which criticized fan charts as difficult to interpret and weak in conveying the narratives underlying risk assessments, the Bank began publishing explicit scenarios in November 2025; in the April 2026 *Monetary Policy Report*, amid exceptional uncertainty about the Middle East energy shock, it even replaced the baseline with three scenarios.¹¹ This marks a shift from a statistical summary of forecast risk to a narrative account of the paths the Monetary Policy Committee (MPC) considers most relevant.¹²

OTHER CENTRAL BANKS. In April 2025, the Bank of Canada departed from its conventional single-baseline forecast, communicating its outlook *exclusively* through two scenarios without assigning relative probabilities, on the grounds that deep uncertainty about U.S. trade policy made any baseline spuriously precise. The Riksbank has published scenarios regularly since 2023 and is unusual in reporting an explicit interest-rate path for each, thereby clarifying its reaction function under alternative outcomes. The Central Bank of Armenia has gone further, adopting a purely scenario-based framework and replacing the single baseline altogether.

INTERNATIONAL INSTITUTIONS. The IMF’s *World Economic Outlook* has published scenarios around its baseline forecast for nearly a decade, covering risks such as financial stress, trade tensions, commodity-price shocks, and geopolitical disruptions. These scenarios are typically presented as deviations from the global baseline, built with the IMF’s global macroeconomic model, and accompanied by narratives describing the shock assumptions and transmission channels. The OECD similarly publishes risk assessments and alternative scenarios in its *Economic Outlook*. Both institutions have also developed Growth-at-Risk frameworks, so scenario-based and distributional approaches increasingly coexist within the same institution, closely mirroring the parallel structure observed in the Tealbook.

PRIVATE INSTITUTIONS. Private forecasting firms—Oxford Economics, Moody’s Analytics, S&P Global, and others—have long offered scenario analysis as a commercial product. Their scenarios provide probability-weighted alternative paths for the global economy, individual countries, and sectors, and are used for portfolio stress testing, strategic planning, and risk management. Unlike most central-bank scenarios, they often carry explicit probability weights, reflecting client demand for a single risk-adjusted metric.

SUPERVISORY BANK STRESS TESTING. A separate tradition applies scenario-based stress testing to financial institutions. Originating at the IMF in the late 1990s, this practice subjects bank balance sheets to adverse macroeconomic scenarios—severe recessions, asset-

¹¹The May 2020 projection was also presented as a “scenario,” but then, between May 2020 and November 2025, the Bank did not publish any other scenario.

¹²The shift is one of emphasis, not abandonment: fan charts remain in the public databank, now calibrated solely on past forecast errors.

price declines, and credit-spread spikes—and assesses the resulting capital shortfalls (see [Adrian et al., 2020](#)). After the Global Financial Crisis, it became a regulatory requirement in most major jurisdictions. This tradition differs from central-bank forecasting because it focuses on institution-level solvency rather than economy-wide dynamics, but it rests on the same logic: concrete, narratively coherent scenarios are often more decision-relevant than abstract probability distributions alone.

3.3 Predictive Densities Across Institutions

JUDGMENTAL APPROACHES. The Bank of England was the first central bank to communicate macroeconomic risk explicitly through predictive densities. Its fan charts, introduced in the *Inflation Report* in 1996, reported quantiles of the predictive distribution of inflation—and later GDP growth and unemployment—using models and judgment ([Britton et al., 1998](#)). The European Central Bank has employed a different judgmental approach through the *Quantitative Risk Assessment* (QRA), an internal survey-based tool that summarizes ECB staff views on the main risk events surrounding the baseline, including their upside or downside nature and their implications for uncertainty and skewness ([Nickel et al., 2025](#)). The QRA is not a stand-alone survey reference like the SPF, but one element of the internal risk-assessment process.

SURVEY-BASED APPROACHES. The ECB has operated a Survey of Professional Forecasters since its inception, modeled in part on the Philadelphia Fed SPF. It asks respondents for probabilistic assessments of GDP growth, inflation, and unemployment, providing survey-based predictive distributions that complement staff risk assessments. The ECB also draws on other surveys, including the Survey of Monetary Analysts and the Consumer Expectations Survey, which contain probabilistic information relevant for assessing expectations and risks. Similar surveys are conducted by several other central banks, including the Bank of Japan and the Bank of Canada.

STATISTICAL APPROACHES. The ECB also reports uncertainty ranges based on past projection errors, a backward-looking benchmark that complements forward-looking density tools. The most important recent development in quantitative risk analysis has been the worldwide adoption of Growth-at-Risk (GaR) models. The IMF introduced GaR in the October 2017 *Global Financial Stability Report* ([IMF, 2017](#)), applying it across 21 economies; [Adrian et al. \(2022\)](#) later developed its term structure, documenting that loose financial conditions reduce near-term downside risk but increase it at longer horizons. GaR-based forecasts now complement risk toolkits at the ECB ([Figueres and Jarociński, 2020](#); [Nickel et al., 2025](#); [Lenza et al., 2025](#)), the Bank of England ([Aikman et al., 2019](#); [Lloyd et al., 2022, 2024](#); [Anesti et al., 2023](#); [Eguren-Martin et al., 2024](#)), the Banque de France ([Jondeau](#)

et al., 2022; Ferrara et al., 2022; Lhuissier, 2022), and several other central banks.¹³

Evaluating these tools against ease of communication, narrative content, forward-looking nature, and probability assessment, Nickel et al. (2025) find that statistical models score high on probability assessment and forward-looking content but low on narrative—the mirror image of scenarios. Across institutions, then, the toolkit for risk analysis is broad but fragmented, with different tools emphasizing different dimensions of risk. We now turn to the disconnect this produces.

4 The Disconnect Between Scenarios and Predictive Densities

Across institutions, scenarios and predictive densities coexist within the same forecasting and policy process, yet they are rarely embedded in a common framework. They are often produced using different models and information sets: scenarios deliver narratives without probabilities, while predictive densities deliver probabilistic risk assessments with limited economic interpretation.

The disconnect persists because their strengths are complementary. Scenarios are rich in narrative and broad in scope—large structural models, causal stories, joint forecasts—but limited in their treatment of risk: linear or linearized models make uncertainty largely symmetric and time-invariant, with little room for fat tails, skewness, or state-dependent amplification. Predictive densities are richer in risk—capturing asymmetry, nonlinearity, and time-varying tails—but reduced-form and narrower in scope, offering less on mechanisms and typically covering fewer variables.

Scenario analysis also serves distinct purposes, which shape what a scenario set should contain. Some scenarios communicate the balance of risks around the outlook, as in the Tealbook’s “Risks and Uncertainty” chapter. Others communicate the central bank’s reaction function, as Bernanke (2025) and Garga et al. (2025) emphasize, making the policy path itself an object of interest. Supervisory stress tests, by contrast, deliberately probe the tail, whereas monetary-policy scenarios typically capture large but not extreme risks. Scenario Synthesis is agnostic across these uses: it evaluates any scenario set against the reference density that encodes the relevant notion of risk.

The balance is shifting. As macroeconomic risks have intensified—through the Global Financial Crisis, the COVID-19 pandemic, the Russia-Ukraine war, and tariff shocks—institutions that previously emphasized predictive densities have increasingly turned to scenarios. The ECB’s expansion since 2020, the Bank of England’s published scenarios in 2025,

¹³See, among others, applications at the Bank of Italy (Alessandri et al., 2019), the Central Bank of Ireland (O’Brien and Wosser, 2021), the Bundesbank (Plaasch and Röthig, 2025), Banco de España (Galán Camacho, 2020), the Bank of Japan (Makabe and Norimasa, 2022), and Norges Bank (Bowe et al., 2023).

and [Bernanke](#)’s proposal that the Fed publish scenarios after each FOMC meeting all point in this direction. Surveying 25 central banks, [Bell et al. \(2026\)](#) find that the share using scenario analysis rose from about 25% before the pandemic to about 45% in 2025.

As central banks increasingly use scenarios and predictive densities side by side, the risk of conflicting signals grows. The natural question is how to combine the narrative richness of scenarios with the quantitative discipline of predictive densities. The Scenario Synthesis does so by choosing weights on the baseline and alternative scenarios so that their mixture best approximates a reference predictive distribution under a concordance criterion. The resulting Synthesis links the structural and communicative strengths of scenarios with the probabilistic rigor of densities.

The framework also connects to the debate on risk communication. [Bernanke \(2024\)](#) recommended de-emphasizing fan charts in favor of qualitative risk descriptions and narrative scenarios. Scenario Synthesis bridges these approaches by preserving scenario narratives while attaching transparent weights, and is complementary to approaches that attach policy-rate paths to scenarios ([Laxton et al., 2025](#); [Bernanke, 2025](#); [Garga et al., 2025](#)).

A central implication is that weights depend on the reference density. Different predictive densities—that is, different assessments of risk—imply different scenario weights. When downside risk is small, adverse scenarios receive little weight and the set may appear adequate; when downside risk is large, adverse scenarios receive greater weight and shortcomings become visible. Sections 7.1 and 7.2 show that this dependency matters in practice.

The Scenario Synthesis framework is beginning to enter policy practice. At the Banque de France, [Lhuissier \(2026\)](#) applies it to the Eurosystem’s June 2025 projections, using the ECB Survey of Professional Forecasters as the reference distribution to construct a synthesis-based optimal monetary policy path. At the Bank of England, a cross-divisional team applies the framework to the three scenarios in the April 2026 *Monetary Policy Report*. At the three-year horizon, they find that the Synthesis captures about 87% of the inflation risk in the Decision Maker Panel survey and spans roughly the 25th–75th percentiles of the one-year-ahead market-implied Bank Rate distribution. In Section 7, we apply the framework to Federal Reserve Tealbook scenarios in December 2007 and December 2018.

5 The Framework

The economy is complex, nonstationary, and possibly indeterminate. Its dynamics depend on observable variables y (GDP, inflation, unemployment, etc.) and on a large set of largely hidden conditions z (financial frictions, expectations, global conditions, regime indicators, etc.) that no single model can fully capture. Central bank forecasting is therefore an inference

problem under deep uncertainty: policymakers cannot agree on a single representation of the economy, nor do they observe the full economic state directly; they observe only aggregates such as GDP, inflation, unemployment, spreads, and surveys, together with a finite historical sample.

As we have documented in Sections 2 and 3, central banks confront and communicate this uncertainty using a multi-layered forecasting apparatus. The toolkit typically includes a *baseline projection*, $p_0(y) = p(y \mid A_0, M_0)$, that delivers a coherent economic narrative under a set of conditioning assumptions A_0 describing “normal” states, and the baseline model M_0 ; a collection of *alternative scenarios*, $p_j(y) = p(y \mid A_j, M_j)$, $j = 1, \dots, J$, that examine specific risks under alternative conditioning assumptions A_j , and model M_j chosen to capture the mechanisms deemed relevant for the risk under consideration; and a *reference predictive density*, $p(y)$, that aims to characterize the full range of possible outcomes regardless of which conditioning assumptions or model happen to be relevant via flexible statistical methods and/or expert judgment. We discuss each in turn before showing, in the next section, how the Scenario Synthesis links them.

5.1 Baseline projection

As documented in Sections 2.2 and 3.1, most major central banks maintain a large-scale structural model that plays the role of the baseline model (M_0), providing a coherent economic narrative for the institution’s point forecast. Producing the baseline forecast combines M_0 with baseline assumptions A_0 about variables, shocks, and mechanisms treated as exogenous: paths for global demand, foreign output, commodity prices, fiscal policy, and shock distributions. The baseline predictive density is thus

$$p_0(y) \equiv p(y \mid A_0, M_0).$$

Baseline models are typically large macroeconomic systems, with hundreds of equations covering consumption, investment, housing, labor markets, prices, fiscal and monetary policy, and international linkages. They combine micro-founded components with reduced-form elements, such as empirical Phillips curves, and build in several practical approximations. First, they rely on linear or log-linear approximations around a “normal” baseline, limiting their ability to capture strong nonlinearities, regime shifts, and binding constraints. Second, shock variances are typically fixed, so risk is treated as time-invariant and effects are often symmetric. Third, financial frictions—runs, fire sales, and balance-sheet interactions—are represented only partially. Fourth, they are optimized for “normal” business-cycle environments rather than stress episodes or structural breaks. Even so, they remain central: they provide a shared normal-times benchmark, support causal narratives, and organize communication and decomposition exercises.

5.2 Alternative scenarios

Central banks complement the baseline with alternative scenarios designed to explore specific risks or alternative mechanisms. A scenario is defined by a particular combination of assumptions A_j and model M_j :

$$p_j(y) = p(y \mid A_j, M_j), \quad j = 1, \dots, J.$$

Here $A_j \neq A_0$, while M_j may differ from M_0 .

The assumptions A_j are understood broadly. They include paths or ranges for exogenous variables, shock distributions, qualitative states, and the assumed monetary policy reaction function. Examples include “oil prices are elevated,” “financial stress is high,” or “foreign demand is weak.” These assumptions define regions of the state space rather than points, so the associated conditional densities can receive non-negligible weight even when the scenario is communicated as a central path. The policy assumption also matters for interpretation. A scenario is a conditional density given an assumed policy response, so the synthesis weights are conditional on those policy assumptions as well; see Sections 9 and 10.

Because M_0 is a local approximation around normal conditions, assumptions and model structure are often intertwined. In stressed or nonlinear environments, the local approximation may become unreliable and a different model M_j may be needed: a credit-accelerator block for severe financial stress, a nonlinear Phillips curve for large supply shocks. By contrast, M_0 can handle a mild foreign-growth slowdown. The case $M_j \neq M_0$ is not confined to extreme stress. Some first-order risks concern structural change—for example, greater intrinsic persistence of wage- and price-setting after high inflation—and are naturally captured by a model M_j that better reflects the mechanism. The principle is to match the model to the assumptions under consideration.

