Entropic Tilting of Forecasts to SPF Histograms: Analytics & Applications*

Todd E. Clark,[1] Elmar Mertens[2]

[1]*Johns Hopkins University and Federal Reserve Bank of Cleveland,* [2]*European Central Bank.*

November 9, 2025

Abstract

This paper develops a new, direct approach to entropic tilting of model-based predictive distributions to match histogram forecasts provided in the U.S. Survey of Professional Forecasters (SPF). We focus on tilting to histogram probabilities directly, rather than to moments of fitted distributions. We reformulate the single-histogram tilting problem and derive a novel analytic characterization for the multiple-histogram case, with iterative solutions via Iterative Proportional Fitting. Application to quarterly real-time forecasts of major macroeconomic aggregates from a Bayesian vector autoregression with time-varying volatility shows that tilting to SPF histograms significantly improves on the model's baseline forecasts, particularly during periods around the Great Recession and the COVID-19 pandemic.

*We gratefully acknowledge helpful suggestions and comments received from Dean Croushore, Domenico Giannone, Michele Lenza, Simon van Norden, Bernd Schwaab, and conference and seminar participants at the Real Time Econometrics conference 2025 (Eltville) and the ECB. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the European Central Bank, the Eurosystem, the Federal Reserve Bank of Cleveland, or the Federal Reserve System.

1 Introduction

An immense literature has explored the predictive content of forecasts from the U.S. Survey of Professional Forecasters (SPF) or its counterpart in the euro area. Within this body of work, many studies have focused on survey point forecasts, resulting in these forecasts being widely recognized for their real-time accuracy. While less studied, the SPF (both in the U.S. and euro area) also includes density forecasts, expressed as histograms over forecast bins. The merits of these density forecasts have been more widely debated. Some research has found value in these projections. For example, Clements (2004, 2014) finds that such histograms can offer informative density forecasts, often competitive with or superior to parametric model-based alternatives during periods of heightened uncertainty. On the other hand, studies such as Clements (2008) and Engelberg, Manski, and Williams (2009) have noted several limitations of histogram forecasts, including issues of coherence, underdispersion, and resolution due to fixed binning and finite responses.[1] However, it should be noted that some of these issues are more prevalent in individual forecasters' histograms than in the aggregate histograms that average across forecasters.

One of the ways that survey forecasts have been explored or used is to try to improve forecasts from time series models, such as vector autoregressions, by combining the alternative projections. For example, Faust and Wright (2013) use current-quarter point forecasts from surveys as jumping off points to enhance model-based forecasts. More formally, point forecasts from surveys can be brought to inform model forecasts through (Gaussian) conditional forecasting methods, with model forecasts conditioning on survey forecasts. A closely related, less parametric approach is to use the method of entropic tilting introduced into economic forecasting by Robertson, Tallman, and Whiteman (2005). In the tilting approach, the model's predictive distribution is modified so that the relevant moments of the modified distribution match the chosen moments from the survey forecasts. Our

[1] Relatedly, see also Knüppel and Pavlova (2023). In addition, Lenza, Moutachaker, and Paredes (2025) develop a quantile regression forecast model that improves — at short horizons — on the density forecast accuracy of the ECB's SPF for the euro area.

paper rather emphasizes the application of entropic tilting to directly match probabilistic forecasts, as opposed to matching the moments of an underlying distribution.

Most studies that entropically tilt model-based forecasts toward SPF information focus on a survey's point forecasts, and interpret these as targets for matching the predictive mean. As examples, Krueger, Clark, and Ravazzolo (2017) entropically tilt forecasts from a Bayesian vector autoregression (BVAR) to (point) nowcasts from the SPF, and Tallman and Zaman (2020) tilt BVAR forecasts to (point) long-run survey forecasts.[2] The existing work that uses density forecasts from the SPF to inform model-based forecasts does so by extracting a limited set of moments from the SPF histograms. Galvão, Garratt, and Mitchell (2021) and Banbura, et al. (2021) proceed by fitting continuous distributions (parametrically and non-parametrically, respectively) to summarize histogram information in the form of first or second moments from the fitted distributions.[3] Ganics and Odendahl (2021) impute first and second moments under the assumption that all mass within a given histogram bin is concentrated at the center of the bin. All of this work on combining survey and model forecasts uses aggregate (average or median) forecasts from the SPF.

This paper develops and examines a different, direct approach to combining information in density forecasts of the SPF with model-based forecasts. Our approach takes the SPF histograms as given and avoids the need to fit distributions to the histograms and rely on moments from the fitted distribution. Accordingly, our approach does not require taking a stand on how the SPF forecaster distributes predictive mass within each histogram bin or taking a stand on a parametric or non-parametric distribution to characterize the histogram and then in turn estimating the distribution. Instead, our approach directly tilts the predictive distribution of the BVAR to the SPF histogram bins, using information directly in the SPF and nothing more to set the tilting targets. Throughout, we focus

[2]Additional applications of entropic tilting to BVAR forecasts include Cogley, Morozov, and Sargent (2005), and Bobeica and Hartwig (2023).

[3]Specifically, Galvão, Garratt, and Mitchell (2021) extract moment parameters from fitted generalized beta distributions (Engelberg, Manski, and Williams, 2009), whereas Banbura, et al. (2021) employ random sampling from multinomial distributions and kernel-smoothed estimates from the resulting densities.

our application to histograms of the average SPF forecaster, also known as "consensus" histograms.[4]

Our paper has two primary contributions. The first is methodological: We develop a new analytical characterization for the tilting problem when targeting histogram bins. As noted already by Tallman and West (2022), the tilting solution is particularly simple when only a single histogram is targeted, and reweighs model draws within each bin to match SPF-provided bin probabilities. Compared to their work, we reformulate the solution in the single-histogram case, and extend the solution to the case of multiple histograms. In addition, we discuss implementation details.

Our second contribution is empirical: Focusing on the U.S., we examine the information content of SPF histogram forecasts for the accuracy of forecasts of major macroeconomic aggregates produced by a Bayesian vector autoregression. Using our analytical solution for the tilting problem, we entropically tilt the BVAR's quarterly forecast distributions to SPF histograms, using the available multi-year forecast horizons from the SPF for three variables, including GDP growth, the unemployment rate, and core PCE inflation.[5] The BVAR includes stochastic volatility with fat-tailed innovations to the VAR and volatility outliers, a specification recently developed by Carriero, et al. (2024) and found to forecast well in a long historical sample, including the high volatility induced by the COVID-19 pandemic. We provide evidence that, relative to the baseline model forecasts, SPF histogram forecasts add the most information content around the Great Recession and its aftermath during the COVID-19 pandemic, with most of the information coming from the SPF's unemployment rate forecasts from one through three years ahead. As our baseline already allows for some flexibility in adapting to changing economic conditions, in particular during times of turbulence, the gains from tilting densities from less flexible

[4]For a discussion of potential difficulties in handling individual SPF histograms see, for example, Bassetti, Casarin, and Del Negro (2023) and Clark and Mertens (2024).

[5]Our development and use of an analytical solution to the problem of tilting to histograms and application to BVAR forecasts distinguishes this paper from the earlier work of Clark, Ganics, and Mertens (2022), which instead relied on the standard numerical solution of the tilting problem and tilted SPF-consistent point forecasts obtained from a model specified in state-space form to the SPF histograms.

models are, or course, larger.[6]

Our examination of the point and density accuracy of forecasts from the tilted distribution to the baseline distribution of the BVAR shows that, on average over the period from 1996 through 2025, entropically tilting BVAR forecasts to the annual SPF histograms consistently improves the BVAR's quarterly forecast accuracy, with the sizes of the gains varying over horizons and across variables. Drilling into accuracy over time shows that the benefits of entropic tilting BVAR-based forecasts to SPF forecast histograms are driven by the period since 2007 (i.e., since the Great Recession), perhaps related to the fact that unemployment rate forecasts and longer-horizon forecasts did not become available to inform the tilted forecasts until 2009, when their SPF publication began.[7] While helpful to forecast accuracy on average for the 2007-2025 sample, the benefits of entropically tilting the BVAR forecasts to the SPF histograms are driven by the years around the Great Recession and around COVID's outbreak in 2020. In more normal periods, the BVAR's forecasts are not improved by bringing information from the SPF to bear through tilting to the survey's histogram forecasts.

The paper proceeds as follows. Section 2 details the connection of our approach to other work on the method of entropic tilting. Section 3 develops our analytical solution to tilting to histogram forecasts. Section 4 describes the data used and details the forecasting BVAR. Section 5 presents empirical results of tilting BVAR forecasts to SPF density forecasts. Section 6 concludes. Additional details and results are provided in a supplementary online appendix (pending).

[6]Corresponding results for tilting a standard Gaussian BVAR are available upon request. See also the results of Bobeica and Hartwig (2023) related to robustness of BVARs tilted to the ECB's SPF during the COVID-19 pandemic.

[7]Relatedly, the SPF has collected probabilistic forecasts for core PCE inflation only since 2007.

2 Related literature

Our use of entropic tilting to align model-implied forecast densities with survey histograms particularly across multiple forecast horizons can be viewed as a modern and applied extension of the classical *Iterative Proportional Fitting* (IPF) algorithm. In our methodology, IPF serves to iteratively reweight a joint distribution so that its marginals match externally imposed constraints (e.g., SPF histogram bins), while minimizing the Kullback-Leibler (KL) divergence from a prior distribution. This structure of iterative KL projection is rooted in foundational work on IPF (Deming and Stephan, 1940), later formalized as an I-projection by Csiszár (1975) and given geometric interpretation by Fienberg (1970). Our approach also connects naturally to the literature on empirical likelihood and exponential tilting. Schennach (2007) demonstrates how exponential tilting arises in moment-based estimation problems as a dual representation of KL minimization.

A related strand develops the theoretical and structural underpinnings of exponential-tilting adjustments. Giacomini and Ragusa (2014) apply the same Kullback-Leibler-projection principle to impose nonlinear theoretical moment conditions, such as Euler-equation restrictions, on model-based forecast densities. Their "theory-coherent forecasting" approach and our framework share identical entropic geometry but differ in purpose: theirs enforces structural coherence with economic theory, ours empirical coherence with survey-implied probability mass functions. Antolín-Díaz, Petrella, and Rubio-Ramírez (2021) further show that conditional-forecasting and structural-scenario analyses in SVARs can likewise be derived as KL-minimizing projections, providing a unifying information-theoretic view of constrained forecasting across theoretical, structural, and empirical contexts.