Scenarios examine regions of the state space relevant for policy but not central enough for A_0 . They are not competing forecasts but conditional “if-then” explorations.¹⁴ They help policymakers reason under deep uncertainty.¹⁵

5.3 Reference predictive density

The reference predictive density $p(y)$ serves a distinct purpose from both the baseline and the scenarios: it approximates the unconditional predictive distribution as accurately as possible, without necessarily providing a structural narrative. Researchers typically build it with flexible reduced-form methods designed for predictive accuracy—high-dimensional

¹⁴As Crump et al. (2025) emphasize, conditional forecasting links the central-bank scenario tradition to structural analysis based on estimated impulse responses.

¹⁵We use “deep uncertainty” in the sense of the operations-research literature (Walker et al., 2013): situations in which decision-makers cannot agree on the model, the relevant outcomes, or the probability distributions over inputs. We develop this notion and its implications for policy in Section 10.

quantile regressions, machine learning, Bayesian model averaging, time-varying-parameter models, and Growth-at-Risk frameworks. Its strengths are out-of-sample accuracy, nonlinearities, time-varying risk, and large information sets; its weaknesses are limited structural interpretation and communicability.¹⁶

Two points bear on how the synthesis should be read. First, the framework is agnostic about which reference is used: $p(y)$ can be any density the user regards as a good unconditional predictive distribution—a Growth-at-Risk model, a time-varying-parameter model, a survey-based density, or a judgmental assessment. If a user believes a statistical reference understates nonlinear “dark-corner” tail risks—for example, convex-Phillips-curve or credibility-loss dynamics—then the user should supply a reference that does not. A scenario designed to capture such a risk will receive low weight against a reference that underweights the tail, but this signals a limitation of the reference rather than an implausible scenario—exactly the type of diagnostic the synthesis is meant to surface.

Second, when policymakers disagree about the reference, the response is not to force a single choice but to report the synthesis under each candidate, treating the spread of weights as a robustness band. This is the approach we take in Section 7.

6 The Scenario Synthesis

The baseline model M_0 provides interpretability and causal structure, but delivers an incomplete characterization of the range of outcome omitting macro-financial linkages, nonlinearities, and time-varying risk that shape the tails. Scenarios complement it by articulating “what-if” risks not embedded in the baseline. The reference density $p(y)$ delivers a richer characterization of the distribution of outcomes, but at the cost of interpretability: we know that elevated credit growth fattens the left tail, but the mechanisms remain debated.

Scenario Synthesis connects what we have learned empirically (the reference) with what we understand structurally (the baseline and scenarios). We present the framework here and refer to [Adrian et al. \(2025, 2026\)](#) for more statistical details. The synthesis explores and quantifies how well the reference predictive distribution $p(y)$ can be approximated by a statistical mixture of the model-based conditional views:

$$f(y|\boldsymbol{\alpha}) = \sum_{j=0}^J \alpha_j p_j(y), \quad \alpha_j \geq 0, \quad \sum_{j=0}^J \alpha_j = 1, \quad (1)$$

where the weights $\boldsymbol{\alpha}$ quantify the contribution of each conditional view to the synthesis.

¹⁶[Schröder \(2025\)](#) addresses this limitation by decomposing inflation-at-risk densities into underlying risk drivers.

This mixture scenario synthesis resembles predictions using traditional Bayesian Model Averaging (BMA) under model uncertainty. In BMA, the $p_j(y)$ are predictions from a set of assumed models and the α_j are model probabilities. These probabilities are based on past data that have informed the models’ relative predictive performance. The more general framework of Bayesian predictive synthesis (BPS: Tallman and West, 2023; Johnson and West, 2025) allows the α_j to be specified otherwise, while maintaining the traditional mixture form. Note that there is no notion that $f(y|\alpha)$ is the data-generating mechanism: the α_j are wholly relative across the model set, with models weighted relative to one another based solely on predictive (and, in some cases, decision) outcomes. There is no assumption that there is a “true model” within the set (West and Harrison, 1997, sect. 12.2).

Scenario Synthesis adopts the mixture form inspired by BPS/BMA, but the interpretation differs. Each scenario is a distinct conditional distribution $p(y \mid A_j, M_j)$: a complementary exploration of an alternative state of the world, not a competing unconditional forecast. The synthesis weights are chosen so that $f(y|\alpha)$ best approximates the reference density $p(y)$. While the weights have the mathematical form of probabilities on the finite scenario set, they are not generally interpretable as subjective or data-based probabilities that the scenarios will occur. Rather, they are diagnostic importance measures relative to both the scenario set and the reference being matched: a high weight means that a scenario better approximates the reference than the other scenarios considered, not that the scenario is likely in an absolute sense.

There is one special case in which the weights have a direct probabilistic interpretation. Suppose that the baseline, scenarios, and reference density are all generated from the same model M , so that $M_j = M$ for all j , and that the conditioning assumptions $\{A_j\}$ form a mutually exclusive and exhaustive partition of the state space. Then

$$p(y \mid M) = \sum_{j=0}^J p(y \mid A_j, M)p(A_j \mid M),$$

so that the synthesis weights coincide with the marginal probabilities of the conditioning assumptions, $\alpha_j^* = p(A_j \mid M)$. Outside this case, we refer to the α_j^* as weights rather than probabilities.

6.1 Optimizing the weights

The synthesis weight vector $\alpha = (\alpha_0, \dots, \alpha_J)'$ is chosen such that the mixture is as “close” as possible to the reference. To formalize closeness, we use a concordance measure from statistical classification: the expected misclassification rate (EMR: see Adrian et al., 2025, 2026 for statistical background). This is given by

$$\pi_{pf}(\boldsymbol{\alpha}) = \int_y \frac{f(y|\boldsymbol{\alpha})p(y)}{\{f(y|\boldsymbol{\alpha}) + p(y)\}} dy.$$

Higher values of $\pi_{pf}(\boldsymbol{\alpha})$ indicate greater probabilistic concordance between the mixture $f(y|\boldsymbol{\alpha})$ and the reference $p(y)$. Since $\pi_{pf}(\boldsymbol{\alpha}) \in (0, 0.5]$, values close to 0.5 indicate high concordance.

We regularize the optimization by placing a weak Dirichlet prior on the weights, setting $\epsilon = 0.05$ as in (Adrian et al., 2025). Specifically, with $\boldsymbol{\alpha} \sim \text{Dir}(\mathbf{1}(1 + \epsilon))$, the posterior mode solves

$$\{\alpha_j^*\}_{j=0}^J = \arg \max_{\substack{\alpha_j > 0, \alpha_0 \geq \alpha_j \\ \sum_{j=0}^J \alpha_j = 1}} \left[\log\{\pi_{pf}(\boldsymbol{\alpha})\} + \epsilon \sum_{j=0}^J \log(\alpha_j) \right]. \quad (2)$$

The Dirichlet term is not needed to define concordance; it regularizes the weights, avoiding unstable zero weights when scenarios are nearly redundant. The restriction $\alpha_0 \geq \alpha_j$ is likewise not required mathematically; we impose it to reflect the institutional role of the baseline as the central projection around which scenarios are constructed. It prevents any single alternative scenario from receiving more weight than the baseline while still allowing the scenario set as a whole to dominate the baseline when the reference calls for it.

The residual divergence

$$\tilde{\pi}_{pf} = 0.5 - \pi_{pf}(\boldsymbol{\alpha}^*) \quad (3)$$

measures how far the optimally weighted scenario set is from spanning the reference distribution. A value of $\tilde{\pi}_{pf}$ close to zero indicates that the scenario set captures the reference well, whereas larger values reflect missing or redundant scenarios. The diagnostic is asymmetric: a large $\tilde{\pi}_{pf}$ indicates that the scenario set does not span the reference, while a small $\tilde{\pi}_{pf}$ establishes numerical fit but not necessarily economic coverage. For the latter interpretation, the scenarios must represent meaningful regions of the state space.

7 The Scenario Synthesis at Work

7.1 December 2007

We apply the Scenario Synthesis to the alternative scenarios presented in the December 2007 Greenbook (Federal Reserve Board, 2007), one year before the Great Recession. The episode is a natural application because two predictive densities with sharply different risk assessments were available: the SPF consensus and the NY Fed Blackbook’s judgmental density (BB).¹⁷ Holding the Tealbook scenario set fixed, we vary the reference distribution—

¹⁷Real-time Outlook-at-Risk estimates do not exist for 2007, as the underlying quantile-regression methodology postdates it; applying a method introduced a decade later would be inappropriate. The Blackbook

a controlled comparison isolating how alternative risk assessments shape the implied weights and diagnostics.

Although our methodology applies to multiple variables and horizons, we focus on one-year-ahead GDP growth in this section to clarify the mechanics. In Section 7.3, we extend the analysis to the joint distribution of GDP growth and core PCE inflation. Appendix A summarizes the main computational steps.

In the empirical applications, we report both the EMR and the Effective Sample Size statistic (ESS). The two are monotonically related. We report both because the EMR is the optimization objective, while the ESS is easier to read; equivalently, $100 - \text{ESS}$ provides an intuitive measure of residual non-overlap. Since the EMR flattens at high concordance, small EMR differences can correspond to economically meaningful differences in spanning.

7.1.1 Scenario Set and Modeling Framework

The December 2007 Greenbook features seven domestic alternative scenarios.¹⁸ The “Greater Housing Correction” scenario ($\mathcal{S}_1$) assumes a more severe deterioration in the housing market than in the Baseline. The “Credit Crunch” scenario ($\mathcal{S}_2$) examines a situation in which financial market turbulence leads to a sharp tightening of credit conditions, significantly restricting lending to businesses and households.

Scenarios $\mathcal{S}_3$ and $\mathcal{S}_4$ represent upside demand risks. The “Stronger Domestic Demand” scenario ($\mathcal{S}_3$) considers the possibility that financial stress exerts less drag on spending than in the Baseline. Building on this, the “Stronger Domestic Demand with Better Export Performance” scenario ($\mathcal{S}_4$) adds the assumption of stronger export growth.

The “More Room to Grow” scenario ($\mathcal{S}_5$) assumes faster potential output growth, while the “Greater Cost Pressures” scenario ($\mathcal{S}_6$) assumes that firms raise prices more aggressively in response to rising costs. Finally, the “Market-Based FFR” scenario ($\mathcal{S}_7$) assumes that the federal funds rate evolves according to the path implied by futures markets.

All scenarios are constructed using the same baseline model (FRB/US). In all but one, monetary policy responds via an estimated Taylor rule; the exception is the “Market-Based FFR” scenario ($\mathcal{S}_7$), in which the federal funds rate follows the path implied by futures markets. In the notation of Section 5, this implies that $M_j = M_0$ for all j , so that variation across scenarios arises from the conditioning assumptions A_j —including, for $\mathcal{S}_7$, the assumed policy path. The scenarios are therefore intended to capture distinct economic mechanisms

and SPF densities, by contrast, reflect contemporaneous assessments, avoiding hindsight bias.

¹⁸The December 2007 Greenbook also included a separate set of international alternative simulations; presenting domestic and international scenarios separately was standard practice until 2008 (Herbst et al., 2026). We do not incorporate the international simulations into our analysis. We flag them, however, because one included a downside risk to the U.S. outlook. This qualifies, but does not overturn, our conclusion that the domestic scenario set understated downside risk to GDP growth.

through alternative assumptions, although in practice some overlap remains.

We exclude $\mathcal{S}_4$ (“Better Export Performance”) from the analysis because, as shown in Table 1, its one-year-ahead GDP growth projection is identical to that of $\mathcal{S}_5$ (“More Room to Grow”). The two scenarios differ along other dimensions—inflation, unemployment rate, and the federal funds rate—but they are indistinguishable in the univariate analysis focused on GDP growth. We there-

fore drop $\mathcal{S}_4$ under the exclusion criterion used throughout: when two scenarios imply the same projection in the variable space under consideration, one is redundant for the optimization and the individual weights are not uniquely identified (Section 6.1). Removing $\mathcal{S}_4$ also leaves a set with non-overlapping conditioning assumptions in the univariate space, since $\mathcal{S}_4$ combines stronger domestic demand with better export performance and therefore nests $\mathcal{S}_3$.

Table 1: BASELINE AND ALTERNATIVE SCENARIOS
Dec. 2007 Tealbook — One-year ahead GDP growth projection

j	Scenario $\mathcal{S}_j$	P15	P50	P85
0	Baseline	0.1	1.3	2.5
1	Greater housing correction		0.9	
2	Credit crunch		−0.4	
3	Stronger domestic demand		1.7	
4	Better export performance		1.9	
5	More room to grow		1.9	
6	Greater cost pressure		1.2	
7	Market-based FFR		1.6	

NOTE: The baseline projection for 2008 is reported on page I-21; scenario projections appear on page I-17. Scenario values for 2008 are obtained by averaging 2008:H1 and 2008:H2. P50 is the point forecast (Baseline and Scenarios), while P15 and P85 are the bounds of the 70% interval. The Tealbook reports only the point forecast of the alternative scenarios, so their P15 and P85 are left blank.

7.1.2 Reference and Baseline Distributions

We now present the two references and compare them with the Baseline. We use the Crump et al. (2025) Large BVAR to convert the Blackbook and SPF forecasts for annual-average GDP growth into quantile forecasts for Q4/Q4 growth, as detailed in Appendix D.¹⁹ We then fit a skew- t density (Azzalini and Capitanio, 2003) to the resulting quantiles, choosing its four parameters to minimize the squared distance between target and fitted quantiles, following Adrian et al. (2019).

For the Baseline, the Tealbook reports a point forecast and a 70% interval (Table 1). We treat the point forecast as the median and the interval bounds as the 15th and 85th percentiles, and fit a skew- t with location at the point forecast, zero skewness, and degrees of freedom fixed at 50.

Figure 3 plots the Blackbook reference distribution (black), the SPF reference distribution (red), the Baseline (teal dash-dotted), and the alternative scenarios (shown as point forecasts in the left panel and as predictive densities in the right panel). The Baseline differs substantially from the BB Reference in both location and shape. By contrast,

¹⁹Because the Blackbook reports no numeric percentiles, we extract the histogram values using Optical Character Recognition; see Appendix C.

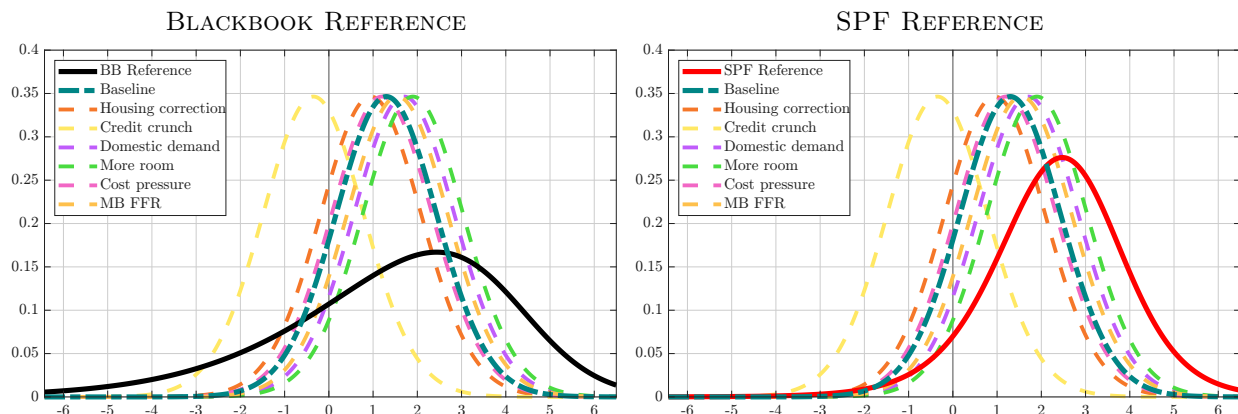
it primarily diverges from the SPF Reference in terms of location; see also the estimated skew- t parameters reported in Table 2. Visually, the figure already suggests that the scenario set is too concentrated around the Baseline to span either reference distribution well.

Table 2: SKEW- t PARAMETERS

	lc	sc	sk	df
BB Reference	4.2	3.6	-2.3	14.3
SPF Reference	3.0	1.5	-0.6	8.8
Baseline	1.3	1.1	0.0	50.0

NOTE: This table shows the skew- t fitted parameters for the Blackbook and SPF References and the Baseline. The parameters reported in the table correspond to the location (lc), scale (sc), skewness (sk), and degrees of freedom (df).

Figure 3: REFERENCE – BASELINE – SCENARIOS
2008 Q4/Q4 GDP GROWTH – DEC. 2007 TEALBOOK



NOTE: In the left panel, the solid black line denotes the Blackbook Reference p.d.f. In the right panel, the solid red line denotes the SPF Reference p.d.f. In both panels, the teal dash-dotted line denotes the Baseline p.d.f., while the dashed lines show the scenario p.d.f.s.

The two references convey sharply different risk assessments. The BB Reference is more dispersed and strongly negatively skewed, with substantial left-tail mass: it assigns 25% probability to negative GDP growth, consistent with FRBNY staff concern about financial conditions. The FRB staff expressed related concerns in the Tealbook, replacing the previous “Greater housing correction with larger fallout” scenario with the more severe “Credit Crunch” scenario. By contrast, the SPF Reference is more concentrated, closer to symmetric, and shifted to the right, assigning only 7% probability to negative growth. The two references thus emphasize different regions of the outcome space—downside tail risks versus relatively favorable outcomes—providing a natural setting to assess whether the Tealbook scenario set can match alternative risk assessments.

7.1.3 Scenario Synthesis

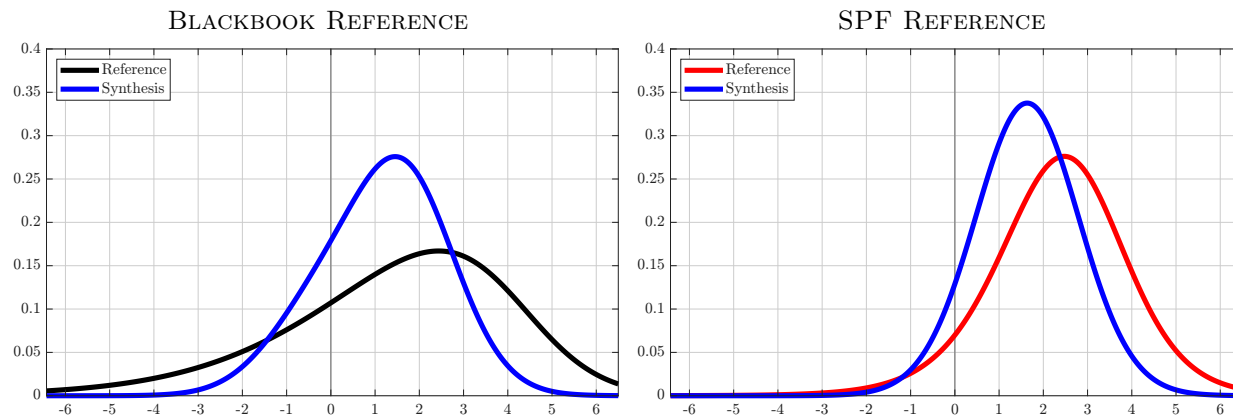
We now evaluate the Tealbook scenario set against each reference. Because the December 2007 Tealbook reports scenarios only as point forecasts, we construct scenario densities by shifting the location of the Baseline density to match each point forecast while keeping scale,

skewness, and degrees of freedom fixed (see also Section 8.2).

Using the Blackbook density as the reference $p(y)$, the Synthesis improves only modestly on the Baseline. The EMR rises from 0.40 to 0.43, and the ESS from roughly 58% to 69% (left column of Figure 4 and Blackbook panel of Table 3). The scenario set provides limited coverage of macroeconomic risk: the Blackbook Reference has substantial left-tail risk, while the scenarios remain concentrated around the Baseline.

This shortfall reflects both scenario design and model structure. FRB staff recognized downside risks, but the two downside scenarios were mild relative to the Blackbook density and, ultimately, to the downturn that followed. Moreover, because all scenarios were built with FRB/US, none incorporated the nonlinear amplification mechanisms of a systemic financial crisis. This does not mean that a linear framework such as FRB/US makes severe outcomes impossible: a large enough shock can generate a deep recession. The difficulty is that doing so may force implausible co-movements among inflation, interest rates, and other variables because the model lacks the nonlinear amplification mechanisms—fire sales, collateral spirals, and bank runs—through which a crisis propagates. For such scenarios, a nonlinear model is therefore preferable; see Section 9.

Figure 4: SCENARIO SYNTHESIS — PROBABILITY DENSITY FUNCTIONS
2008 Q4/Q4 GDP GROWTH – DEC. 2007 TEALBOOK



NOTE: In the left panel, the solid black line denotes the Blackbook Reference p.d.f.; in the right panel, the solid red line denotes the SPF Reference p.d.f. In both panels, the blue line denotes the Scenario Synthesis p.d.f.

Next, we use the SPF consensus predictive density as the reference $p(y)$. As shown in the SPF panel of Table 3, the same Tealbook scenario set receives markedly different weights. The synthesis assigns most weight to the upside scenarios, while the adverse scenarios receive essentially none. By contrast, under the Blackbook Reference, the Credit Crunch scenario receives a weight of approximately 0.24. This reflects how the two references allocate probability mass: the Blackbook calls for the most severe available downside scenario, while the SPF emphasizes relatively optimistic scenarios. It is worth noting that, regardless of the ref-

erence being matched, the “Greater housing correction” scenario receives little weight. This does not imply that housing risks were unimportant; rather, the scenario is too close to the Baseline to help match the left tail of the Blackbook Reference.

Table 3: SCENARIO SYNTHESIS — SUMMARY STATISTICS AND WEIGHTS
2008 Q4/Q4 GDP growth – Dec. 2007 Tealbook

j Scenario $\mathcal{S}_j$	Blackbook Reference			SPF Reference		
	ESS%	EMR	α_j^*	ESS%	EMR	α_j^*
0 Baseline	57.9	0.396	0.238	61.1	0.419	0.313
1 Greater housing correction	54.0	0.383	0.009	47.4	0.379	0.007
2 Credit crunch	37.1	0.313	0.238	14.3	0.221	0.013
3 Stronger domestic demand	61.5	0.407	0.238	76.7	0.456	0.313
4 With better export performance						
5 More room to grow	62.7	0.410	0.238	83.2	0.469	0.313
6 Greater cost pressure	56.9	0.393	0.011	57.1	0.408	0.009
7 Market-based FFR	60.3	0.404	0.027	71.0	0.443	0.031
----- Synthesis	----- 69.3	----- 0.430	----- -----	----- 74.5	----- 0.452	----- -----

NOTE: ESS% denotes the effective sample size as a percentage of the total sample. EMR is the Expected Misclassification Rate for each scenario. α_j^* are the optimal synthesis weights, which sum to one across the scenarios included in the synthesis. Scenario 4 is grayed out because it is excluded from the synthesis.

The scenario set also fails to span the SPF Reference well: $\tilde{\pi}_{pf} = 0.5 - 0.452 = 0.048$, or 25.5% when computed as $100 - \text{ESS}$. This residual divergence is smaller than with the Blackbook Reference, where $\tilde{\pi}_{pf} = 0.070$ (30.7% as $100 - \text{ESS}$), but this should not be interpreted as evidence that the SPF provided a better risk assessment. Concordance is only as informative as the reference being matched: the SPF is easier to span precisely because it assigns little probability to adverse outcomes. The more relevant diagnostic is where the mismatch occurs. In contrast to the Blackbook case, the right column of Figure 4 shows that the incompleteness is concentrated in the right tail, as the available scenarios do not generate sufficiently optimistic outcomes to match the SPF Reference.

WHY THE TWO SYNTHESSES DIVERGE. The divergence arises because the same scenario set is evaluated against references that place probability mass in different regions: the Blackbook assigns much more mass to the left tail, while the SPF assigns little probability to adverse outcomes. The scenario set spans both references imperfectly, but in different regions. The residual divergence is larger against the Blackbook (0.070) than against the SPF (0.048), but the more informative distinction is the location of the gap: the scenarios are not severe enough for the Blackbook reference and not optimistic enough for the SPF reference. The diagnostic value lies less in the small numerical difference than in the location of the mismatch, which the Synthesis makes transparent by linking reference distributions to weights and fit diagnostics.

This comparison also highlights the value of models that incorporate financial conditions when assessing macroeconomic risks. As discussed in Section 2.3, Growth-at-Risk approaches formalize the link between financial variables and downside risk. The point is not that a statistical model would necessarily have outperformed expert judgment in real time: the NY Fed’s judgmental Blackbook density captured the elevated downside risk that the SPF missed. Rather, a reference informed by financial conditions, as the Blackbook density was, makes the scenario-set gap visible. Against such a reference, the Scenario Synthesis flags systematically and quantitatively that the available scenarios understate left-tail risk.

Overall, this exercise illustrates how the Scenario Synthesis provides a unified framework for evaluating whether a given scenario set adequately captures alternative risk assessments and diagnosing the sources of any mismatch.

7.2 December 2018: Univariate Analysis

In this section, we apply the Scenario Synthesis to the alternative scenarios reported in the December 2018 Tealbook (Federal Reserve Board, 2018). The December 2018 Tealbook provides a useful “normal-times” benchmark, as macroeconomic and financial conditions were relatively stable. Two statistical reference densities were available at the time: the Time-Varying Macroeconomic Risk (TVMR) density reported in the Tealbook and the New York Fed’s Outlook-at-Risk (OaR) density.²⁰ We begin with one-year-ahead GDP growth and then extend the analysis to the joint distribution of GDP growth and core PCE inflation.

7.2.1 Scenario Set and Modeling Framework

Unlike December 2007, the December 2018 Tealbook lets us compare the same scenario set with both the NY Fed’s OaR density and the Board’s own TVMR density. It therefore provides a setting in which the narrative and statistical approaches to risk can be evaluated jointly within the Federal Reserve System.

The December 2018 Tealbook features five alternative scenarios (Table 4). The “Financial-based recession” scenario ($\mathcal{S}_1$) examines a situation in which a correction in financial market valuations, combined with constraints on financial intermediaries, triggers a recession. The “Stronger Supply Side” scenario ($\mathcal{S}_2$) assumes more favorable supply conditions than in the Baseline, with a smaller output gap and faster potential growth. In contrast, the “Supply constraints” scenario ($\mathcal{S}_3$) considers the consequences of prolonged labor-market tightness. In this scenario the growth is the same as the baseline, but as we will discuss later, inflation is higher.²¹ The “Greater Interest Rate Sensitivity” scenario ($\mathcal{S}_4$) assumes that household

²⁰Although OaR was first published in 2023, historical estimates are available, and the underlying methodology was already well established in 2018, which is why we will use it in this section.