Our methodology builds directly on and extends the literature on entropic tilting and constrained Bayesian forecasting, including key contributions by Mike West and coauthors. In particular, Tallman and West (2022) develop a foundational treatment of entropic tilting as a Bayesian decision-analytic method for modifying forecast distributions un-

der expectation constraints, including cases with quantile or indicator function targets. Their analytic treatment of tilting under histogram- or quantile-based constraints shares the same structure as our own KL projection approach, though our implementation emphasizes iterative reweighting across multiple marginals, leveraging computational tools from iterative proportional fitting. The related work of West (2024) further generalizes constrained forecasting within a decision-theoretic framework, advocating entropic tilting as a principled alternative to naive probabilistic conditioning. This distinction is particularly relevant in our context, where survey histograms represent external, judgment-based constraints not generated within the probabilistic system of the model itself. Our approach complements the scenario synthesis framework of Adrian, et al. (2025), which employs entropic tilting to reconcile a baseline forecast with externally specified scenarios. While their emphasis lies in evaluating the concordance of narrative scenarios with a statistical reference using expected misclassification rates, our contribution focuses on reconciling high-frequency density forecasts with survey-based marginal distributions for multiple variables and multiple forecast horizons.

More broadly, our paper also contributes to the literature that evaluates the predictive content of probabilistic survey forecasts. A large part of this literature has examined heterogeneity, consistency, and performance characteristics across individual respondents.[8] By contrast, and similar to the literature surveyed in the previous section, we investigate the predictive value of aggregated probability forecasts, where issues of coherence and consistency are fairly negligible, as we document below.

Our focus on directly tilting model-based forecasts to match histogram probabilities complements this literature by providing a practical method to leverage the information content of survey histograms without imposing additional parametric assumptions. As such, our out-of-sample analysis differs from that of Giacomini and Ragusa (2014), who evaluate

[8] See, for example, Engelberg, Manski, and Williams (2009), Clements (2010), Rich, Song, and Tracy (2012), Kenny, Kostka, and Masera (2014), Kenny, Kostka, and Masera (2015), Meyler (2020), Bassetti, Casarin, and Del Negro (2022, 2023), Coroneo, Iacone, and Profumo (2024), Allayioti, et al. (2024), Clements, Rich, and Tracy (2025), and Clements (2025).

entropic tilting as a means of imposing theory-coherent moment restrictions—such as Euler equations—on model-based predictive densities. Their exercise tests whether such structural coherence improves forecast performance in a frequentist rolling-window setting. By contrast, we evaluate empirical coherence: our tilting aligns Bayesian predictive densities with survey-implied probability distributions from the SPF and assesses the resulting forecasts using standard density-forecast scoring rules. Thus, whereas Giacomini and Ragusa's analysis validates theoretical consistency, our framework tests calibration gains from survey consistency within a Bayesian, analytically tractable implementation.

By evaluating the forecast performance of tilted distributions, we contribute empirical evidence on the value of survey-based density information in enhancing model-based forecasts of key macroeconomic variables. Moreover, we document the added value of survey histograms particularly during periods of economic turbulence, such as the Great Recession and the COVID-19 pandemic, relative to state-of-the-art time-series models highlighting the practical relevance of our approach for real-world forecasting challenges.

3 Analytical tilting to histograms

Here, we present our new analytical solution to the problem of tilting to histograms. We begin with describing the basic setup of the tilting problem. We then take up the solution to the problem. Given its particular simplicity, we first discuss the case of tilting to a single histogram. We then describe the generalization to two and more histograms. In addition, we discuss our computational implementation of the tilting solution in the occasional cases in which the original predictive density from the model does not have mass in some of the target bins from SPF histogram forecasts.

3.1 Setup of the tilting problem

Let y_t denote the variable of interest, such as the quarter-on-quarter log growth rate of GDP. At a forecast origin of period t, estimates of a BVAR (some other model could be used) provide a baseline predictive density for future outcomes, which we denote $p(\boldsymbol{X}_t)$, where $\boldsymbol{X}_t$ is a (typically) multivariate vector of outcomes, such as GDP growth in different quarters.

As detailed with examples in Section 4, for a selected set of variables and forecast horizons, the SPF collects and publishes forecasts of probability distributions, in the form of histograms with pre-specified bins of outcome ranges. More formally, for N histograms, each with J bins, A_j^n, the SPF collects and publishes

$$\widehat{\text{Prob}}_t \left(b_{j-1}^n < f_t^n \left(\boldsymbol{X}_t \right) \leq b_j^n \right) \forall \; j, n,$$

where the b terms refer to lower and upper limits of the SPF-specified bins and $f_t^n \left(\boldsymbol{X}_t \right)$ denotes an outcome that is a (non-)linear transformation of the vector of underlying forecasts $\boldsymbol{X}_t$. Let $A_{t,j}^n$ denote the set of outcomes inside the j-th bin of histogram n:

$$A_{t,j}^n \equiv \{ \boldsymbol{X}_t \; : \; b_{t,j-1}^n < f_t^n \left(\boldsymbol{X}_t \right) \leq b_{t,j}^n \}, \tag{1}$$

where $b_{t,j}^n$ denotes the upper limit for the j bin, the lower limit of the first bin is $b_{t,0}^n$, and open-ended bins are represented as $b_{t,0}^n = -\infty$ and $b_{t,J}^n = \infty$.

If histogram n targets the unemployment rate, $f_t^n \left(\boldsymbol{X}_t \right)$ is the linear average of the quarterly unemployment rates for the targeted year contained in $\boldsymbol{X}_t$. If histogram n targets GDP growth, $f_t^n \left(\boldsymbol{X}_t \right)$ is the simple percentage change in the annual average level of real GDP, which involves the ratio of average annual levels of real GDP between the targeted and the previous year.

Turning back to the general setup, for simplicity we drop the t subscripts on the set of

outcomes, bin limits, $\boldsymbol{X}$, and the transformation function $f(\cdot)$. Each histogram provides probabilities for a set of J mutually exclusive and exhaustive events $\{A_j\}_{j=1}^J$. For the baseline density, $p(A_j) \equiv p(\boldsymbol{X} \in A_j)$ describes the probabilities for $\boldsymbol{X}$ to fall in the jth bin of the histogram. Since the events are mutually exclusive and exhaustive we have $\sum_{j=1}^J p(A_j) = 1$. The tilted density $q(\boldsymbol{X})$ is required to match the histogram probabilities, so that $q(A_j) \equiv q(\boldsymbol{X} \in A_j)$ are the target probabilities from the SPF histogram, with $\sum_{j=1}^J q(A_j) = 1$. We seek to construct a tilted density, $q(\boldsymbol{X})$, that minimizes the relative entropy (or Kullback-Leibler divergence measure)

$$D\left(q(\boldsymbol{X}) \,||\, p(\boldsymbol{X})\right) = E_x^q \, \log\left(\frac{q(\boldsymbol{X})}{p(\boldsymbol{X})}\right) = \int_x q(\boldsymbol{X}) \, \log\left(\frac{q(\boldsymbol{X})}{p(\boldsymbol{X})}\right) dx \tag{2}$$

subject to matching histogram probabilities. Note that relative entropy is non-negative, $D\left(q(\boldsymbol{X}) \,||\, p(\boldsymbol{X})\right) \geq 0$, and equal to zero if and only if $p(\boldsymbol{X}) = q(\boldsymbol{X})$.[9]

3.2 Tilting to a single histogram

In the case of tilting to a single histogram with J bins, $\{A_j\}_{j=1}^J$, the entropic tilting problem amounts to finding a density $q(\boldsymbol{X})$ that matches a given set of marginal distributions, $\{q\left(A_j\right)\}_{j=1}^J$, and otherwise minimizes the relative entropy $D_x(q \,||\, p)$. The densities $q(\boldsymbol{X})$ and $p(\boldsymbol{X})$ can generally be factorized as

$$q(\boldsymbol{X}) = q\left(A_j\right) q\left(\boldsymbol{X}|A_j\right), \; p(\boldsymbol{X}) = p\left(A_j\right) p\left(\boldsymbol{X}|A_j\right), \; \forall \, j, \boldsymbol{X} \in A_j \,, \tag{3}$$

leading to a corresponding decomposition of relative entropy:

$$D\left(q(\boldsymbol{X}) \,||\, p(\boldsymbol{X})\right) = D\left(q\left(A\right) \,||\, p\left(A\right)\right) + \sum_j q\left(A_j\right) \cdot \left\{D\left(q\left(\boldsymbol{X}|A_j\right) \,||\, p\left(\boldsymbol{X}|A_j\right)\right)\right\},$$

[9] Notation for relative entropy has been adopted from Cover and Thomas (2006).

where

$$D\left(q\left(A\right)\ ||\, p\left(A\right)\right) = \sum_{j} q\left(A_j\right) \cdot \log\left(\frac{q\left(A_j\right)}{p\left(A_j\right)}\right)$$
$$D\left(q\left(\boldsymbol{X}|A_j\right)\ ||\, p\left(\boldsymbol{X}|A_j\right)\right) = \int_{x \in A_j} q\left(\boldsymbol{X}|A_j\right) \cdot \log\left(\frac{q\left(\boldsymbol{X}|A_j\right)}{p\left(\boldsymbol{X}|A_j\right)}\right) dx.$$

Notice that $D\left(q\left(A\right)\ ||\, p\left(A\right)\right)$ is solely determined by the bin probabilities under the baseline density p and the targeted density q. This term is fixed by the choice of bins and proposal density, and it represents the relative entropy between the targeted histogram and its counterpart from the baseline. Crucially, as argued next, the relative entropy between histograms from baseline and the SPF targets is also the *minimized* relative entropy between the baseline density $p(\boldsymbol{X})$ and its tilted counterpart $q(\boldsymbol{X})$. To see this, note that we can freely choose the *conditionals* of the tilted density $q\left(\boldsymbol{X}|A\right)$, to obtain zero relative entropy within each bin. Thus, we can achieve the optimum by simply leaving the densities *inside each bin* unchanged relative to the baseline:

$$D\left(q\left(\boldsymbol{X}|A_j\right)\ ||\, p\left(\boldsymbol{X}|A_j\right)\right) = 0 \quad \Leftrightarrow \quad q\left(\boldsymbol{X}|A_j\right) = p\left(\boldsymbol{X}|A_j\right) \quad \forall\ j. \tag{4}$$

Because relative entropy is nonnegative, this is the minimizing solution to the tilting problem. Intuitively speaking, the minimizing solution reweights draws of $\boldsymbol{X}$ that fall into the same bin A_j with identical weight $q\left(A_j\right)/p\left(A_j\right)$, and minimizes relative entropy by preserving the shape of the original distribution f inside each bin.