²¹In this univariate analysis, $\mathcal{S}_3$ is indistinguishable from the Baseline and hence we drop it. We will

and business spending respond more strongly to interest rate increases than in the Baseline. Finally, the “Foreign Slowdown” scenario ($\mathcal{S}_5$) simulates the effects of weaker foreign activity and an appreciation of the dollar.

Scenarios $\mathcal{S}_2$ and $\mathcal{S}_4$ are constructed using FRB/US (so $M_2 = M_4 = M_0$). Scenario $\mathcal{S}_1$ uses a Gertler–Karadi-type model (Gertler and Karadi, 2011) with a richer financial sector, $\mathcal{S}_3$ relies on a calibrated New Keynesian DSGE model with labor-market search frictions similar to Gertler et al.

(2008), and $\mathcal{S}_5$ is based on SIGMA, a calibrated multicountry DSGE model maintained at the Federal Reserve Board. The assumptions underlying the scenarios are non-overlapping and correspond to distinct economic mechanisms.

Table 4: BASELINE AND ALTERNATIVE SCENARIOS
Dec. 2018 Tealbook — One-year-ahead GDP growth projection

j	Scenario $\mathcal{S}_j$	P15	P50	P85
0	Baseline	1.2	2.4	3.9
1	Financial-based recession		−0.7	
2	Stronger supply side		3.1	
3	Supply constraints		2.4	
4	Greater interest rate sensitivity		1.5	
5	Foreign slowdown		1.6	

NOTE: The baseline projection for 2019 is on page 88, while the alternative scenarios are on page 84. P50 is the point forecast (Baseline and Scenarios), while P15 and P85 are the bounds of the 70% interval.

7.2.2 Reference and Baseline Distributions

For the December 2018 Tealbook, we consider two alternative reference distributions: the TVMR density reported in the Tealbook and the OaR density produced by the New York Fed. In practice, both the Tealbook and the New York Fed report predictive percentiles (P5, P15, P50, P85, and P95 for TVMR; P10, P25, P50, P75, and P90 for OaR). To obtain full reference distributions, we fit a skew- t density to these percentiles.

Table 5: SKEW- t PARAMETERS – DEC. 2018

	lc	sc	sk	df
OaR Reference	2.5	1.3	−0.3	3.0
TVMR Reference	2.2	1.1	0.4	30.8
Baseline	2.4	1.3	0.0	50.0

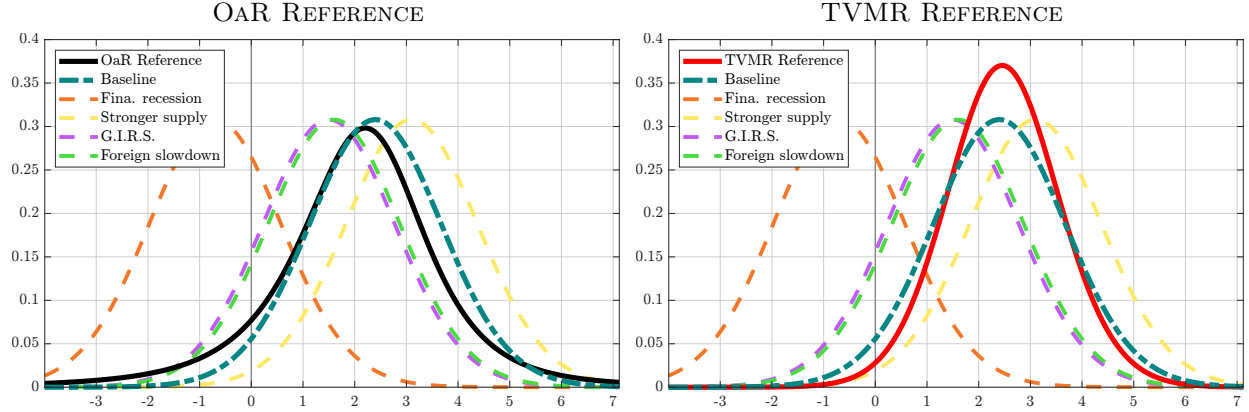
NOTE: This Table shows the skew- t fitted parameters for the OaR and TVMR References and the Baseline. The parameters reported in the table correspond to the location (lc), scale (sc), skewness (sk), and degrees of freedom (df).

The Baseline is constructed in the same way as in Section 7.1.

Figure 5 shows that the Baseline is broadly aligned with both References. However, Table 5 reveals meaningful differences in shape. Relative to the Baseline, the OaR Reference exhibits fatter tails and modest negative skewness, whereas the TVMR Reference is somewhat less dispersed and displays mild positive skewness. Thus, even in a relatively tranquil period, the two statistical approaches imply different assessments of the risks around the baseline outlook.

be able to attach weight when considering growth and inflation jointly in Section 7.3. The Tealbook also presents a sixth scenario, “Lower Oil Prices,” which we exclude because its implications for 2019 GDP growth are identical to the Baseline and differ only marginally (0.1 percentage point) for core PCE inflation.

Figure 5: REFERENCE – BASELINE – SCENARIOS
2019 Q4/Q4 GDP GROWTH – DEC. 2018 TEALBOOK

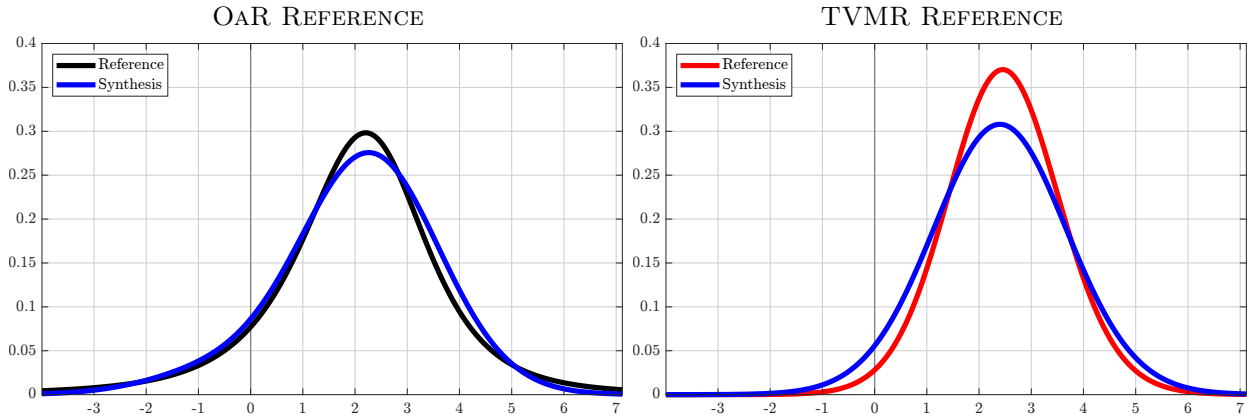


NOTE: In the left panel, the solid black line denotes the OaR Reference p.d.f. In the right panel, the solid red line denotes the TVMR Reference p.d.f. In both panels, the teal dash-dotted line denotes the Baseline p.d.f., while the dashed lines show the scenario p.d.f.s.

7.2.3 Scenario Synthesis

Figure 6 reports the resulting syntheses. When the Synthesis is constructed to match the OaR Reference, the fit is excellent, with only a minor underestimation of probability mass near the center of the distribution. When the Synthesis is constructed to match the TVMR Reference, the fit is also strong, but the decomposition is much less informative: because the TVMR Reference lies so close to the baseline, the Synthesis places most of the weight on the Baseline and only limited weight on the alternative scenarios.

Figure 6: SCENARIO SYNTHESIS — PROBABILITY DENSITY FUNCTIONS
2019 Q4/Q4 GDP GROWTH – DEC. 2018 TEALBOOK



NOTE: In the left panel, the solid black line denotes the OaR Reference p.d.f.; in the right panel, the solid red line denotes the TVMR Reference p.d.f. In both panels, the blue line denotes the Scenario Synthesis p.d.f.

The summary statistics in Table 6 confirm these visual impressions. In both cases, the scenario set spans the reference distribution well. However, the decomposition implied by

the two references differs markedly. Because the TVMR Reference is extremely close to the Baseline, the Baseline receives a weight of 0.87, while all scenarios other than “Stronger Supply Side” receive negligible weight. By contrast, when the reference is OaR, whose fitted density exhibits fatter tails, the resulting weights provide a richer characterization of risks around the outlook.

Under the OaR Reference, we can summarize the December 2018 risk assessment as follows: “Risks to the outlook are modestly skewed to the downside. The Synthesis assigns almost half of its weight (about 48%) on the Baseline materializing one year ahead, about 15% weight on a modest upside scenario (for example, due to stronger supply conditions), roughly 30% on mild downside scenarios, and about 7% to a financially driven recession.”

Table 6: SCENARIO SYNTHESIS — SUMMARY STATISTICS AND WEIGHTS
2019 Q4/Q4 GDP growth – Dec. 2018 Tealbook

j	Scenario $\mathcal{S}_j$	OaR Reference			TVMR Reference		
		ESS%	EMR	α_j^*	ESS%	EMR	α_j^*
0	Baseline	90.6	0.481	0.476	91.8	0.493	0.871
1	Financial-based recession	15.5	0.229	0.069	0.2	0.137	0.003
2	Stronger supply side	62.7	0.434	0.153	62.9	0.464	0.086
3	Supply constraints						
4	Greater interest rate sensitivity	79.2	0.465	0.133	33.3	0.426	0.018
5	Foreign slowdown	83.0	0.471	0.169	39.9	0.438	0.022
	Synthesis	96.0	0.492		88.6	0.491	

NOTE: ESS% denotes the effective sample size as a percentage of the total sample. EMR is the Expected Misclassification Rate for each scenario. α_j^* are the optimal synthesis weights, which sum to one across scenarios 0–5. Scenario 3 is grayed out as it is excluded from the synthesis.

Compared to the December 2007 case, the December 2018 scenario set spans the reference distribution well, consistent with a period in which alternative risk assessments were broadly aligned.

7.3 December 2018: Multivariate Analysis

The analysis so far has focused on GDP growth alone. This is sufficient to build intuition, but policymakers care about multiple variables at once. In particular, the Fed’s dual mandate makes monetary policy a trade-off between output and inflation risks, requiring a joint assessment. We therefore extend the analysis to a bivariate setting with one-year-ahead GDP growth and core PCE inflation.

This extension has an important implication: all five scenarios become potentially informative. The “Supply constraints” scenario, excluded in the univariate case because it coincided with the Baseline GDP median, introduces a distinct inflationary configuration once inflation is included and therefore contributes new information to the synthesis (Ta-

ble 7).

Table 7: BASELINE AND ALTERNATIVE SCENARIOS
2019 Q4/Q4 GDP growth and core PCE price inflation projections – Dec. 2018 Tealbook

j Scenario $\mathcal{S}_j$	GDP			Core PCE price inflation		
	P15	P50	P85	P15	P50	P85
0 Baseline	1.2	2.4	3.9	1.2	2.0	2.8
1 Financial-based recession		−0.7			1.9	
2 Stronger supply side		3.1			2.0	
3 Supply constraints		2.4			2.6	
4 Greater interest rate sensitivity		1.5			2.0	
5 Foreign slowdown		1.6			1.7	

NOTE: The baseline projection for 2019 is on page 88, while the alternative scenarios are on page 84. P50 is the point forecast (Baseline and Scenarios), while P15 and P85 are the bounds of the 70% interval.

7.3.1 Reference and Baseline Distributions

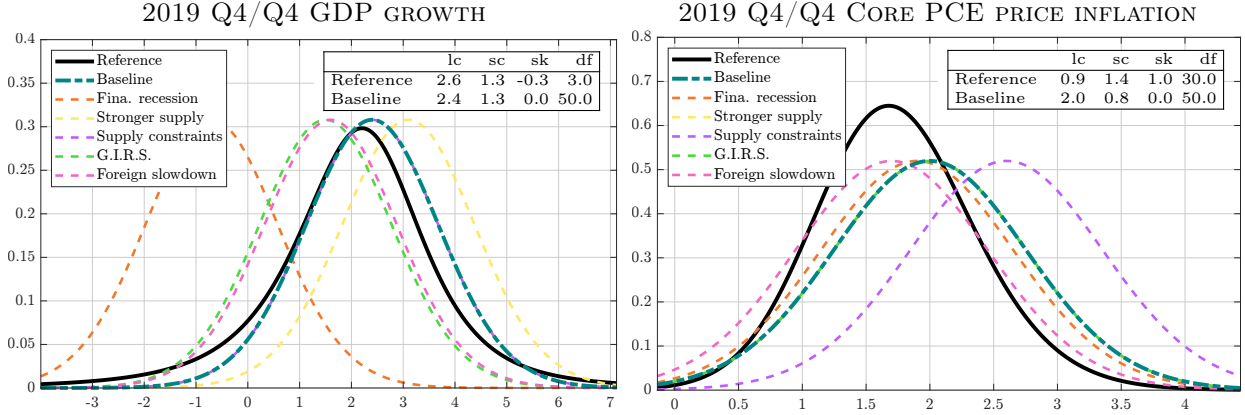
For the multivariate analysis, we focus on the OaR density because Section 7.2 shows that TVMR is extremely close to the Baseline and therefore less informative. We construct joint Reference and Baseline distributions by combining the GDP growth and core PCE marginals with a copula dependence structure, calibrating the correlation matrix from a VAR and imposing no tail dependence (see Appendix A.2). The marginals are constructed as in the univariate exercise; for core inflation, however, we had to convert the OaR headline CPI density into a core PCE density as described in Appendix B.