To make the implementation concrete in the single histogram case, suppose we have a simple histogram with just two bins, for inflation next year, with one bin for inflation to exceed 2 percent and the other bin for inflation to be 2 percent or less. Assume that, for the particular year ahead, survey respondents attach equal probabilities of 50 percent to these two bins. Now suppose that, in the baseline forecast density from the BVAR that is represented by Markov Chain Monte Carlo draws from the posterior

predictive distribution, the baseline density has 60 percent of forecast draws that exceed 2 percent and 40 percent of draws of inflation of 2 percent. In this case, our entropic tilting solution takes all draws of inflation that are 2 percent or less and reweights (or overweights) them by a factor of 5/4 and takes all draws of inflation exceeding 2 percent and reweights by them by a factor of 5/6. With this solution, the tilted weights integrate to 1, $(5/4 \times 0.4) + (5/6 \times 0.6) = 0.5 + 0.5 = 1$, while satisfying the targeted histogram probabilities of 50 percent in each bin. Moreover, within each bin, the relative weights of the original draws are preserved, so that the shape of the baseline density within each bin is unchanged.

3.3 Analytical solution to tilting with two or more histograms

The multiple-histogram case is a bit more involved than the one-histogram case, because the targeted histograms restrict multiple marginal densities. In practical terms, the histograms tell us the probability of $\boldsymbol{X}$ falling into either set of bins, but nothing about the interdependence between $\boldsymbol{X}$ falling into bins from either histogram.

In the analytics for multiple histograms, we index the histograms by $n = 1, 2, \ldots, N$ and N refers to the number of histograms. Each histogram contains $J(n)$ bins that represent a set of mutually exclusive events. Probabilities of falling into the jth bin of the nth histogram under proposal and targeted histogram are denoted by $p\left(A_j^n\right)$ and $q\left(A_j^n\right)$, respectively.

3.4 The case of two histograms

Considering the case of two histograms, we denote the probabilities pertaining to joint occurrences in bins i and j of the two histograms as follows:[10]

$$p\left(A_i^1, A_j^2\right) \equiv p\left(\boldsymbol{X} \in A_i^1 \cap \boldsymbol{X} \in A_j^2\right) , \tag{5}$$

[10]The definition applies to every $0 < i \leq J(1)$ and $0 < j \leq J(2)$.

The SPF prescribes targeted values for, $q\left(A_i^1\right)$ and $q\left(A_j^2\right)$, instead of the joint density $q\left(A_i^1, A_j^2\right)$. To decompose the the relative entropy objective, we factorize the densities:

$$\begin{aligned} q(\boldsymbol{X}) &= q\left(A_i^1, A_j^2\right)\ q\left(\boldsymbol{X}|A_i^1, A_j^2\right) \\ &= q\left(A_i^1\right)\ q\left(A_j^2\right)\ m^q\left(A_i^1, A_j^2\right)\ q\left(\boldsymbol{X}|A_i^1, A_j^2\right)\ , \qquad (6) \\ p(\boldsymbol{X}) &= p\left(A_i^1, A_j^2\right)\ p\left(\boldsymbol{X}|A_i^1, A_j^2\right) \\ &= p\left(A_i^1\right)\ p\left(A_j^2\right)\ m^p\left(A_i^1, A_j^2\right)\ p\left(\boldsymbol{X}|A_i^1, A_j^2\right)\ , \qquad (7) \end{aligned}$$

where we have introduced the following "interaction coefficients," that capture the dependence structure between the two histograms.

$$m^q\left(A_i^1, A_j^2\right) \equiv \frac{q\left(A_i^1, A_j^2\right)}{q\left(A_i^1\right)\ q\left(A_j^2\right)}\ , \qquad m^p\left(A_i^1, A_j^2\right) \equiv \frac{p\left(A_i^1, A_j^2\right)}{p\left(A_i^1\right)\ p\left(A_j^2\right)}\ , \qquad (8)$$

These interaction coefficients are measures of association and capture deviations from independence. Formally, the interaction coefficients are also Radon-Nikodym derivatives that describe the change of measure between the product of the marginal distributions $q\left(A_i^1\right)$ and $q\left(A_j^2\right)$ and the joint distribution $q\left(A_i^1, A_j^2\right)$, and the same for $p(\cdot)$.[11] A positive association, $m^q\left(A_i^1, A_j^2\right) > 1$, reflects the case when outcomes tend to occur together (and vice versa for negative associations). From the factorizations in (6) and (7), we obtain a linear decomposition of relative entropy:

$$\begin{aligned} D\left(q(\boldsymbol{X})\,||\,p(\boldsymbol{X})\right) = D\left(q\left(A^1\right)\ ||\,p\left(A^1\right)\right) + D\left(q\left(A^2\right)\ ||\,p\left(A^2\right)\right) \\ + \Delta_q\left(m^q\left(A^1, A^2\right)\ ||\,m^p\left(A^1, A^2\right)\right) \\ + E_q\left\{D\left(q\left(\boldsymbol{X}|A^1, A^2\right)\ ||\,p\left(\boldsymbol{X}|A^1, A^2\right)\right)\right\}\ , \qquad (9) \end{aligned}$$

[11] The product of the marginal distributions characterizes a (hypothetical) joint distribution that would arise if events in both histograms were independent. Note that $m^q\left(A_i^1, A_j^2\right) = 1$ if and only if events in both histograms are independent under $q(\cdot)$.

with $\quad \Delta_q \left(m^q \left(A^1, A^2 \right) \, || \, m^p \left(A_i^1, A_j^2 \right) \right) \equiv$

$$\sum_i \sum_j q \left(A_i^1 \right) \, q \left(A_j^2 \right) \, m^q \left(A_i^1, A_j^2 \right) \, \log \left(\frac{m^q \left(A_i^1, A_j^2 \right)}{m^p \left(A_i^1, A_j^2 \right)} \right), \quad (10)$$

and $\quad E_q \left\{ D \left(q \left(\boldsymbol{X} | A^1, A^2 \right) \, || \, p \left(\boldsymbol{X} | A^1, A^2 \right) \right) \right\} \equiv$

$$\sum_i \sum_j q \left(A_i^1, A_j^2 \right) \cdot D \left(q \left(\boldsymbol{X} | A^1, A^2 \right) \, || \, p \left(\boldsymbol{X} | A^1, A^2 \right) \right) . \quad (11)$$

As in the one-histogram case, the first term in (9) is fixed by the choice of bins and the targeted histogram probabilities; we refer to this term as "marginal entropy," since it captures the relative entropy between the marginal histogram densities. The last term in (9) captures an average of the relative entropies between the conditional densities inside each of the joint cross-histogram events, as defined in (11), which are not directly restricted by the histograms. As before, the within-cross-event relative entropies, $\forall \, i, j$, can be set to zero by leaving the conditional densities unchanged relative to the baseline:[12]

$$q \left(\boldsymbol{X} | A^1, A^2 \right) = p \left(\boldsymbol{X} | A^1, A^2 \right) \iff D \left(q \left(\boldsymbol{X} | A^1, A^2 \right) \, || \, p \left(\boldsymbol{X} | A^1, A^2 \right) \right) = 0, \quad (12)$$

This leaves the second term in (9), $\Delta_q \left(m^q \left(A^1, A^2 \right) \, || \, m^p \left(A^1, A^2 \right) \right)$, as the only remaining component of relative entropy that can be influenced by choice of $q(\boldsymbol{X})$. We refer to $\Delta_q \left(\cdot || \cdot \right)$ as "interaction entropy," and it measures the entropic cost of changing the interaction (or association) structure between the two histograms. We thus need to choose the joint bin probabilities $q \left(A_i^1, A_j^2 \right)$, or equivalently the interaction coefficients $m^q \left(A_i^1, A_j^2 \right)$, to minimize interaction entropy while satisfying the constraints imposed by

[12] Applying the usual definition for relative entropy to the within-cross-event case, we have:

$$D \left(q \left(\boldsymbol{X} | A_i^1, A_j^2 \right) \, || \, p \left(\boldsymbol{X} | A_i^1, A_j^2 \right) \right) = \int_X q \left(\boldsymbol{X} | A_i^1, A_j^2 \right) \, \log \left(\frac{q \left(\boldsymbol{X} | A_i^1, A_j^2 \right)}{p \left(\boldsymbol{X} | A_i^1, A_j^2 \right)} \right) dx \, ,$$

where the integral is taken over $\boldsymbol{X} \in A_i^1 \cap A_j^2$.

the targeted marginal histogram probabilities:[13]

$$\min_{m^q(A^1,A^2)} \Delta_q \left(m^q \left(A^1, A^2 \right) \, || \, m^p \left(A^1, A^2 \right) \right) \tag{13}$$

$$\text{subject to} \quad \sum_j q \left(A_i^1 \right) \, m^q \left(A_i^1, A_j^2 \right) = \sum_i q \left(A_i^1 | A_j^2 \right) = 1 \, , \qquad \forall \, j \, , \tag{14}$$

$$\text{and} \quad \sum_i q \left(A_j^2 \right) \, m^q \left(A_i^1, A_j^2 \right) = \sum_j q \left(A_j^2 | A_i^1 \right) = 1 \, , \qquad \forall \, i \, , \tag{15}$$

The first order conditions for minimizing relative entropy imply the following:

$$m^q \left(A_i^1, A_j^2 \right) \quad = \quad m^p \left(A_i^1, A_j^2 \right) \quad \cdot \phi_i^1 \cdot \phi_j^2 \qquad \forall i, j \tag{16}$$

$$\Leftrightarrow \quad \frac{q \left(A_i^1, A_j^2 \right)}{p \left(A_i^1, A_j^2 \right)} \quad = \quad \frac{q \left(A_i^1 \right)}{p \left(A_i^1 \right)} \cdot \frac{q \left(A_j^2 \right)}{p \left(A_j^2 \right)} \cdot \phi_i^1 \cdot \phi_j^2 \, , \tag{17}$$

where ϕ_i^1 and ϕ_j^2 are (elements of vectors of) Lagrange multipliers attached to events from each histogram that need to be set such as to assure that (14) and (15) hold. Notice that $J(1) \times J(2)$ values for $q \left(A_i^1, A_j^2 \right)$ are pinned down by $p \left(A_i^1, A_j^2 \right)$ and the $J(1) + J(2)$ Lagrange multipliers ϕ_i^1 and ϕ_j^2. At the optimum, the Lagrange multipliers satisfy the following set of conditions:

$$\phi_i^1 \cdot \left(\sum_j q \left(A_j^2 \right) \cdot m^p \left(A_i^1, A_j^2 \right) \cdot \phi_j^2 \right) = 1 \, , \qquad \forall \, i \tag{18}$$

$$\text{and} \quad \phi_j^2 \cdot \left(\sum_i q \left(A_i^1 \right) \cdot m^p \left(A_i^1, A_j^2 \right) \cdot \phi_i^1 \right) = 1 \, , \qquad \forall \, j \tag{19}$$

Without convenient closed-form solutions for ϕ_i^1 and ϕ_j^2 available, we obtain the solution by iterating over (18) and (19).[14] As discussed in Section 2, our procedure is an

[13] Note that, as a consequence of (14) and (15), we have $\sum_j q \left(A_j^2 | A_i^1 \right) q \left(A_i^1 \right) = q \left(A_j^2 \right)$ and $\sum_i q \left(A_i^1 | A_j^2 \right) q \left(A_j^2 \right) = q \left(A_i^1 \right)$. Moreover, $\sum_j \sum_i q \left(A_i^1, A_j^2 \right) = 1$ also holds, since the targeted (marginal) histogram probabilities sum to one.

[14] Formally, index each step of the iteration by s and for $s = 0$ set initial values $\phi_i^1\{0\} = 1$ and $\phi_j^2\{0\} = 1$. Then iterate over the following: First, given $\phi_j^2\{s-1\} \; \forall \; i$, for each event i in A^1, set $\phi_i^1\{s\}$ to satisfy $\phi_i^1\{s\} \cdot \left(\sum_j q \left(A_j^2 \right) \, m^p \left(A_i^1, A_j^2 \right) \, \phi_j^2\{s-1\} \right) = 1, \; \forall \; j$. Second, given $\phi_i^1\{s\} \; \forall \; i$, for each event i in A^2, set $\phi_j^2\{s\}$ to satisfy $\phi_j^2\{s\} \cdot \left(\sum_i q \left(A_i^1 \right) \, m^p \left(A_i^1, A_j^2 \right) \, \phi_i^1\{s\} \right) = 1, \; \forall \; j$. until convergence at a solution

instance of "iterative proportional fitting" (IPF) which has a longstanding tradition in the statistics literature.

Remark: When the original density, $p\left(A_i^1, A_j^2\right)$, is identical to the product density of its marginals, and the interaction coefficient equals one:

$$p\left(A_i^1, A_j^2\right) = p\left(A_i^1\right) \cdot p\left(A_j^2\right) \Leftrightarrow m^p\left(A_i^1, A_j^2\right) = 1\,, \quad \forall\, j, i$$

In this case, the entropy-minimizing Lagrange multipliers ϕ_i^1 and ϕ_j^2 are all identical to one. [15] When, under the baseline, the histogram events are mutually independent, jointly tilting to all histograms is identical to sequentially targeting the individual histograms.

3.5 The general case of two and more histograms

Our approach easily generalizes to case of three and more histograms. As in the two-histogram case, relative entropy can be decomposed into (a) the sum of relative entropies for the marginal histograms, (b) relative entropy of the cross-product densities ("interaction entropy"), and (c) the relative entropy of the within-cross-event densities. As before, the marginal entropy is fixed by the problem, and the within-cross-event entropies can always be set to zero, so that only the minimization of interaction entropy remains.

With $N = 3$, the optimal solution is characterized by the following:

$$\frac{q\left(A_i^1, A_j^2, A_k^3\right)}{p\left(A_i^1, A_j^2, A_k^3\right)} \quad = \quad \frac{q\left(A_i^1\right)}{p\left(A_i^1\right)} \cdot \frac{q\left(A_j^2\right)}{p\left(A_j^2\right)} \cdot \frac{q\left(A_k^3\right)}{p\left(A_k^3\right)} \cdot \phi_i^1 \cdot \phi_j^2 \cdot \phi_k^3\,, \tag{20}$$

where ϕ_i^1, ϕ_j^2, ϕ_k^3 are (elements of vectors of) Lagrange multipliers attached to bin events from each histogram that need to be set such as to assure that probabilities sum to one,

that satisfies the constraints (14) and (15).

[15] To verify this, note that with $m^p\left(A_i^1, A_j^2\right) = 1$, the interaction entropy is equal to the relative entropy between $q\left(A^1, A^2\right)$ and the independence distribution $q\left(A^1\right) \cdot q\left(A^2\right)$: $\Delta_q\left(m^q\left(A^1, A^2\right) \,||\, 1\right) = D\left(\, q\left(A^1, A^2\right) \,||\, q\left(A^1\right) \cdot q\left(A^2\right)\right)$, which in turn is minimized (to zero) by setting the interaction coefficient for $q(\cdot)$ to unity, $m^p\left(A_i^1, A_j^2\right) = 1$.

analogously to (14) and (15) in the case of $N = 2$.

At the optimum, the Lagrange multipliers satisfy the following set of conditions:

$$\phi_i^1 \cdot \left(\sum_i \sum_k q\left(A_j^2\right) \cdot q\left(A_k^3\right) \cdot m^p\left(A_i^1, A_j^2, A_k^3\right) \cdot \phi_j^2 \cdot \phi_k^3 \right) = 1\,, \qquad \forall\, i\,, \tag{21}$$

$$\phi_j^2 \cdot \left(\sum_j \sum_k q\left(A_i^1\right) \cdot q\left(A_k^3\right) \cdot m^p\left(A_i^1, A_j^2, A_k^3\right) \cdot \phi_i^1 \cdot \phi_k^3 \right) = 1\,, \qquad \forall\, j\,, \tag{22}$$

$$\text{and} \quad \phi_k^3 \cdot \left(\sum_j \sum_i q\left(A_i^1\right) \cdot q\left(A_j^2\right) \cdot m^p\left(A_i^1, A_j^2, A_k^3\right) \cdot \phi_i^1 \cdot \phi_j^2 \right) = 1\,, \qquad \forall\, k\,, \tag{23}$$

which are a direct extension of the case of two-histogram case and straightforward to generalize to case of arbitrary N. Starting from an initialization that sets all Lagrange multipliers to one, we solve for the solution by iterating over (21)-(23).

3.6 Computational implementation

This section details our computational implementation of the tilting solution in the occasional cases in which the (simulated) predictive density from the model does not have mass in some of the target bins from SPF histogram forecasts.[16] In the context of our application to histograms, the sufficient condition for a solution to the tilting problem is to have

$$p(A_i^1, A_j^2, \ldots) \equiv p(A_i^1 \cap A_j^2 \cap A_k^3 \cap \ldots) > 0\,, \qquad \forall\, j, i, k\,. \tag{24}$$

However, the proposals simulated from our BVAR may not always satisfy this condition.[17] To address this computational issue we construct an augmented proposal density that is a mixture between the original density and an equally weighted density over all joint

[16] For our purpose, where the original density is represented as draws from a Monte Carlo simulation, the terms "generate draws" and "has mass" are understood to be synonyms.

[17] In principle, since the model's forecast densities are unbounded, there is always non-zero mass on any event. However, for practical purposes, the probability of landing in a given joint cross-histogram event can be very small, leading to zero draws in a given simulation of arbitrary (but finite) length. As a point of reference, consider an application, similar to ours, with $N = 8$ histograms having $J = 10$ bins each, leading to 100,000,000 events.

cross-histogram event events (leaving the within-joint-event density unspecified), placing only minimal weight on the latter:

$$\hat{p}(A_i^1, A_j^2, \ldots) = \big(1-\omega\big)\, p(A_i^1, A_j^2, \ldots) + \omega/(J^N) \tag{25}$$

where $J(n)$ denotes the number of bins in the nth histogram, so that $\prod_n J(n)$ denotes the total number of joint cross-histogram events. (In our application, we set $\omega = 1e^{-6}$.) The idea of adding uninformative "noise" to the proposal is related (though distinct) to the use of tempering by Montes-Galdon, Paredes, and Wolf (2022) in estimating the underlying model density while ensuring coverage of events that the survey places mass on.

The augmented proposal describe in (25), leaves the within-joint-event density, $\hat{p}(X|A_i^1, A_j^2, \ldots)$ unspecified. For the results reported below, we derive tilting weights $q(A_i^1, A_j^2, \ldots)/\hat{p}(A_i^1, A_j^2, \ldots)$ based on the augmented proposal, but focus the evaluation only on (reweighted) draws from the original BVAR proposal. This procedure amounts to conditioning the tilted density on outcomes for which the BVAR has generated draws with positive probability. Thus, the results reported below reflect $q(X|X \in \mathcal{P})$ where $\mathcal{P}$ is the set of outcomes that fall into joint cross-histogram events for which the BVAR simulation has generated draws.[18]

4 Data and model specification

This section details the SPF forecasts used in entropic tilting, the real-time data used for model estimation and forecasting, and the forecasting model.

4.1 Survey data from the SPF

Our analysis uses forecasts of real GDP growth, the unemployment rate, and core PCE inflation from the U.S. Survey of Professional Forecasters (SPF). The SPF collects forecasts

[18]Formally, let $\mathcal{P} \equiv \{X|X \in (A_i^1 \cap A_j^2 \cap \ldots)$ with $p(X \in A_i^1 \cap A_j^2 \cap \ldots) > 0$, for some $i, j \ldots\}$.

quarterly, and they are publicly available on the Federal Reserve Bank of Philadelphia's website. In particular, we consider the SPF's density forecasts, which are collected in the form of probability histograms for fixed-event, calendar-year forecasts, for (variously) the current year up through three years ahead. These forecasts refer to percent changes in annual average levels of GDP; the annual average level of the unemployment rate; and the Q4/Q4 percent change in the core PCE price index. Figures 1-3 show the histogram forecasts published for 2025Q3. The histograms are constructed from individual forecaster responses, which are aggregated into bins. While our BVAR specification and forecast evaluation will include the 3-month Treasury bill rate, the SPF does not collect or publish density forecasts of any interest rates.

[Figure 1 about here.]

The availability and specifics of histogram forecasts have varied over time. Since 1981, the SPF has always included current- and next-year density forecasts of GDP growth. In 2007Q1, the survey added current- and year-ahead histogram forecasts of core PCE inflation, and has maintained this coverage through the present. In 2009Q2, the survey extended the GDP horizons to include 2- and 3-year-ahead histograms and added histogram forecasts for the unemployment rate for the current year through 3 years ahead. Reflecting the availability of real-time data on core PCE inflation (more below), the earliest histogram forecasts we use are for 1996Q1.[19]

[Figure 2 about here.]

[Figure 3 about here.]

[Figure 4 about here.]

The ranges underlying the histogram forecasts (in most cases) have also changed over time, leading to variation in the bins and their widths. Figure 4 shows time series of the

[19]The Federal Reserve Bank of Philadelphia took over management of the SPF in 1992.

bins for the histograms of real GDP growth, unemployment rate, and core PCE inflation forecasts used by the SPF. For example, the SPF lowered the histogram ranges for the unemployment a few years after the Great Recession ended, widened them sharply with the COVID-19 pandemic, and narrowed the unemployment ranges a few years later.

[Figure 5 about here.]