Figure 7 shows the marginals. For GDP growth, the comparison is unchanged from the univariate case: the Reference and Baseline are centered close together, differing mainly in asymmetry and tails. For core inflation, the contrast is much stronger: the Reference is shifted left, much more dispersed, and strongly positively skewed, while the Baseline remains centered near 2 percent and symmetric. The inflation marginal therefore reflects disagreement about both the center and the balance of risks, which affects which scenarios help match the joint Reference.

Figure 8 adds the bivariate view. The left panel places the Baseline and scenario point forecasts on the Reference contour map: the scenarios are dispersed along GDP but compressed along inflation, so only a subset is well positioned to match the joint shape. The right panel overlays the Baseline joint density, which captures the broad center but not the full geometry, especially along inflation.²²

²²Scenario marginals are constructed as in the univariate exercise and combined using the same correlation structure as the Baseline and Reference.

Figure 7: REFERENCE, BASELINE, AND SCENARIOS – DEC. 2018 TEALBOOK



NOTE: In both panels, the solid black line denotes the OaR Reference p.d.f. and the teal dash-dotted line denotes the Baseline p.d.f. Dashed lines show the scenario p.d.f.s for GDP growth (left panel) and core PCE price inflation (right panel). The skew- t parameters reported in the inset boxes (upper right of each panel) correspond to the location (lc), scale (sc), skewness (sk), and degrees of freedom (df).

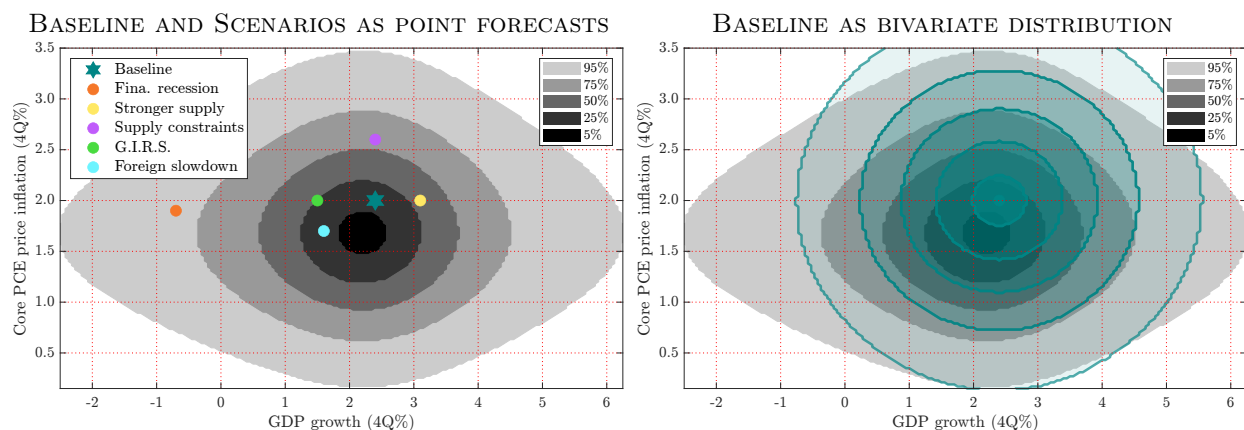
7.3.2 Scenario Synthesis

Figure 9 and Table 8 report the resulting Scenario Synthesis. The contour plot shows that the Synthesis improves visibly upon the Baseline and tracks the overall shape of the Reference reasonably well, although the fit is far from exact. This improvement is evident both graphically—by comparing Figure 9 with the right panel of Figure 8—and in the summary statistics: the EMR rises from 0.459 for the Baseline to 0.477 for the Synthesis, while the ESS increases from 69.5% to 79.9%.

The estimated weights show that this improvement is achieved through a selective reweighting of the scenario set rather than through broad use of all scenarios. The Baseline and the “Foreign slowdown” scenario receive equal weight, 0.376 each, and together account for more than three quarters of the synthesis. The next most important component is “Stronger supply side,” with weight 0.152, followed by “Financial-based recession,” with weight 0.060. By contrast, “Greater interest rate sensitivity” receives only 0.03, and “Supply constraints” is essentially irrelevant, with weight 0.006.

This pattern is intuitive given the joint Reference. Once output and inflation risks matter jointly, the most useful scenarios match not only the GDP distribution but also the inflation asymmetry. The “Foreign slowdown” scenario does this well: weaker growth with lower inflation reproduces the left-shifted inflation distribution without moving GDP far from the Reference mass. By contrast, “Supply constraints,” which mainly raises inflation, becomes much less useful. Adding inflation thus makes the scenario set more informative but exact coverage more demanding.

Figure 8: REFERENCE, BASELINE, AND SCENARIOS – DEC. 2018 TEALBOOK



NOTE: Both panels show joint probability density for GDP growth (x -axis) and core PCE inflation (y -axis). Contour lines show regions of equal probability density. Color intensity indicates probability concentration. The left panel displays the OaR Reference p.d.f. with colored points showing the Baseline and alternative scenario forecasts as (GDP, inflation) pairs. Contour regions denote the 5%, 25%, 50%, 75%, and 95% highest density regions. The right panel shows only the Baseline joint p.d.f. (teal contours) overlaid on the Reference p.d.f. (gray shaded regions).

8 Practical Considerations

The empirical exercises highlight that the Scenario Synthesis is not just a tool for ex post evaluation. It also raises a set of practical questions for real-time policy work: Where do scenarios come from? What should be done when scenarios are reported only as point forecasts or narratives rather than as full predictive distributions? And how should institutions respond when the available scenario set is unable to span the relevant risks? This section discusses these issues.

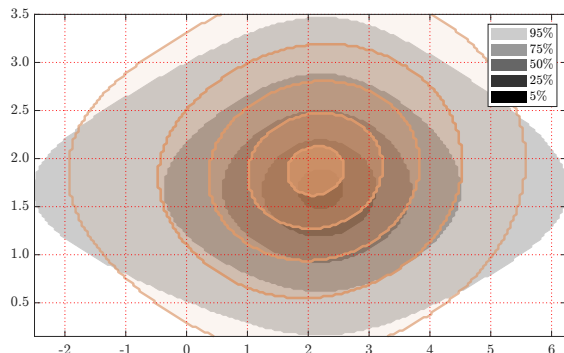
8.1 Where do scenarios come from?

A practical question is how the scenario set is selected. As discussed in Section 2.2, scenarios change from round to round: new risks emerge, old concerns fade, and the selection criteria are usually implicit. Staff ask which risks are salient, which mechanisms the baseline may miss, and which questions policymakers are raising, but these criteria are rarely formalized.

This matters because the scenario set changes from round to round: one cannot track the weight on a named scenario if it may not exist next round. But the Synthesis also reduces the need to communicate risk shifts through scenario turnover. Without it, declining optimism is conveyed by adding downside scenarios or removing upside ones; with it, the same shift appears in the weights on scenarios already on the table. What matters is not only which scenarios are present, but how weight is distributed across them. Turnover remains relevant but less central, and the Synthesis can inform selection by diagnosing missing and redundant scenarios.

A related question is reverse stress testing, which derives the scenarios that would produce

Figure 9: SCENARIO SYNTHESIS
PROBABILITY DENSITY FUNCTION
GDP GROWTH AND CORE INFLATION
DEC. 2018 TEALBOOK



NOTE: This figure shows probability density contours for the joint p.d.f. of GDP growth (x -axis) and core PCE price inflation (y -axis) under the Scenario Synthesis distribution (orange contours), overlaid on the joint Reference p.d.f. (gray shaded regions). Contour regions denote the 5%, 25%, 50%, 75%, and 95% highest density regions.

Table 8: SCENARIO SYNTHESIS
SUMMARY STATISTICS AND WEIGHTS
GDP GROWTH AND CORE INFLATION
DEC. 2018 TEALBOOK

j Scenario $\mathcal{S}_j$	ESS%	EMR	α_j^*
0 Baseline	69.5	0.459	0.376
1 Financial-based recession	13.1	0.225	0.060
2 Stronger supply side	48.2	0.415	0.152
3 Supply constraints	17.3	0.347	0.006
4 Greater interest rate sensitivity	60.8	0.444	0.030
5 Foreign slowdown	70.1	0.462	0.376
Synthesis	79.9	0.477	

NOTE: ESS% denotes the effective sample size as a percentage of the total sample. EMR is the Expected Misclassification Rate for each scenario. α_j^* are the optimal synthesis weights, which sum to one across scenarios 0–5.

a given outcome from exposures, balance sheets, or realized losses. Scenario design is often iterative in this spirit: construct, assess, adjust. A reverse-engineered scenario can be treated as another conditional density $p_j(y)$ once finalized, and the diagnose-and-re-synthesize loop of Section 8.4 follows the same logic. A full treatment is beyond our scope, but the framework is compatible with it.

8.2 Scenario predictive distributions

In our theoretical framework, each scenario provides a full predictive distribution $p_j(y)$, constructed using the appropriate model M_j and assumptions A_j . This is the first-best case, because the synthesis can then be implemented directly using the scenario densities.

In practice, scenarios are often presented only as point forecasts or narratives. In this second-best case, the framework can still be implemented by constructing scenario-consistent densities via entropic tilting or related methods that embed the scenario path in the baseline’s uncertainty—a feasible workaround, used in Section 7, but not ideal. The limitation bites when the scenario requires $M_j \neq M_0$: the reason for a different model is usually that the baseline cannot represent the nonlinear mechanism, so a scenario from a nonlinear accelerator or kinked Phillips curve implies asymmetric uncertainty or fat tails that tilting the baseline density may miss. For internal policy analysis, institutions should produce full predictive distributions for scenarios whenever possible, even if external communication emphasizes central paths.

8.3 Joint risk assessment

All the predictive distributions considered so far are marginal densities rather than full joint densities. In Section 7.3, we constructed a joint density using a normal copula, that is, by assuming no tail dependence and calibrating the linear correlation across variables. This is a useful workaround, but at best a second-best solution. Ideally, the reference would provide a full joint risk assessment rather than marginals stitched together through auxiliary assumptions; but the literature has progressed far more on marginal tail risks than on their joint distribution. Multivariate applications thus remain feasible and informative but require dependence assumptions beyond the original reference densities.

The bivariate exercise of Section 7.3 extends directly to more variables and multiple horizons in principle. In practice, however, the dependence assumptions become more demanding in higher dimensions. The statistical reference densities we use are also predominantly one-year-ahead, which limits the horizons we can study. Finally, beyond the bivariate case, the results become harder to display, since the contour plots used above are no longer available. These are constraints of data and communication, not of principle.

8.4 What to do if the scenario set is incomplete?

As shown in the December 2007 case study in Section 7.1, the scenario set was unable to recover the Reference distribution. In that exercise, we treated the scenario set as fixed. In practice, however, the scenario set is not fixed, and the relevant question becomes what to do when it is incomplete.

A natural starting point is statistical. Scenario-set incompleteness is the analogue of “model-set incompleteness” in Bayesian econometrics. Bayesian Predictive Synthesis (BPS)—and its decision-guided extension BPDS (Tallman and West, 2023; Chernis et al., 2024)—augments the initial densities with a *backstop*: a deliberately over-dispersed distribution centered on the mixture, designed to capture outcomes outside the original support, often implemented as an over-dispersed average of the existing densities. The same idea can be applied to scenario analysis by constructing a backstop scenario. One concrete formulation is:

$$P50_{backstop} = \text{median}_{j=1,\dots,J} P50_j, \quad P15_{backstop} = \min_j P15_j, \quad P85_{backstop} = \max_j P85_j.$$

Such a scenario is grounded in BPS theory and guarantees improved spanning, but lacks a clear economic interpretation, limiting its communicative value; Adrian et al. (2025) provide a concrete implementation of this backstop construction to which we refer the reader for details.

An alternative, and often more economically meaningful, approach is to diagnose which

risks are missing and design new scenarios—or recalibrate existing ones—to span those regions of the predictive distribution. In practice, the framework is most useful iteratively: start with the existing scenario set, compute the weights and $\tilde{\pi}_{pf}$, diagnose the gaps, refine the scenarios, and re-synthesize. The loop can also run in reverse: under-covered regions of the reference can guide the conditioning assumptions for new scenarios. Thus, the Synthesis serves not only as an evaluation tool, assessing an existing scenario set, but also as a design tool, guiding improved scenarios.

9 Alternative Economic Models

Baseline models—FRB/US, NAWM II, and COMPASS—are local approximations calibrated for normal times. They linearize around a steady state and represent financial frictions only partially. In calm periods this is often adequate; in stress episodes it can become misleading. The nonlinear amplification mechanisms that define crises—fire sales, collateral spirals, and wage-price spirals—are precisely the mechanisms that scenario analysis is designed to illuminate, and they require models the baseline excludes.

This is the operational meaning of $M_j \neq M_0$ in Section 5. In December 2007, the “Credit Crunch” scenario required endogenous credit feedback that FRB/US could only partially capture. In December 2018, the “Financial-based recession” scenario used a Gertler–Karadi-type model (Gertler and Karadi, 2011), while the “Supply constraints” scenario had different inflation implications depending on whether the Phillips curve was flat or steep. We discuss these two mechanisms briefly and then introduce the Macroeconomic Model Database as a tool for understanding model uncertainty.

THE NONLINEAR FINANCIAL ACCELERATOR. Standard baseline models include a simplified financial sector: credit spreads respond linearly, while the feedback loops that amplify distress—falling collateral tightening credit, contracting output, and further depressing collateral—are absent. Bernanke et al. (1999) embed these loops structurally; Kiyotaki and Moore (1997) formalize a related mechanism through land prices and collateral constraints; and Gertler and Karadi (2011) and Gertler and Kiyotaki (2010) extend the accelerator to constrained intermediaries, providing the structural basis for asset purchases and macroprudential regulation (Adrian and Shin, 2010). The implication is direct: when financial conditions are tight and the accelerator is operative, the reference distribution should place more mass in the left tail, and adverse scenarios should receive higher weights—the pattern captured by the Blackbook density in December 2007 and missed by the SPF.

THE NONLINEAR PHILLIPS CURVE. Standard models embed a linear Calvo Phillips curve: inflation responds proportionally to slack regardless of state. The early-2020s inflation surge

challenged this assumption, and evidence for a state-dependent slope has grown. In menu-cost models, large shocks raise adjustment frequency and steepen the curve (Costain and Nakov, 2019; Alvarez et al., 2022); Benigno and Eggertsson (2023) propose an “Inverse-L” curve, flat below a labor-market threshold and steep above; and Harding et al. (2023) reach a complementary conclusion through quasi-kinked demand. For scenario design, this changes the stakes of supply scenarios. The December 2018 “Supply constraints” scenario examined prolonged tightness: under a linear curve, the inflation implications are moderate; under a convex curve, they are sharply higher, calling for a different model. Assigning probability to the scenario therefore requires taking a stand on which portion of the curve the economy occupies—squarely an $M_j \neq M_0$ disagreement (Section 5). The 2021–22 run-up illustrates how a mechanism once confined to a scenario can become central to the baseline.

THE MACROECONOMIC MODEL DATABASE. The mechanisms above are well-understood departures from the baseline. In practice, however, the space of plausible models is much larger, differing in price and wage rigidity, expectations, openness, and policy rules. For a given shock, these differences can generate forecast divergences that rival those between the baseline and the most extreme scenario. The Macroeconomic Model Database (MMB; Wieland et al., 2012, 2016) helps navigate this space: it is an open-source archive of more than 150 structural models from academia and policy institutions, re-coded into a common format so impulse responses can be compared like-for-like. For Scenario Synthesis, it makes the M_j dimension concrete: staff can identify which models predict meaningfully different outcomes for a given shock, and therefore which disagreements merit distinct scenarios. The MMB thus provides a principled way to populate $\{M_j\}$ and turn implicit model uncertainty into explicit weights.

10 Policy Considerations: Consensus Under Deep Uncertainty

Policymakers often disagree about both the near-term outlook and long-run concepts such as the natural rate of interest, the equilibrium unemployment rate, or the slope of the Phillips curve. In the U.S., dispersion in the FOMC’s Summary of Economic Projections makes this disagreement visible. Similar disagreements persist at other central banks such as within the ECB’s Governing Council and the Bank of England’s MPC despite common staff briefings, forecasts, and risk assessments.

Such disagreement is natural under *deep uncertainty*—a term used in the operations research and management science literature to describe situations in which decision-makers cannot agree on the model of the system, the relevant outcomes, or the probability distributions over uncertain inputs (Walker et al., 2013). Central banks face deep uncertainty:

there is no agreement on how best to represent the economy, transmission mechanisms shift over time, and the distributions governing future outcomes cannot be pinned down with confidence.

Disagreement under deep uncertainty has three sources:

- i) *Shock uncertainty* concerns the assumptions A_j . Policymakers may agree on the model but disagree about which shocks have hit the economy, how to interpret recent developments, or which shocks are likely ahead.
- ii) *Transmission uncertainty* concerns the model component M_j . Policymakers may disagree about how the economy works and which mechanisms are operative. As discussed in Section 9, the same shock path can generate different outcomes depending on whether financial acceleration, nonlinear price adjustment, or another form of state dependence is relevant.
- iii) *Ranking uncertainty* concerns the loss function. Policymakers may agree on shocks and transmission but still weigh outcomes differently: one may worry more about inflation, another about unemployment. Point forecasts obscure this distinction by combining beliefs about the economy with the weighting of risks; predictive densities help separate what is believed from what is prioritized.

Once these sources of disagreement are explicit, the question becomes how to aggregate them into a consensus risk assessment. Scenario Synthesis offers a natural device. Each committee member’s view can be represented as a conditional predictive density $p(y \mid A_j, M_j)$, just as scenarios are. Under this view, disagreement is structured information about which models and assumptions members consider most relevant, not noise to be averaged away.

Rather than reporting point forecasts that combine beliefs with risk rankings, members could work with staff to design scenarios reflecting their preferred model and assumptions. Scenario Synthesis would then aggregate these views into a committee-level risk assessment. In our applications, we use a relatively flat prior and let the reference density discipline the weights through the concordance criterion. More generally, members could bring priors over scenarios, reflecting their subjective probabilities that a given (M_j, A_j) combination is the relevant approximation.

As views and the reference density evolve, the synthesis can be recomputed, providing a transparent record of how the committee’s collective assessment changes over time. The residual divergence $\tilde{\pi}_{pf}$ is especially useful: a large residual indicates that the scenario set does not span the range of committee views and that new scenarios are needed. The framework thus serves not only as an evaluation tool but also as a discipline for scenario design under deep uncertainty.

Finally, it is important to note that Scenario Synthesis is a tool for risk assessment, not

a decision rule. It does not prescribe a policy action or close the feedback loop from risk assessment to policy response and realized outcomes: easing in response to downside risk may itself reduce that risk. Embedding scenario analysis in a decision framework that closes this loop is the subject of a complementary literature. Cairó et al. (2025) and Garga et al. (2025), for example, develop risk-adjusted optimal-control policies in which policymakers weigh scenario probabilities and update them as shocks arrive. A key lesson is that the policy path is not a probability-weighted average of scenario-specific policy paths, because beliefs, policy, and outcomes interact.²³

11 Conclusions

Central banks have developed two approaches to monitoring macroeconomic risk—scenario analysis and predictive distributions—that evolved in parallel yet separately. Scenarios provide narratives without probabilities; predictive distributions provide probabilities without narrative. We argue that the two are complements, not substitutes. We provide a historical account of their emergence and introduce the Scenario Synthesis, a framework that connects them by assigning scenario weights consistent with a reference predictive distribution.

The Scenario Synthesis evaluates existing scenario sets and helps improve them. Scenarios are necessarily subjective and lack continuity across rounds; the Synthesis adds quantitative discipline by comparing the weighted combination of the scenarios with the reference distribution. This makes it possible to assess whether scenarios span the relevant risks, identify what is missing, and guide the construction of new scenarios.

Beyond scenario design, the framework offers a structured way to represent disagreement within policy committees. Members may differ not only in how they weight outcomes, but also in their views of shocks and transmission. This helps explain why disagreement persists even under shared information, and why scenario weights can serve as a disciplined summary of heterogeneous views.

References

- Adams, P. A., T. Adrian, N. Boyarchenko, and D. Giannone (2021). Forecasting macroeconomic risks. *International Journal of Forecasting* 37(3), 1173–1191.
- 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

²³Relatedly, the scenarios entering the synthesis are conditional on an assumed policy reaction function, so the weights inherit that conditioning. Pure monetary-policy scenarios—those that vary only the assumed rule—should therefore be interpreted accordingly.

- risk. Finance and Economics Discussion Series 2025-036, Board of Governors of the Federal Reserve System.
- Adrian, T., D. Giannone, M. Luciani, and M. West (2026). Predictive concordance for parameter optimisation and mixture synthesis. arXiv 2606.14382.
- Adrian, T., F. Grinberg, N. Liang, S. Malik, and J. Yu (2022). The term structure of Growth-at-Risk. *American Economic Journal: Macroeconomics* 14(3), 283–323.
- Adrian, T., J. Morsink, and L. B. Schumacher (2020). Stress testing at the IMF: A framework for macroprudential analysis. Departmental Paper 2020/016, International Monetary Fund.
- Adrian, T. and H. S. Shin (2010). Financial intermediaries and monetary economics. Staff Reports 398, Federal Reserve Bank of New York.
- Aikman, D., J. Bridges, S. Hacioglu Hoke, C. O’Neill, and A. Raja (2019). Credit, capital and crises: a gdp-at-risk approach. Working Paper 824, Bank of England.
- Albuquerque, D., J. Chan, D. Kanngiesser, D. Latto, S. Lloyd, S. Singh, and J. Žáček (2025). Decompositions, forecasts and scenarios from an estimated dsge model for the uk economy. Macro Technical Paper 1, Bank of England.
- Alessandri, P., L. D. Vecchio, and A. Miglietta (2019). Financial conditions and ‘Growth at Risk’ in Italy. Economic Working Paper 1242, Bank of Italy, Economic Research and International Relations Area.
- Alvarez, F., F. Lippi, and A. Oskolkov (2022). The macroeconomics of sticky prices with generalized hazard functions. *Quarterly Journal of Economics* 137(2), 989–1038.
- Amburgey, A. and M. W. McCracken (2025, August). Growth-at-risk is investment-at-risk. Working Paper 2023-020C, Federal Reserve Bank of St. Louis.
- Anesti, N., M. Garofalo, S. Lloyd, E. Manuel, and J. Reynolds (2023). Unknown measures: Assessing uncertainty around UK inflation using a new Inflation-at-Risk model. Bank Underground, Bank of England.
- Angelini, E., N. Bokan, K. P. Christoffel, M. Ciccarelli, and S. Zimic (2019). Introducing ECB-Base: The blueprint of the new ECB semi-structural model for the Euro Area. Working Paper 2315, European Central Bank.
- Angelini, E., N. Bokan, M. Ciccarelli, M. Lalik, and S. Zimic (2026). The ECB-Multi Country Model. a semi-structural model for forecasting and policy analysis for the largest euro area countries. Working Paper 3119, European Central Bank.
- Anobile, F., F. Frangiamore, M. M. Matarrese, and J. Saadaoui (2025). Investment-at-risk of geopolitical tensions. Working Papers 2025.19, International Network for Economic Research - INFER.
- Azzalini, A. and A. Capitanio (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society (Ser. B)* 65(2), 367–389.
- Bauer, M., T. Berge, G. Fiori, F. Loria, and M. Zhong (2025). Accounting for uncertainty and

- risks in monetary policy. Technical Report 2025-073, Board of Governors of the Federal Reserve System.
- Bell, S., M. Chavaz, B. Hofmann, D. Rees, and M. Rottner (2026, March). Evolving approaches to monetary policy communication in the face of uncertainty: Fan charts, scenarios and guidance. *BIS Quarterly Review*, 17–30.
- Benigno, P. and G. B. Eggertsson (2023). It’s baaack: The surge in inflation in the 2020s and the return of the non-linear Phillips curve. Working Paper 31197, National Bureau of Economic Research.
- Bernanke, B. (2025). Improving Fed communications: A proposal. Working Paper 102, The Brookings Institution Hutchins Center.
- Bernanke, B. S. (2024). Forecasting for monetary policy making and communication at the Bank of England: A review. Technical report, Bank of England.
- Bernanke, B. S., M. Gertler, and S. Gilchrist (1999). The financial accelerator in a quantitative business cycle framework. In J. B. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics*, Volume 1C, pp. 1341–1393. Elsevier.
- Bowe, F., S. J. Kirkeby, I. H. Lindalen, K. A. Matsen, S. S. Meyer, and Ø. Robstad (2023). Quantifying macroeconomic uncertainty in norway. Technical Report 13/2023, Norges Bank.
- Boyarchenko, N., R. K. Crump, L. Elias, and I. Lopez Gaffney (2026). Outlook-at-risk. In D. Aikman and P. Gai (Eds.), *The Research Handbook of Macroprudential Policy*. Edward Elgar. Forthcoming.
- Brayton, F., T. Laubach, and D. Reifschneider (2014). The FRB/US model: A tool for macroeconomic policy analysis. FEDS Notes 2014-04-03, Board of Governors of the Federal Reserve System.
- Brayton, F., A. Levin, R. Tryon, and J. Williams (1997). The evolution of macro models at the federal reserve board. *Carnegie-Rochester Conference Series on Public Policy* 47, 43–81.
- Brayton, F. and P. Tinsley (1996). A guide to FRB/US: A macroeconomic model of the united states. Technical Report 1996-42, Federal Reserve Board.
- Britton, E., P. Fisher, and J. Whitley (1998). The inflation report projections: Understanding the fan chart. Quarterly Bulletin Q1, Bank of England.
- Burgess, S., E. Fernandez-Corugedo, C. Groth, R. Harrison, F. Monti, K. Theodoridis, and M. Sherlock (2013). The Bank of England’s forecasting platform: COMPASS, MAPS, EASE and the suite of models. *Bank of England Working Paper* (471).
- Cairó, I., C. Gust, J. Hebden, E. Herbst, S. Konzem, and G. Nicolò (2025). Risk-adjusted optimal policy for scenario analysis. Manuscript, Board of Governors of the Federal Reserve System, Division of Monetary Affairs.
- Caldara, D., C. Scotti, and M. Zhong (2021). Macroeconomic and financial risks: A tale of mean and volatility. International Finance Discussion Papers 1326, Board of Governors of the Federal Reserve System.