Throughout, we collect probabilistic forecast for the average SPF forecaster ("consensus histogram"), which is consistent with treating the ensemble of individual forecasts as an equally weighted linear prediction pool. Equally weighted prediction pools have been shown to often lead to good forecast performance (for example, see Clark and Mertens (2024) for a recent survey).[20] Moreover, as argued next, the SPF consensus histograms display good coherence with SPF point forecasts. By contrast, Engelberg, Manski, and Williams (2009) and Clements (2025, 2010) have documented notable inconsistencies between point and histogram predictions of individual SPF participants. For the case of GDP growth forecasts, Figure 5 compares SPF point forecasts against potential mean and median values consistent with the SPF histograms. Following Engelberg, Manski, and Williams (2009) and Clements (2010), a range of SPF-consistent means is computed while taking the lower and upper bounds of each bin as the location of the probability mass within that bin. As can be seen, the SPF point forecasts lie within or very close to the range of SPF-consistent means for all forecast horizons and for most of the times.[21]

For sake of comparison, we also compute SPF-consistent medians from the midpoints of the bins containing the median, which is found by accumulating bin probabilities until reaching 50%. (Of course, the true medians are not known and could lie within the range of upper and lower limits of the respective median bins which we omit for sake of brevity). Not surprisingly, the SPF-consistent medians are less sensitive to quarterly changes in

[20] Studies such as Chernis, et al. (2025) and Diebold, Shin, and Zhang (2023) (and other references therein) have developed alternative pooling schemes for individual SPF histogram forecasts.

[21] As shown in Figures 11 and 12 in the appendix, similar results apply also to the cases of of SPF histograms for core inflation and the unemployment rate.

the outlook, in particular at longer forecast horizons, which reflects inertia in the identity of the median bin, as well as the infrequent nature of changes in bin limits. That said, SPF-consistent medians broadly track the point-forecast trajectories and often lie within the range of histogram-consistent mean forecasts.[22]

Overall, this coherence between point and histogram forecasts for the consensus SPF histograms suggests that tilting to histogram information will not produce results that are wildly at odds with point forecast information. Of course, point forecast information from the SPF has a well-established track record of predictive value. For the remainder of this paper, we will ignore information on point forecasts, which also allows us to sidestep the question as to whether those should be interpreted as means or modal forecasts or other metrics. Instead, we solely focus on the predictive value contained in histograms.

4.2 Real-time data for model estimation and forecast evaluation

In keeping with common macroeconomic forecasting, we obtain baseline forecasts from models of major macroeconomic aggregates, at the quarterly frequency: real GDP growth, the unemployment rate, core PCE inflation, and the 3-month Treasury bill rate.[23] Real-time data vintages for GDP, the unemployment rate, and the core PCE price index are taken from the Federal Reserve Bank of Philadelphia's Real-Time Data Set for Macroeconomists (RTDSM). We obtained the T-bill rate from the FRED database of the Federal Reserve Bank of St. Louis.

Our analysis of out-of-sample forecasts uses real-time data vintages from 1996Q1 (the first for which core PCE data are available) through 2025Q3. As described in Croushore and Stark (2001), the vintages of the RTDSM are dated to reflect the information available around the middle of each quarter. For each forecast origin t starting with 1996:Q1, we use

[22]Nevertheless, the SPF-consistent median values for GDP growth typically run below the actual point forecasts, suggestive of skew in the growth distribution (when point forecasts are interpreted as means).

[23]Quarterly levels of the unemployment and T-bill rates are monthly averages. GDP growth and PCE inflation are constructed as 400 times log differences of the levels of GDP and the price index.

the real-time data vintage t containing data through $t-1$ to estimate the forecast models and construct forecasts for periods t and beyond. To evaluate the accuracy of the real-time forecasts, we use the second available (in the quarterly vintages of the RTDSM) estimates of the real-time measured variables as actuals in evaluating forecast accuracy. Our models are estimated using data starting in 1959.

4.3 BVAR-SVO-t model

We produce baseline macroeconomic forecasts using the BVAR-SVO-t model developed in Carriero, et al. (2024), but modified so that stochastic volatility follows an AR(1) rather than random walk process. The BVAR takes the form::

$$y_t = \Pi_0 + \Pi(L)y_{t-1} + v_t\,, \qquad\qquad v_t \sim N(0, \Sigma_t), \tag{26}$$

where y_t is a vector of N observables, $\Pi(L) = \sum_{i=1}^{p} \Pi_i \; L^{i-1}$ is a pth-order lag polynomial of VAR coefficients, and v_t denotes the VAR's residuals. We denote the vector of stacked coefficients contained in $\{\Pi_i\}_{i=0}^{p}$ as Π. All models are specified with non-conjugate priors for Π and Σ_t. In the results presented below, the innovation variance matrix Σ_t varies over time. The online appendix shows that a homoskedastic (constant Σ) BVAR yields very similar impacts of tilting to SPF histograms.

The innovation vector in the model consists of a vector of Gaussian innovations scaled by three time-varying variance components, for: (i) stochastic volatility; (ii) rare, transitory, and large — i.e., outlier — changes in volatility; and (iii) transitory changes in volatility that are more frequent but less extreme in impact (consistent with draws from the tails of a fat-tailed distribution). With this stochastic structure, the VAR's innovation v_t can be seen as mixed Gaussian instead of Gaussian. More specifically, the innovation vector and its

covariance matrix take the forms:

$$v_t = A^{-1}\,\Lambda_t^{0.5}\,O_t\,Q_t\,\varepsilon_t,\;\; \varepsilon_t \sim N(0,I),\;\; \Sigma_t = A^{-1}\,O_t\,Q_t\Lambda_t\,Q_t'\,O_t'(A^{-1})'.$$

A^{-1} is a unit-lower-triangular matrix. $\Lambda_t^{0.5}$ is a diagonal matrix of stochastic volatilities; the elements of the vector of logs of the diagonal elements of Λ_t, denoted $\log\lambda_t$, evolve as AR(1) processes with correlated errors:

$$\log\lambda_t = \Psi_0 + \Psi_1\log\lambda_{t-1} + e_t \quad \text{with} \quad e_t \sim N(0,\Phi),$$

where Ψ_0 is a vector of intercepts and Ψ_1 is a diagonal matrix of AR(1) coefficients.

The outliers enter the model in a diagonal matrix of scale factors, denoted O_t, with diagonal elements $o_{j,t}$ that are mutually $i.i.d.$ over all j and t. The outlier $o_{j,t}$ has a two-part distribution that distinguishes regular observations with $o_{j,t} = 1$ from outliers for which $o_{j,t} \geq 2$.[24] Outliers in variable j, $j = 1,\ldots,N$, occur with probability p_j and the distribution:

$$o_{j,t} = \begin{cases} 1 & \text{with probability} \quad 1-p_j \\ U(2,20) & \text{with probability} \quad p_j, \end{cases}$$

where $U(2,20)$ denotes a uniform distribution with support between 2 and 20. Stock and Watson (2016) first used this outlier specification with stochastic volatility in a multivariate unobserved component model of inflation.

Finally, the frequent-change volatility states consistent with fat-tailed innovations enter the model in the diagonal matrix Q_t with elements $q_{j,t}$ that are mutually $i.i.d.$ over all j and t. This component of the model is equivalent to having t-distributed VAR residuals (conditional on Λ_t and O_t). Following Jacquier, Polson, and Rossi (2004), we let the

[24]The lower bound of 2 on the scale shift in outliers is motivated by seeing outliers as events firmly outside the typical mass of their otherwise Gaussian distribution (conditional on $o_{j,t}$).

squares of the diagonal elements of Q_t, $q_{j,t}$, have inverse-gamma distributions:

$$q_{j,t}^2 \sim IG\left(\frac{d_j}{2}, \frac{d_j}{2}\right).$$

The j^{th} residual $q_{j,t} \cdot \varepsilon_{j,t}$ (adjusted for the rotation by A^{-1} and scaling by $\Lambda_t^{0.5} O_t$) has a student-t distribution with d_j degrees of freedom, since $\varepsilon_{j,t} \sim N(0,1)$ and $d_j/q_{j,t} \sim \chi^2_{d_j}$.

4.4 Model estimation

Our models are estimated with a Gibbs sampler, based on the methods of Carriero, Clark, and Marcellino (2019) (henceforth "CCM") for estimating large BVARs, but as corrected in Carriero, et al. (2022). As in CCM, we use a Minnesota prior for the VAR coefficients Π and follow their other choices for priors as far as applicable, too. Throughout, our BVARs include $p = 4$ lags.

For the infrequent outlier component of the BVAR-SVO-t model, we follow Stock and Watson (2016) in placing a beta prior on the outlier probability p_j. The prior is set to imply a mean outlier frequency of once every 10 years in quarterly data for SVO-t estimates, with precision set to be consistent with 10 years' worth of prior observations. For the t-distributed component of the BVAR-SVO-t model, we follow Jacquier, Polson, and Rossi (2004) and estimate the degrees of freedom d_j for each variable using a uniform discrete prior with a range of 3 to 40. Carriero, et al. (2024) provide details on the Gibbs sampler steps for these and other volatility components of the model.

Parameter estimates from the BVAR-SVO-t model are based on 5,000 retained draws, obtained by sampling a total of 7,5000 draws and discarding the first 2,500. For each retained draw of parameters in the Gibbs sampler, we produce 50 draws of forecasts.[25] Accordingly, all results in the paper are based on a total of 250 thousand draws of forecasts

[25]This approach builds on the ideas of Waggoner and Zha (1999). At every Gibbs sampler node m, we employ J draws of the volatility components and VAR shocks over the forecast horizons, to better balance computational costs against a high degree of accuracy in the Monte Carlo approximation of the VAR's predictive likelihood.

from the BVAR's posterior predictive distribution.

5 Effect of SPF-tilting on model forecasts

This section evaluates the effects of entropic tilting when applied to predictive densities for real GDP growth, the unemployment rate, core PCE inflation, and the 3-month T-bill rate, comparing the predictive densities of the BVAR-SVO-t model against histogram forecasts reported by the SPF. As described in Section 4.1, as available we include SPF histograms for GDP growth, unemployment, and inflation, for horizons from the current year through three years ahead. While we do not tilt to survey forecast information for the T-bill rate, tilting to forecasts of other variables will impact the model's forecasts of the interest rate.

Entropic tilting is implemented using the analytical solution detailed in the preceding section. Of course, a single given histogram (one variable at one horizon) has multiple bins to hit. Histograms for additional variables or horizons adds to the bins to hit. The timing and horizons of BVAR forecasts are aligned with SPF survey rounds.

Our tilting implementation appropriately bridges between the underlying quarterly forecasts from the BVAR and the annual forecasts of the SPF histograms. Specifically, starting from the quarterly draws, we perform the necessary transformations to obtain BVAR draws of calendar year forecasts as defined by the SPF, to which we apply the tilting targets.[26]

[Figure 6 about here.]

5.1 Visual fit to SPF histograms

We begin by illustrating the differences between raw model output and SPF histogram forecasts for selected forecast origins, of 2009Q1 and 2022Q2. In both instances, reflecting

[26] Reflecting the SPF's variable-specific convention, these transformation can be linear (e.g., taking annual averages of quarterly forecasts of the unemployment rate) or nonlinear (e.g. calculating growth rates of annual average levels of GDP).

somewhat unusual economic circumstances associated with the Great Recession (2009Q1) and the recovery following the COVID-19 pandemic (2022Q2), the SPF forecasts differ from the BVAR forecasts, such that entropic tilting to the SPF histograms has an impact on the model's forecasts. For brevity, we defer to Section 5.2 the details of the choice of SPF histograms used as targets.

Figures 6 and 7 provide predictive densities for annual forecasts corresponding to SPF horizons, from the current year through two years ahead (we omit three years ahead for chart readability). Each panel provides the BVAR forecasts as magenta (untilted) and dotted black (tilted) lines. For variables and horizons for which SPF histograms are available, the relevant panels show the SPF histograms with black bars and untilted BVAR histograms with red bars.

Starting with basic features of the SPF histograms and the BVAR's untilted forecasts, the baseline model, the BVAR-SVO-t specification, produces smooth predictive densities. In contrast, the SPF histograms in some instances exhibit mass concentrated in specific bins, reflecting survey participants' discrete expectations. Most starkly, in 2009Q1, when the economy was still in recession, the SPF predictive density for GDP growth in 2009 was sharply skewed to the left, with substantial probability on GDP growth of -2 percent or less and an 89 percent probability of GDP declining in 2009 (panel 6a). While the BVAR's predictive density shows a significant chance of a decline in GDP for the year, the model had a much more favorable outlook for GDP growth. In this instance, tilting the BVAR's forecasts to the SPF histograms substantially shifts the model's predictive density for GDP growth in 2009 to the left. In general, though, as noted above, tilting the model output to match the SPF histogram redistributes probability mass within bins, preserving the overall shape while aligning with survey responses.

Tilting to the SPF histograms also impacts the model's forecasts for unemployment and core PCE inflation (in 2009Q1, there were not yet published unemployment histograms to incorporate as tilting targets). As expected, with tilting reducing the GDP growth forecast,

it also raises the unemployment rate forecast, shifting the model's predictive density to the right (panel 6b). The available SPF forecast of core inflation tends to reduce the dispersion of the BVAR's forecast for inflation in 2009 (panel 6c). Forecasts for the year-ahead horizon, the year of 2010, show patterns similar to those for the current year forecasts, although with tilting in this instance also tending to shift to the right some the model's forecast for core PCE inflation (panel 6f). The BVAR's forecasts for two years ahead continue to show some impact of tilting to SPF histograms, but smaller impacts than for shorter horizons (panels 6g, 6h, and 6i).

[Figure 7 about here.]

In 2021Q2, with the economy continuing to recover from the COVID-19 pandemic, the SPF predictive density for GDP growth in 2021 is somewhat skewed to the right, as shown in Figure 7. The histogram's mode put a probability of 51 percent on growth between 4 and 6.9 percent and a probability of 24 percent on growth between 7 and 9.9 percent, compared to a probability of 16 percent on growth between 2.5 and 3.9 percent. The SPF's density forecast for the unemployment rate in 2021 is largely symmetric, but fairly dispersed. Entropically tilting the BVAR's forecasts to match these histogram shapes results in predictive densities for GDP growth and unemployment with two peaks (panels 7a and 7b). The same occurs with the BVAR's forecasts of core PCE inflation, in this case also shifting the predictive distribution toward higher inflation (panel 7c). At longer horizons, the impacts of tilting to the SPF histograms are comparable for year-ahead GDP growth (panel 7d) and more modest two years out (panel 7g), and some tilting impacts remain evident for longer-horizon forecasts of unemployment and core PCE inflation (panels 7e, 7f, 7h, and 7i).

[Figure 8 about here.]

As the preceding examples indicate, entropic tilting the BVAR's forecasts to SPF histograms can materially alter the model's forecasts in order to closely follow the SPF

bin probabilities. To provide a broader time perspective on how much such tilting impacts the baseline model forecasts, we examine the time series of the KLIC distance between the untilted and tilted distributions from 1996 into 2025. The KLIC distance provides a measure of how much (at each forecast origin) the conditions provided by the SPF histograms alter the baseline BVAR's predictive density. The model's ability to adapt to these conditions is evident in the tilting process, which effectively redistributes probability mass to align with the survey data.

Figure 8 provides the time series of KLIC distances. This history includes some periods characterized by significant uncertainty, particularly during the GFC and the COVID-19 pandemic. Tilting the model outputs to the SPF around the aforementioned periods of greater uncertainty also leads to more notable differences in relative entropy, as measured by the KLIC, between the model and the SPF histogram forecasts as illustrated in Figure 8. The figure shows the total KLIC (black line) and contributions (colored bars) from tilting to the selected set of targets, for unemployment (by each year of available SPF forecast), inflation (also by year), and current-year GDP growth. The total KLIC is typically lower than the sum of KLIC's obtained from separately tilting to individual targets, since the information conveyed by different targets tends to overlap.

As shown in the figure, the information content conveyed by the SPF (as measured by the KLIC) has varied over time. In the period of (relatively speaking, by historical standards) macroeconomic stability from 1996 through 2006 — a period in which the only SPF histograms available for tilting were current- and next-year GDP growth — the KLIC distance between the model's raw forecast and the tilted forecast density was relatively small and stable, with varying (over time) contributions from both the current- and next-year GDP forecasts. Subsequently, the information content conveyed by the SPF (as measured by the KLIC) rose particularly around the GFC and its aftermath for a few years and then again during the COVID-19 pandemic. Around the GFC, the KLIC contribution bars are largest for the SPF's unemployment rate forecasts from one through

three years ahead. The current-year forecasts of unemployment, the current- and next-year forecasts of GDP growth, and the available core PCE forecasts make some contribution, but by less so than the unemployment forecasts one or more years ahead. Together, around the GFC, the multi-year unemployment rate histograms from the SPF seem to provide the most information content for the BVAR forecasts. The shorter period of COVID-19 volatility around 2020 shows broadly similar patterns, but with a larger contribution from the current year unemployment rate forecast.

5.2 Forecast accuracy evaluation

We assess the performance of the tilted and untilted forecasts using real-time forecast evaluations, of forecasts made starting in 1996Q1 and continuing into 2025. As indicated above, drawing on the histogram forecasts available in the SPF, our tilting targets include GDP growth, unemployment, and core PCE inflation. For the unemployment rate, we use all of the SPF horizons available, from the current year through three years ahead, starting with the 2009Q2 forecast origin. For core PCE inflation, we use the current and next year forecast, starting with 2007Q1. In the case of GDP growth, we use only the current year and next-year forecast histograms in the entropic tilting, starting with 1996Q1. Our rationale for only using the current year and next-year GDP forecasts reflects two considerations. First, at longer horizons, the SPF's view of GDP is likely connected to its view of unemployment, such that the unemployment outlook by itself has information for the GDP forecast. Second, it is well known that GDP growth has only low-order serial correlation, and that predictability declines quickly with a few quarters (see, e.g., Breitung and Knüppel (2021)). Of course, one could consider tilting the BVAR forecasts to other, more limited combinations of the available SPF histogram forecasts. We view our chosen, larger set of targets to have the advantage of being comprehensive in capturing the SPF's view of the macroeconomic outlook and providing a simple and broad look at the potential efficacy of information BVAR forecasts with survey-based density forecasts in the form of

the SPF histograms.

5.3 Forecast accuracy of annual histogram predictions

We begin by gauging the potential for improved forecast accuracy from tilting to the SPF by comparing the SPF histogram forecasts to the BVAR's predictive densities for corresponding annual forecasts. That means we compare the quality of predictions for outcomes to fall into a given SPF-defined bin as predicted by the model (without tilting) and the SPF histogram forecasts. As summary measure, we computed discrete rank probability score (DRPS) for both the BVAR-SVO-t model and the SPF histogram forecasts. The DRPS is defined as $\text{DRPS}_t = \sum_{k=1}^{K} \left(P_t^k - D_t^k\right)^2$, with K indicating the number of bins, P_t^k denoting the cumulative forecast probability from bins 1 through k, and D_t^k is the cumulation of an indicator variable with value for k of 1 if the outcome falls in the bin and value of 0 otherwise. The lower the score, the better the forecast. Past studies of survey forecasts using the DRPS include Boero, Smith, and Wallis (2011), Clements (2018), and Krueger and Pavlova (2024).

[Figure 9 about here.]

The real-time contributions of realized outcomes to the DRPS score are shown in Figure 9, for the samples over which annual forecasts from the SPF are available at a given horizon. For chart readability, we report results for the current year and year-ahead horizons; results for longer horizons are qualitatively similar. As can be seen from the figure, for the period before the Great Recession (i.e., 1996-2006), the SPF's density forecasts of GDP growth at the current- and next-year horizons were less accurate than the BVAR's annual forecasts (on average, by roughly 20 to 30 percent). However, for the period since the Great Recession, the SPF's density forecasts are generally more accurate than the BVAR's. For real GDP growth, the SPF's advantage is sizable (roughly 50 to 60 percent) and consistent over the period since the Great Recession. In the case

of the unemployment rate and core PCE inflation, the SPF's forecasts show the largest average gains relative to the BVAR through about 2016 and then smaller gains for the remainder of the sample (with average gains over the entire sample of about 20 percent for unemployment and 10 percent for inflation). Overall, these DRPS results for annual density forecasts from the SPF compared to the BVAR indicate that, since the Great Recession, the SPF histogram forecasts provide valuable information for forecasting GDP growth, unemployment, and inflation.

5.4 Forecast accuracy of BVARs tilted to SPF histograms

[Table 1 about here.]

From this evidence of predictive content in the SPF's histograms, we turn to examining the potential forecast accuracy gains that may be achieved by using entropic tilting to inform BVAR forecasts with SPF histogram forecasts. More specifically, we compare tilted BVAR forecasts to an untilted baseline in terms of quarterly forecasts of growth, unemployment, inflation, and the T-bill rate at horizons ranging from horizons of $h = 0$ through 15 (so from the current quarter through four years ahead). Table 1 provides results for the full sample of 1996-2025. For each variable and horizon, the table reports the ratio of the RMSE or CRPS for the tilted forecast to the corresponding score for the untilted BVAR-SVO-t forecast, with values below one indicating improved performance from tilting. The tables also indicate statistical significance of the relative RMSE and CRPS as measured by Diebold-Mariano-West tests.