- Carriero, A., T. E. Clark, and M. Marcellino (2024, August). Capturing macro-economic tail risks with Bayesian vector autoregressions. *Journal of Money, Credit and Banking* 56(5), 1099–1127.
- Chernis, T., G. Koop, E. Tallman, and M. West (2024). Decision synthesis in monetary policy. Bank of Canada, Staff Working Paper 2024-30. arXiv:2406.03321.
- Christoffel, K., G. Coenen, and A. Warne (2008). The new Area-Wide Model of the euro area: A micro-founded open-economy model for forecasting and policy analysis. *ECB Working Paper Series* (944).
- Ciccarelli, M., M. Darracq Pariès, R. Priftis, E. Angelini, M. Bańbura, N. Bokan, G. Fagan, J. E. Gumiel, A. Kornprobst, M. Lalik, C. Montes-Galdón, G. Müller, J. Paredes, S. Santoro, A. Warne, S. Zimic, R. Dinis Rigato, H. Kase, I. Koutsoulis, S. Brunotte, S. Cocchi, A. Giammaria, M. Invernizzi, and E. Von-Pine (2024). ECB macroeconomic models for forecasting and policy analysis. Occasional Paper Series 344, European Central Bank.
- Coenen, G., P. Karadi, S. Schmidt, and A. Warne (2018). The New Area-Wide Model II: An extended version of the ECB’s micro-founded model for forecasting and policy analysis with a financial sector. *ECB Working Paper Series* (2200).
- Costain, J. and A. Nakov (2019). Logit price dynamics. *Journal of Money, Credit and Banking* 51(1), 43–78.
- Crump, R. K., S. Eusepi, D. Giannone, E. Qian, and A. M. Sbordone (2025). A large Bayesian VAR of the United States economy. *International Journal of Central Banking* 21, 352–409.
- Del Negro, M. and M. Giannoni (2017). Using dynamic stochastic general equilibrium models at the New York Fed. In R. S. Gürkaynak and C. Tille (Eds.), *DSGE Models in the Conduct of Policy: Use as Intended*. VoxEU.org eBook.
- Domit, S., F. Monti, and A. Sokol (2019). Forecasting the UK economy with a medium-scale Bayesian VAR. *International Journal of Forecasting* 35(4), 1669–1678.
- Edge, R. M., M. T. Kiley, and J.-P. Laforge (2008). An estimated DSGE model of the US economy with an application to natural rate measures. *Journal of Economic Dynamics and Control* 32, 2512–2535.
- Eguren-Martin, F., S. Kösem, G. Maia, and A. Sokol (2024). Targeted financial conditions indices and Growth-at-Risk. Staff Working Paper 1084, Bank of England.
- Elliott, G. and A. Timmermann (2008). Economic forecasting. *Journal of Economic Literature* 46(1), 3–56.
- Engstrom, E. and M. Gonzalez-Astudillo (2017). Time variation in upside and downside risks to the staff baseline forecast. Staff Memo to the Federal Open Market Committee, Board of Governors of the Federal Reserve System.
- Erceg, C. J., L. Guerrieri, and C. Gust (2006, March). SIGMA: A new open economy model for policy analysis. *International Journal of Central Banking* 2(1).
- Federal Open Market Committee (2007). Meeting transcript, August 7, 2007. pp. 60–63.
- Federal Reserve Bank of New York (2007). Blackbook. December 7, 2007, Federal Reserve Bank of

New York.

Federal Reserve Board (2007). Report to the FOMC on Economic Conditions and Monetary Policy. Part 1– Current Economic and Financial Conditions: Summary and Outlook. December 5, 2007, Board of Governors of the Federal Reserve System.

Federal Reserve Board (2018). Report to the FOMC on Economic Conditions and Monetary Policy. Book A– Economic and Financial Conditions: Outlook, Risks, and Policy Strategies. December 7, 2018, Board of Governors of the Federal Reserve System.

Federal Reserve Board (2019a). Report to the FOMC on Economic Conditions and Monetary Policy. Book A– Economic and Financial Conditions: Outlook, Risks, and Policy Strategies. April 19, 2019, Board of Governors of the Federal Reserve System.

Federal Reserve Board (2019b). Report to the FOMC on Economic Conditions and Monetary Policy. Book A– Economic and Financial Conditions: Outlook, Risks, and Policy Strategies. July 19, 2019, Board of Governors of the Federal Reserve System.

Ferrara, L., M. Mogliani, and J.-G. Sahuc (2022). High-frequency monitoring of growth at risk. *International Journal of Forecasting* 38(2), 582–595.

Figueres, J. M. and M. Jarociński (2020). Vulnerable growth in the Euro area: Measuring the financial conditions. *Economics Letters* 191(C), 109–126.

Fischer, S. (2017). I’d rather have Bob Solow than an econometric model, but Speech at the Warwick Economics Summit, Board of Governors of the Federal Reserve System.

Galán Camacho, J. E. (2020). The benefits are at the tail: Uncovering the impact of macroprudential policy on growth-at-risk. Technical Report 2007, Banco de España.

Garga, V., E. Herbst, A. McKay, G. Nicolò, and M. Paustian (2025). Monetary policy, uncertainty, and communications. Finance and Economics Discussion Series 2025-074, Board of Governors of the Federal Reserve System.

Gertler, M. and N. Kiyotaki (2010). Financial intermediation and credit policy in business cycle analysis. In B. M. Friedman and M. Woodford (Eds.), *Handbook of Monetary Economics*, Volume 3A, pp. 547–599. Elsevier.

Gertler, M. L. and P. Karadi (2011). A model of unconventional monetary policy. *Journal of Monetary Economics* 58(1), 17–34.

Gertler, M. L., L. Sala, and A. Trigari (2008). An estimated monetary DSGE model with unemployment and staggered nominal wage bargaining. *Journal of Money, Credit and Banking* 40, 1713–1764.

Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association* 106(494), 746–762.

González-Astudillo, M. and D. Vilán (2019). A new procedure for generating the stochastic simulations in FRB/US. FEDS Notes 2019-03-07, Board of Governors of the Federal Reserve System.

Granger, C. W. J. and M. J. Machina (2006). Forecasting and decision theory. In G. Elliott, C. W. J. Granger, and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Volume 1, pp. 81–98.

Elsevier.

- Harding, M., J. Lindé, and M. Trabandt (2023). Understanding post-COVID inflation dynamics. *Journal of Monetary Economics* 140(S), 101–118.
- Herbst, E., S. Konzem, and C. Scofield (2026). Alternative scenarios at the Federal Reserve from 1968 to 2020: Data, interpretation, and evaluation. Technical Report 2026.033, Board of Governors of the Federal Reserve System.
- IMF (2017). *Global financial stability report: Is growth a risk?* International Monetary Fund.
- Johnson, M. C. and M. West (2025). Bayesian predictive synthesis with outcome-dependent pools. *Statistical Science* 40(1), 109–127.
- Jondeau, E., P. Poncet, and C. Rebillard (2022). Are financial variables useful to complement GDP nowcasting? Eco Notepad, Banque de France.
- Keijsers, B. and D. van Dijk (2025). Does economic uncertainty predict real activity in real time? *International Journal of Forecasting* 41(2), 748–762.
- Kiley, M. T. (2022). Unemployment risk. *Journal of Money, Credit and Banking* 54(5), 1407–1424.
- Kiyotaki, N. and J. Moore (1997). Credit cycles. *Journal of Political Economy* 105(2), 211–248.
- Laxton, D., H. Igityan, and S. Mkhatrshvili (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, 452–483.
- Lenza, M., I. Moutachaker, and J. Paredes (2025). Density forecasts of inflation: A quantile regression forest approach. *European Economic Review* 178, 105079.
- Lhuissier, S. (2022). Financial conditions and macroeconomic downside risks in the euro area. *European Economic Review* 143, 104046.
- Lhuissier, S. (2026). Towards a method for guiding monetary policy in times of great uncertainty. Eco Notepad Post No. 439, Banque de France.
- Lloyd, S., G. Mantoan, and E. Manuel (2022). When growth-at-risk hits the fan: Comparing quantile-regression predictive densities with committee fan charts. Bank of England, preliminary draft, May 2022.
- Lloyd, S., E. Manuel, and K. Panchev (2024). Foreign vulnerabilities, domestic risks: The global drivers of gdp-at-risk. *IMF Economic Review* 72(1), 335–392.
- López-Salido, D. and F. Loria (2024). Inflation at risk. *Journal of Monetary Economics* 145.
- Makabe, Y. and Y. Norimasa (2022). The term structure of inflation at risk: A panel quantile regression approach. Technical Report 22-E-4, Bank of Japan.
- Nickel, C., J. Kilponen, E. Moral-Benito, G. Koester, et al. (2025). ECB monetary policy strategy 2025: A strategic view on the economic and inflation environment in the euro area. Technical Report 371, European Central Bank.
- O’Brien, M. and M. Wosser (2021). Growth at risk & financial stability. Technical Report 2, Central Bank of Ireland.
- Plaasch, J. and A. Röthig (2025). A growth-at-risk model for the German economy. Technical

- Report 05/2025, Deutsche Bundesbank.
- Schröder, M. (2025). Mixing it up: Inflation at risk. Working paper, Norges Bank.
- Scotti, C. (2023). Financial shocks in an uncertain economy. Working Paper 2308, Federal Reserve Bank of Dallas.
- Sims, C. A. (2002). The role of models and probabilities in the monetary policy process. *Brookings Papers on Economic Activity* 33(2), 1–62.
- Tallman, E. and M. West (2023). Bayesian predictive decision synthesis. *Journal of the Royal Statistical Society (Ser. B)* 86(2), 340–363.
- Tetlow, R. J. and B. Ironside (2007). Real-time model uncertainty in the United States: The Fed, 1996–2003. *Journal of Money, Credit and Banking* 39(7), 1533–1561.
- Walker, W. E., R. J. Lempert, and J. H. Kwakkel (2013). Deep uncertainty. In S. I. Gass and M. C. Fu (Eds.), *Encyclopedia of Operations Research and Management Science* (3rd ed.), pp. 395–402. Boston, MA: Springer.
- West, M. and P. J. Harrison (1997). *Bayesian Forecasting and Dynamic Models* (2nd ed.). Springer.
- Wieland, V., E. Afanasyeva, M. Kuete, and J. Yoo (2016). New methods for macro-financial model comparison and policy analysis. In *Handbook of Macroeconomics*, Volume 2, pp. 1241–1319. Elsevier.
- Wieland, V., T. Cwik, G. J. Müller, S. Schmidt, and M. Wolters (2012). A new comparative approach to macroeconomic modeling and policy analysis. *Journal of Economic Behavior and Organization* 83(3), 523–541.

Supplementary material for the paper:

Risks and Uncertainty in Monetary Policy

TOBIAS ADRIAN
International Monetary Fund
tadrian@imf.org

DOMENICO GIANNONE
Johns Hopkins University
domenico.giannone@jhu.edu

MATTEO LUCIANI
Federal Reserve Board
matteo.luciani@frb.gov

MIKE WEST
Duke University
mike.west@duke.edu

Appendix A Scenario synthesis implementation

This appendix summarizes the main theoretical and computational steps behind the empirical implementation of the Scenario Synthesis. We first describe the univariate implementation and then turn to the multivariate case used in the joint analysis of GDP growth and core PCE inflation.

A.1 Univariate analysis

In the univariate applications, the synthesis is implemented as follows:

- S1: Generate a large random sample $\mathbf{y}^i$, ($i = 1:n$), from the reference $p(\mathbf{y})$, to define an importance sample for Monte Carlo evaluation of the baseline $p_0(\mathbf{y})$ and the scenarios $p_j(\mathbf{y})$.
- S2: Evaluate baseline IS weights $w_0^i \propto p_0(\mathbf{y}^i)/p(\mathbf{y}^i)$, subject to normalization.
- S3: Evaluate the scenario p.d.f.s $p_j(\mathbf{y})$ for $j > 0$ as in Step 2 now applied to scenario p.d.f.s $p_j(\mathbf{y})$ instead of the baseline $p_0(\mathbf{y})$. For each $\mathcal{S}_j$ this delivers normalized IS weights w_j^i on the reference sample values.
- S4: Compute synthesis weights by maximising

$$\{\alpha_j^*\}_{j=0}^J = \arg \max_{\substack{\alpha_j > 0, \alpha_0 \geq \alpha_j \\ \sum_{j=0}^J \alpha_j = 1}} \left[\log\{\pi_{pf}(\boldsymbol{\alpha})\} + \epsilon \sum_{j=0}^J \log(\alpha_j) \right]$$

where

$$\pi_{pf}(\boldsymbol{\alpha}) = \frac{1}{n} \sum_{i=1}^n \frac{w_f^i(\boldsymbol{\alpha})}{w_f^i(\boldsymbol{\alpha}) + w_p^i},$$

DISCLAIMER: The views expressed in this paper are those of the authors and do not necessarily reflect the views and policies of the Board of Governors, the Federal Reserve System, or the International Monetary Fund, its Management, or its Executive Directors.

$w_f^i(\boldsymbol{\alpha}) = \sum_{j=0}^J \alpha_j w_j^i$ are the synthesis IS weights and $w_p^i = 1/n$ are the (uniform) reference weights.

A.2 Multivariate analysis

Š1: Generate a large random sample $\mathbf{y}^i$, ($i = 1:n$), from the Reference distribution using the skew- t copula approach, as described in Appendix A.2.1.

Š2: Evaluate baseline IS weights:

$$w_0^i \propto \frac{p_0(\mathbf{y}^i)}{p(\mathbf{y}^i)},$$

subject to normalization, where $p_0(\mathbf{y})$ and $p(\mathbf{y})$ are computed using the exact parametric copula density formula in Appendix A.2.2.