These results indicate that, on average over the full sample, entropically tilting BVAR forecasts to the annual SPF histograms consistently improves the BVAR's quarterly forecast accuracy, both point (RMSE) and density (CRPS). The sizes of the gains vary over horizons and across variables. The benefits of bring the SPF histograms to bear are generally smallest for GDP growth, except at $h = 0$, for which the current-year SPF forecast is particularly helpful for improving the BVAR's nowcast. At medium horizons, tilting to the

SPF histograms yields small improvements in GDP forecast accuracy (roughly 2 percent for RMSE and 5 percent for CRPS), whereas it has little impact at long horizons (with RMSE and CRPS ratios at about 1.0). Tilting to SPF histograms yields larger (albeit not statistically significant) benefits to forecasting the unemployment rate, lowering the RMSE by roughly 20 percent at medium horizons (and CRPS by a bit less) and roughly 10 percent at longer horizons before dwindling away by h = 15. For core PCE inflation and the T-bill rate, the gains to tilting to the SPF histograms fall in between those for GDP growth and unemployment. At medium horizons, forecasts for these variables show tilting gains of roughly 10 percent, commonly with statistical significance. While omitted in the interest of brevity, corresponding results for a sample ending before the pandemic are qualitatively very similar.

[Table 2 about here.]

While the full sample shows gains (on average) to our proposed tilting approach, the benefits of incorporating the information from SPF histograms could vary over time. One specific driver of time variation could be the changing availability of SPF histograms for tilting. As noted above, before 2007 and the availability of inflation or unemployment histograms from the SPF, the tilting impacts are driven entirely by the information in the current- and next-year GDP forecast histograms. SPF histograms for inflation and unemployment only impact the tilted forecasts starting in 2007 and 2009, respectively. Accordingly, Table 2 provides quarterly forecast accuracy results for a sample of 2007-2025. Broadly, the RMSE and CRPS ratios are very similar for the 2007-2025 and 1996-2025 samples. It follows that the benefits of entropic tilting BVAR-based forecasts to SPF forecast histograms are driven by the period since 2007, corresponding to the period since the Great Recession. With only GDP growth histogram forecasts available for tilting from 1996 through 2006, there is little implied benefit to bringing the SPF densities to bear in this pre-Great Recession period.

[Figure 10 about here.]

The benefits of tilting to SPF histograms could also vary within the period since the Great Recession. For example, the judgment of forecasters captured in the SPF might be especially helpful around business cycle turning points and less helpful in the middle of economic expansions. To examine time variation, we consider the RMSEs and CRPS's of the untilted and tilted forecasts computed over an expanding time sample as forecast moves forward and report time series of the difference in the cumulative average scores. For brevity, Figure 10 provides the CRPS results, for selected horizons of $h = 0$, 4, 8, and 12 quarters ahead. RMSE results are qualitatively the same.

Consistent with the comparison of the average results in the tables for the 1996-2025 and 2007-2025 samples, the time series of cumulative average CRPS differences show little consistent advantage to tilting to just the SPF's current- and next-year GDP growth histograms in roughly the first 10 years of the sample. Across horizons, the negative cumulative CRPS differences for inflation and interest rates point to some small gains to tilting in this period, with the opposite for unemployment and little net impact for GDP growth (except at the nowcast horizon of $h = 0$). Following the onset of the Great Recession in late 2007, the cumulative average loss differences dropped, often sharply, continuing for a few years in most cases. This indicates the judgment of professional forecasters embedded in the SPF histograms was especially helpful following the Great Recession. Plots of time series of the BVAR's median forecast with the SPF histogram-implied mean (omitted in the interest of brevity) show that, in this period, the SPF forecasts improve accuracy by pulling down (boosting) the model's too-strong forecasts of GDP growth and the T-bill rate (unemployment rate), while tending to edge up the model's relatively low forecasts of core inflation. Subsequently, from roughly 2014 until the COVID-19 pandemic, score differences were largely flat (in the case of inflation for the first few years of this interval) or drifted up, indicating that the raw BVAR forecasts were on balance more accurate than their SPF-tilted counterparts. Then, with the pandemic and the remainder

of the sample, relative performance showed a similar pattern, with (relative to the raw BVAR) tilted forecast accuracy improving for a time with the pandemic's outbreak and then either edging up or showing little change. Together, this time variation indicates that the information content of the SPF's histogram forecasts is most evident around turning points and otherwise not much of an improvement on BVAR forecasts.

5.5 Tilting to selected SPF histograms

[Table 3 about here.]

As noted above, while we have focused on tilting BVAR forecasts to a relatively comprehensive set of SPF histograms, one could instead choose to tilt to subsets of horizons or variables on the basis of subjective preference or perhaps performance. Many possible permutations are possible. Rather than engage in an exhaustive search and set of results, we consider a basic range of alternative sets of tilting targets, in which the BVAR forecasts are tilted to: (i) all of the available SPF forecasts for growth, unemployment, and inflation for the current and next year; (ii) all of the available SPF forecasts of GDP growth (covering all horizons, not just current- and next-year as in the baseline tilting set); (iii) the same (i.e., all forecast horizons available) for unemployment, and (iv) the same for core PCE inflation. To be clear, when we tilt to forecasts of single variable, we tilt all of the BVAR's forecasts (i.e., all variables) to those single-variable histograms.

Table 3 provides quarterly forecast accuracy results for these comparisons, for the full sample of 1996-2025 (results for 2007-2025 are qualitatively very similar and available in the online appendix). In this table, all tilting forecasts are compared against the raw BVAR forecasts, and to facilitate comparisons, results for the baseline tilting set of all variables is provided in the first column of results. To save space, results are reported for a selected subset of horizons.

Overall, none of the reduced sets of tilting targets exceeds the gains and consistency achieved by tilting to our larger baseline set of SPF histograms. Some of the alternative

approaches are not as good as the baseline tilting specification or match it, but none exceed its accuracy with any consistency. Consider first the alternative to tilting to just current- and next year forecasts of growth, unemployment, and inflation. This approach yields RMSE and CRPS ratios very close to those for the baseline tilting specification. Of course, this means that most or all of the information in SPF histograms forecasts that could improve BVAR forecasts is contained in the current- and next-year forecasts, with less information helpful in the available longer-horizon forecasts of unemployment and core PCE inflation. If we instead consider tilting to forecasts of a single variable, it tends to be the case that, for the BVAR forecasts of the variable being used as tilting information, tilting to its own forecasts from the SPF tend to be about as helpful — or sometimes a little more helpful — as tilting to the full set of SPF histograms we consider. For example, in the case of inflation, RMSE and CRPS ratios show that inflation forecasts at longer horizons can be improved by about 10-15 percent with either tilting to just the SPF's inflation forecasts or tilting to the full set of SPF histograms. However, tilting to SPF forecasts of a single variable usually has smaller benefits to BVAR forecasts of other variables than to BVAR forecasts of that variable. For example, when the BVAR forecasts of unemployment are tilted to the SPF's density forecasts of inflation, the accuracy of the BVAR's unemployment forecasts does not improve, whereas the BVAR's unemployment forecasts are improved by the baseline tilting (as well as by tilting to the SPF's unemployment forecasts).

5.6 Tilting to a few histogram-consistent percentiles

[Table 4 about here.]

With previous studies such as those mentioned in the introduction finding that point forecasts from the SPF can be used to improve forecasts from time series models, a natural question is whether the information content we find in SPF histograms is driven mostly or entirely by matching the center of the histograms, or whether there is information in the entire histogram. To assess this question, we apply our analytical solution to entropic

tilting to tilt the BVAR's forecasts to just the medians of the SPF's predictive histograms. In this comparison, the set of variables and forecast horizons in the tilting is the same as in our baseline tilting to complete SPF histograms. The histogram-implied percentiles are calculated as the midpoints of the histogram bins where the cumulative probability reaches or exceeds the desired percentile level. For example, if the cumulative probability reaches 0.50 within a bin, the median is taken as the midpoint of that bin. If the cumulative probability jumps from below 0.50 in one bin to above 0.50 in the next bin, the median is taken as the midpoint of the second bin. This approach ensures that the calculated percentiles accurately reflect the distribution of probabilities across the histogram bins.

The upper panel of Table 4 provides these results for the 1996-2025 sample (results for 2007-2025 are qualitatively the same), comparing the forecasts with tilting to implied 50th percentiles of SPF histograms with the untilted BVAR forecasts. (To save space, we omit a few longer forecast horizons.) Broadly, the gains to entropic tilting to 50th percentiles reported in this table panel are smaller and more selective than the gains shown in Table 1's baseline results. Tilting to the SPF histogram's medians has some benefit (relative to the baseline BVAR) for unemployment at short horizons and for inflation at most horizons, while harming the RMSE accuracy of GDP growth forecasts at short horizons and the RMSE accuracy of T-bill rate forecasts at most horizons. In contrast, tilting to the full histograms yields more consistent and more sizable gains.

A related question is whether most of the information in the SPF histograms could be captured by the center of the histograms coupled with some limited information on their shape. For this case, we tilt the BVAR forecasts to three percentiles of the same SPF histograms as before: 15th, 50th, and 85th. This set of moments captures some rough information about center and shape of the histograms (width and possible asymmetry).

The lower panel of Table 3 provides these results for 1996-2025 (results for 2007-2025 are qualitatively the same). Tilting the BVAR's forecasts to this set of histogram percentiles improves on tilting to just the histogram median. On average over the full sample, this

tilting to 15th, 50th, and 85th percentiles yields consistent, modest to moderate gains in accuracy for forecasts of unemployment, inflation, and the T-bill rate, with small gains for GDP growth (but sizable for $h = 0$). Qualitatively, this set of tilting targets yields benefits analogous to those achieved by tilting to the full histograms. Based on these results, the shape of SPF histograms appears to have some information content for BVAR forecasts. However, quantitatively, tilting to the full histograms yields slightly bigger gains than does tilting to the 15th, 50th, and 85th percentiles of the histograms. The better performance of full tilting is perhaps most evident for some short-horizon forecasts (e.g., GDP growth and unemployment) and for the T-bill rate across most horizons.

6 Conclusion

This paper develops a new, direct approach to combining information in the density forecasts of the SPF with model-based forecasts. Taking the SPF histograms as given — without needing to fit histogram distributions or first and higher-order moments — our approach directly tilts the predictive distribution of time series model-based forecasts to the SPF histogram bins. Our solution is applicable to simulated densities from variety of time series models, including linear VARs, Quantile VARs and DSGE models

We first develop a new analytical characterization for the tilting problem when targeting histogram bins. To examine the empirical merits of the approach, we then examine the information content of SPF histogram forecasts for the accuracy of forecasts of major macroeconomic aggregates for the U.S. produced by a Bayesian VAR with time-varying volatility.