Š3: Evaluate scenario IS weights:

$$w_j^i \propto \frac{p_j(\mathbf{y}^i)}{p(\mathbf{y}^i)},$$

subject to normalization for each scenario j , using the exact parametric copula density.

Š4: Compute synthesis weights as in S4.

ALTERNATIVE MULTIVARIATE DISTRIBUTION. One could instead use the [Azzalini and Capitanio \(2003\)](#) multivariate skew- t distribution. However, its marginals are not univariate skew- t with the specified parameters $(\xi_k, \omega_k, \alpha_k, \nu)$ —the effective marginal skewness depends on the full parameter vector and correlation structure. Since we require exact control over each marginal to match the quantiles of the Reference and Baseline distributions, the copula approach is necessary.

IGNORING DEPENDENCE. Another alternative is to ignore dependence altogether and work only with the marginals. In this case, the IS weights for scenario $j = 0, \dots, J$ are obtained as $w_j^i \propto \prod_{k=1}^2 \frac{p_{j,k}(\mathbf{y}^i)}{p_k(\mathbf{y}^i)}$. Potentially, one could even place different weights on the importance of matching the reference across variables by estimating the IS weights as $w_j^i \propto \prod_{k=1}^2 \left(\frac{p_{j,k}(\mathbf{y}^i)}{p_k(\mathbf{y}^i)} \right)^{\gamma_k}$, where $\gamma_k \geq 0$ governs the importance assigned to matching that variable—setting $\gamma_k = 0$ excludes variable k from the reweighting; $\gamma_k = 1$ assigns it full weight.

A.2.1 Skew- t copula sampling

Without loss of generality, let $\mathbf{y} = (y_1, y_2)'$ be a 2×1 vector.

- Each marginal y_k follows a univariate skew- t distribution: $y_k \sim \text{ST}(\xi_k, \omega_k, \alpha_k, \nu_k)$, where ξ_k is location, ω_k is scale, α_k is skewness, and ν_k is degrees of freedom.
- The joint distribution is constructed using a copula with correlation matrix $\boldsymbol{\rho}$.
- To generate samples:
 - Generate $\mathbf{z} \sim N(\mathbf{0}, \boldsymbol{\rho})$ from multivariate normal with correlation $\boldsymbol{\rho}$.
 - Transform to uniform: $u_k = \Phi(z_k)$ for $k = 1, 2$, where Φ is the standard normal CDF.
 - Transform to skew- t marginals: $y_k = F_k^{-1}(u_k)$, where F_k is the skew- t CDF with parameters $(\xi_k, \omega_k, \alpha_k, \nu_k)$.
- This procedure guarantees exact marginal distributions $y_k \sim \text{ST}(\xi_k, \omega_k, \alpha_k, \nu_k)$ with specified

correlation structure.

COPULA CHOICE. In the empirical application we used a Gaussian copula for computational efficiency and because in our specific application the correlation structure is the same across all distributions. The t -copula could alternatively be used if tail dependence varies across distributions.

A.2.2 Copula density evaluation

For importance sampling, we require exact density evaluation $p(\mathbf{y})$ at sample points.

- The joint density decomposes as:

$$p(\mathbf{y}) = c(F_1(y_1), F_2(y_2)) \times f_1(y_1) \times f_2(y_2),$$

where f_k are the skew- t marginal densities, F_k are the skew- t marginal CDFs, and c is the copula density.

- For Gaussian copula with correlation $\boldsymbol{\rho}$:

$$c(u_1, u_2) = \frac{1}{\sqrt{|\boldsymbol{\rho}|}} \exp \left\{ -\frac{1}{2} \mathbf{z}' (\boldsymbol{\rho}^{-1} - \mathbf{I}) \mathbf{z} \right\}, \quad \text{where } z_k = \Phi^{-1}(u_k).$$

- In log-space (for numerical stability):

$$\log p(\mathbf{y}) = \log c(F_1(y_1), F_2(y_2)) + \sum_{k=1}^2 \log f_k(y_k).$$

- This exact parametric density is used to compute importance weights $w^i = p(\mathbf{y}^i)/q(\mathbf{y}^i)$ for any two distributions p and q with known parameters.

A.3 Calibrating the copula correlation matrix

Let $\mathbf{y}_{t+h}^{(m)} = (y_{1,t+h}^{(m)}, \dots, y_{K,t+h}^{(m)})'$, $m = 1, \dots, M$, denote the m th draw from the Large BVAR predictive distribution at horizon h for the K variables of interest.

- Stack the simulated vectors row-wise into the $M \times K$ matrix

$$\mathbf{Y}_{t+h} = \begin{pmatrix} (\mathbf{y}_{t+h}^{(1)})' & \dots & (\mathbf{y}_{t+h}^{(M)})' \end{pmatrix}'.$$

- The copula dependence matrix is defined as the sample Pearson correlation matrix across simulation draws:

$$\boldsymbol{\rho} = \text{corr}(\mathbf{Y}_{t+h}).$$

- This matrix is treated as fixed and is used in the Gaussian copula for the Reference, the Baseline, and all scenario distributions.
- Hence, the Large BVAR is used only to discipline the cross-sectional dependence structure, while the marginals are imposed separately using the fitted univariate skew- t distributions.

Appendix B Predictive distribution for Core PCE price inflation

OaR provides distributions for GDP growth and CPI inflation. Because the Tealbook scenarios are defined for core PCE price inflation, we convert the OaR CPI density into a core PCE density. We first shift the OaR CPI median by the difference between the Blue Chip core-PCE and CPI forecasts. We then rescale the CPI quantiles so that, for each symmetric quantile pair, the core-PCE-to-CPI interquantile-range ratio matches the corresponding ratio implied by the VAR forecast. This transformation preserves the asymmetry of the OaR CPI density in the constructed core-PCE reference.

STEP 1: SCALING THE MEDIAN. Let $\mathcal{Q}_{1,\tau}^R$ be the τ -quantile from the Reference for core PCE price inflation, and $\mathcal{Q}_{2,\tau}^R$ be the same object for CPI inflation. Let y_{1t+h}^{BC} be the Blue Chip forecast for core PCE price inflation. Then,

$$\mathcal{Q}_{1,50}^R = \mathcal{Q}_{2,50}^R + (y_{1t+h}^{BC} - y_{2t+h}^{BC})$$

STEP 2: SCALING THE OTHER QUANTILES. Let $\mathcal{Q}_{1,\tau}^V$ be τ -quantile from the empirical distribution of $y_{1t+h}^{(d)}$, where $y_{1t+h}^{(d)}$, $d = 1, \dots, D$ is the d -th draw of the unconditional forecast of core PCE price inflation from the VAR. Then, take the case of the 25-th and 75-th quantile. Let

$$\theta_1 = (\mathcal{Q}_{2,\tau}^R - \mathcal{Q}_{2,(1-\tau)}^R) \frac{\mathcal{Q}_{1,\tau}^V - \mathcal{Q}_{1,(1-\tau)}^V}{\mathcal{Q}_{2,\tau}^V - \mathcal{Q}_{2,(1-\tau)}^V}$$

be the desired inter-quantile range in the core PCE price inflation Reference, and let

$$\theta_2 = \frac{\mathcal{Q}_{2,\tau}^R + \mathcal{Q}_{2,(1-\tau)}^R - 2\mathcal{Q}_{2,50}^R}{\mathcal{Q}_{2,\tau}^R - \mathcal{Q}_{2,(1-\tau)}^R}$$

be the τ -quantile based skewness in the Reference for CPI inflation, then

$$\begin{aligned} \mathcal{Q}_{1,(1-\tau)}^R &= \mathcal{Q}_{1,50}^R + \frac{\theta_1(\theta_2 - 1)}{2} \\ \mathcal{Q}_{1,\tau}^R &= \mathcal{Q}_{1,(1-\tau)}^R + \theta_1. \end{aligned}$$

Appendix C Extracting probabilities from the Blackbook

The probability estimates for the December 7, 2007 FRBNY Blackbook are approximate values obtain through Optical Character Recognition via ClaudeAI from the bar chart visualizations. To refine the first step estimate we provide Claude with the true SPF (Survey of Professional Forecasters) values. Specifically, we used the following three prompts to extract the numbers

PROMPT 1: Can you extract the numbers in the bars in the attached charts?

PROMPT 2: Now I can provide you with the correct numbers for SPF so that you can revise those for FRBNY.

PROMPT 3: Can you extract those numbers with more precision? say 1 decimal?

Table C1: OPTICAL CHARACTER RECOGNITION RESULTS

Range	SPF		FRBNY		
	1 st e.	True	1st e.	2nd e.	3rd e.
<-2.0	0	0.22	4	4	4.5
-2.0 to -1.0	0	0.44	3	3	3.5
-1.0 to 0.0	2	2.13	8	8	8.5
0.0 to 1.0	7	7.28	18	18	18.0
1.0 to 2.0	25	25.00	24	24	24.0
2.0 to 3.0	45	45.02	21	20	20.5
3.0 to 4.0	18	17.18	15	15	15.0
4.0 to 5.0	3	2.10	6	6	5.5
5.0 to 6.0	0	0.47	1	1	0.5
> 6.0	0	0.16	0	1	0.0

NOTE: The table compares probability values extracted from the Blackbook histogram using successive OCR prompts. For the SPF histogram, the column labeled “True” reports the probabilities implied by the published SPF distribution and is used as a benchmark. For the FRBNY histogram, “1st e.”, “2nd e.”, and “3rd e.” report the first, second, and third OCR-based extractions, respectively. All entries are percentages.

Table C1 shows the results we obtained from the different prompts.

Appendix D How to convert annual-average GDP growth into quantile forecasts for Q4/Q4 GDP growth

Let i denote GDP growth and $r \in \{\text{BB}, \text{SPF}\}$ denote the two reference distributions. Each reference provides the set of J quantile pairs shown in Table C1; we denote them as $\mathcal{Q}^{(i,r)} = \{(q_j^{(i)}, p_j^{(i,r)})\}_{j=1}^J$, where $q_j^{(i)}$ is an outcome value (e.g., GDP growth of 2%) and $p_j^{(i,r)} = \Pr(\tilde{y}^{(i)} \leq q_j^{(i)})$ is the corresponding cumulative probability extracted from the Blackbook fan chart or the SPF histogram. Furthermore, let Y_t be the n -dimensional vector of macroeconomic variables used to estimate the VAR. The model is estimated with Bayesian methods with $N = 20,000$ posterior draws. From the VAR, we generate conditional forecasts $\{y_{T+h}^{(i,d)}\}_{d=1}^N$. Then, we converted annual-average GDP growth into quantile forecasts for Q4/Q4 GDP growth as follows:

STEP 1: SKEW- t FIT THE REFERENCE QUANTILES. For each reference r , we fit a skew- t distribution with parameters $\theta^{(r)} = (\xi, \omega, \alpha, \nu)$ on $\mathcal{Q}^{(i,r)}$ as in Adrian et al. (2019). This delivers a smooth fitted c.d.f. $\hat{F}_{\text{ST}}^{(r)}$ and the corresponding p.d.f. $\hat{f}_{\text{ST}}^{(r)}$.

STEP 2: CALENDAR-YEAR YEAR-OVER-YEAR GROWTH RATES We construct annual-average growth rates from the quarterly VAR forecasts. Let $Y_t^{(i)}$ denote the log-level of variable i in quarter t . The annual-average growth rate at quarter t is defined as:

$$\tilde{y}_t^{(i)} = \frac{1}{4} \sum_{k=0}^3 Y_{t-k}^{(i)} - \frac{1}{4} \sum_{k=4}^7 Y_{t-k}^{(i)}. \quad (\text{D1})$$

Applied to the VAR forecasts, this delivers a N -draw sample $\{\tilde{y}_H^{(i,d)}\}_{d=1}^N$ of the CY-YoY growth rate at horizon $H = T + 4$.

STEP 3: OPTIMAL TRANSPORT (QUANTILE MAPPING) We tilt the N VAR draws of the annual-average growth rate $\{\tilde{y}_H^{(i,d)}\}_{d=1}^N$ to match the reference marginal distribution using a rank-preserving (monotone) map. Let $\hat{F}_{\text{VAR}}^{(i)}$ denote the empirical c.d.f. of the VAR draws. For each draw d :

$$\tilde{x}_{\text{ref}}^{(i,d)} = (\hat{F}_{\text{ST}}^{(i,r)})^{-1} \left(\hat{F}_{\text{VAR}}^{(i)} \left(\tilde{y}_H^{(i,d)} \right) \right). \quad (\text{D2})$$

This replaces the VAR marginal distribution with the reference distribution while preserving the rank ordering of draws, so that the joint dependence structure of the VAR is retained across variables.

STEP 4: CONDITIONAL BVAR FORECASTS: The tilted annual draws $\{\tilde{x}_{\text{ref}}^{(i,d)}\}_{d=1}^N$ are imposed as hard constraints on the augmented VAR state space. The VAR is iterated forward draw-by-draw to recover the corresponding 4-quarter percent change:

$$\Delta_4 y_H^{(i,d)} = y_H^{(i,d)} - y_{H-4}^{(i)}. \quad (\text{D3})$$

This yields a draw sample $\{\Delta_4 y_H^{(i,d)}\}_{d=1}^N$ that is internally consistent with the VAR dynamics and anchored to the reference marginal at the annual frequency.

STEP 5: QUANTILES ESTIMATION A small set of quantiles is extracted from the N draws at fixed probability levels $\{p_1, \dots, p_K\}$: $\hat{q}_k^{(i)} = \text{Quantile} \left(\{\Delta_4 y_H^{(i,d)}\}_{d=1}^N, p_k \right), \quad k = 1, \dots, K.$