Using our analytical solution for the tilting problem, we entropically tilt the BVAR's quarterly forecast distributions to SPF histograms, including multi-year forecast horizons from the SPF for GDP growth, the unemployment rate, and core PCE inflation. Relative to the baseline model forecasts, SPF histogram forecasts add the most information content

around the Great Recession and its aftermath during the COVID-19 pandemic, with most of the information coming from the SPF's unemployment rate forecasts from one through three years ahead. On average over samples of 1996-2025 and 2007-2025, entropically tilting BVAR forecasts to the annual SPF histograms consistently improves the BVAR's quarterly forecast accuracy (point and density). The benefits of entropically tilting the BVAR forecasts to the SPF histograms are specifically driven by the years around the Great Recession and around COVID-19's outbreak in 2020. In more normal periods, the BVAR's forecasts are not improved by bringing information from the SPF to bear through tilting to the survey's histogram forecasts.

In future applications we intend to extend the set of tilting targets to cover all available SPF histogram forecasts, specifically to add those for longer horizons of GDP, and to incorporate the SPF's recession-probability forecasts. We also plan to examine the effects of tilting on the joint distribution of outcomes, such as GDP and inflation.

A Additional Figures

This Appendix provides additional figures. Figures 11 and 12 extend the results on coherence between SPF point and histogram, shown in Figure 5 above for the case of GDP growth, to SPF histograms for core inflation and the unemployment rate.

[Figure 11 about here.]

[Figure 12 about here.]

References

Adrian, Tobias, Domenico Giannone, Matteo Luciani, and Mike West (2025), "Scenario synthesis and macroeconomic risk," Manuscript, March 28.

Allayioti, Anastasia, Rodolfo Arioli, Colm Bates, Vasco Botelho, Bruno Fagandini, Luís Fonseca, Peter Healy, Aidan Meyler, Ryan Minasian, and Octavia Zahrt (2024), "A look

back at 25 years of the ECB SPF," ECB Occasional Paper Series 364, European Central Bank.

Antolín-Díaz, Juan, Ivan Petrella, and Juan F. Rubio-Ramírez (2021), "Structural scenario analysis with SVARs," *Journal of Monetary Economics*, 117, 798–815, https://doi.org/10.1016/j.jmoneco.2020.06.001.

Banbura, Marta, Federica Brenna, Joan Paredes, and Francesco Ravazzolo (2021), "Combining Bayesian VARs with survey density forecasts: Does it pay off?" Working Paper Series 2543, European Central Bank.

Bassetti, Federico, Roberto Casarin, and Marco Del Negro (2022), "A Bayesian approach to inference on probabilistic surveys," Staff Reports 1025, Federal Reserve Bank of New York.

——— (2023), "Inference on probabilistic surveys in macroeconomics with an application to the evolution of uncertainty in the survey of professional forecasters during the covid pandemic," in *Handbook of Economic Expectations* eds. by Rüdiger Bachmann, Giorgio Topa, and Wilbert van der Klaauw: Academic Press, chap. 15, 443–476, https://doi.org/10.1016/B978-0-12-822927-9.00023-9.

Bobeica, Elena, and Benny Hartwig (2023), "The COVID-19 shock and challenges for inflation modelling," *International Journal of Forecasting*, 39, 519–539, https://doi.org/10.1016/j.ijforecast.2022.01.002.

Boero, Gianna, Jeremy Smith, and Kenneth F. Wallis (2011), "Scoring rules and survey density forecasts," *International Journal of Forecasting*, 27, 379–393, https://doi.org/10.1016/j.ijforecast.2010.04.003.

Breitung, Jörg, and Malte Knüppel (2021), "How far can we forecast? Statistical tests of the predictive content," *Journal of Applied Econometrics*, 36, 369–392, https://doi.org/10.1002/jae.2817.

Carriero, Andrea, Joshua C. C. Chan, Todd E. Clark, and Massimiliano Marcellino (2022), "Corrigendum to: Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors," *Journal of Econometrics*, 227, 506–512, https://doi.org/10.1016/j.jeconom.2021.11.010.

Carriero, Andrea, Todd E. Clark, and Massimiliano Marcellino (2019), "Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors," *Jour-*

nal of Econometrics, 212, 137–154, https://doi.org/10.1016/j.jeconom.2019.04.024.

Carriero, Andrea, Todd E. Clark, Massimiliano Marcellino, and Elmar Mertens (2024), "Addressing COVID-19 outliers in BVARs with stochastic volatility," *The Review of Economics and Statistics*, 106, 1403–1417, https://doi.org/10.1162/rest_a_01213.

Chernis, Tony, Niko Hauzenberger, Florian Huber, Gary Koop, and James Mitchell (2025), "Predictive density combination using Bayesian machine learning," *International Economic Review*, 66, 1287–1315, https://doi.org/10.1111/iere.12759.

Clark, Todd E., Gergely Ganics, and Elmar Mertens (2022), "What is the predictive value of SPF point and density forecasts?" Working paper 22-37, Federal Reserve Bank of Cleveland, https://doi.org/10.26509/frbc-wp-202237.

Clark, Todd E., and Elmar Mertens (2024), "Survey expectations and forecast uncertainty," in *Handbook of Research Methods and Applications in Macroeconomic Forecasting*, Cheltenham, UK: Edward Elgar Publishing, 305–333.

Clements, Michael P. (2004), "Evaluating the Bank of England density forecasts of inflation," *Journal of Forecasting*, 23, 463–473, https://doi.org/10.1002/for.928.

——— (2008), "Survey-based expectations and the Phillips curve," *Journal of Macroeconomics*, 30, 555–572, https://doi.org/10.1016/j.jmacro.2007.06.001.

——— (2010), "Explanations of the inconsistencies in survey respondents' forecasts," *European Economic Review*, 54, 536–549, https://doi.org/10.1016/j.euroecorev.2009.10.003.

——— (2014), "Probability distributions or point predictions? Survey forecasts of US output growth and inflation," *International Journal of Forecasting*, 30, 99–117, https://doi.org/10.1016/j.ijforecast.2013.07.010.

——— (2018), "Are macroeconomic density forecasts informative?" *International Journal of Forecasting*, 34, 181–198, https://doi.org/10.1016/j.ijforecast.2017.10.004.

——— (2025), "Inconsistent survey histograms and point forecasts revisited," *Journal of Economic Behavior & Organization*, 236, https://doi.org/10.1016/j.jebo.2025.107097.

Clements, Michael P., Robert W. Rich, and Joseph Tracy (2025), "An investigation into the uncertainty revision process of professional forecasters," *Journal of Economic Dynamics and Control*, 173, https://doi.org/10.1016/j.jedc.2025.105060.

Cogley, Timothy, Sergei Morozov, and Thomas J. Sargent (2005), "Bayesian fan charts for U.K. inflation: Forecasting and sources of uncertainty in an evolving monetary system," *Journal of Economic Dynamics and Control*, 29, 1893–1925, https://doi.org/10.1016/j.jedc.2005.06.005.

Coroneo, Laura, Fabrizio Iacone, and Fabio Profumo (2024), "Survey density forecast comparison in small samples," *International Journal of Forecasting*, 40, 1486–1504, https://doi.org/10.1016/j.ijforecast.2023.10.005.

Cover, Thomas A., and Joy A. Thomas (2006), *Elements of Information Theory 2nd Edition*, Wiley Series in Telecommunications and Signal Processing: Wiley, 2nd edition.

Croushore, Dean, and Tom Stark (2001), "A real-time data set for macroeconomists," *Journal of Econometrics*, 105, 111–130, https://doi.org/10.1016/S0304-4076(01)00072-0.

Csiszár, Imre (1975), "I-divergence geometry of probability distributions and minimization problems," *Annals of Probability*, 3, 146–158.

Deming, W. Edwards, and Frederick F. Stephan (1940), "On a least squares adjustment of a sampled frequency table when the expected marginal totals are known," *Annals of Mathematical Statistics*, 11, 427–444.

Diebold, Francis X., Minchul Shin, and Boyuan Zhang (2023), "On the aggregation of probability assessments: Regularized mixtures of predictive densities for Eurozone inflation and real interest rates," *Journal of Econometrics*, 237, p. 105321, https://doi.org/10.1016/j.jeconom.2022.06.008.

Engelberg, Joseph, Charles F. Manski, and Jared Williams (2009), "Comparing the point predictions and subjective probability distributions of professional forecasters," *Journal of Business & Economic Statistics*, 27, 30–41, https://doi.org/10.1198/jbes.2009.0003.

Faust, Jon, and Jonathan H. Wright (2013), "Forecasting inflation," in *Handbook of Economic Forecasting* eds. by Graham Elliott, and Allan Timmermann, 2A of Handbook of Economic Forecasting: Elsevier, chap. 1, 2–56, https://doi.org/10.1016/B978-0-444-53683-9.00001-3.

Fienberg, Stephen E. (1970), "The iterative proportional fitting procedure," *Biometrika*, 57, 617–624.

Galvão, Ana Beatriz, Anthony Garratt, and James Mitchell (2021), "Does judgment improve macroeconomic density forecasts?" *International Journal of Forecasting*, 37, 1247–1260, https://doi.org/10.1016/j.ijforecast.2021.02.007.

Ganics, Gergely, and Florens Odendahl (2021), "Bayesian VAR forecasts, survey information, and structural change in the euro area," *International Journal of Forecasting*, 37, 971–999, https://doi.org/10.1016/j.ijforecast.2020.11.001.

Giacomini, Raffaella, and Giuseppe Ragusa (2014), "Theory-coherent forecasting," *Journal of Econometrics*, 182, 145–155, https://doi.org/10.1016/j.jeconom.2014.04.006.

Jacquier, Eric, Nicholas G. Polson, and Peter E. Rossi (2004), "Bayesian analysis of stochastic volatility models with fat-tails and correlated errors," *Journal of Econometrics*, 122, 185–212, https://doi.org/10.1016/j.jeconom.2003.09.001.

Kenny, Geoff, Thomas Kostka, and Federico Masera (2014), "How informative are the subjective density forecasts of macroeconomists?" *Journal of Forecasting*, 33, 163–185, https://doi.org/10.1002/for.2281.

——— (2015), "Density characteristics and density forecast performance: A panel analysis," *Empirical Economics*, 48, 1203–1231, https://doi.org/10.1007/s00181-014-0815-9.

Knüppel, Malte, and Lora Pavlova (2023), "Assessing the ex-ante uncertainty in the US SPF," Mimeo.

Krueger, Fabian, Todd E. Clark, and Francesco Ravazzolo (2017), "Using entropic tilting to combine BVAR forecasts with external nowcasts," *Journal of Business & Economic Statistics*, 35, 470–485, https://doi.org/10.1080/07350015.2015.1087856.

Krueger, Fabian, and Lora Pavlova (2024), "Quantifying subjective uncertainty in survey expectations," *International Journal of Forecasting*, 40, 796–810, https://doi.org/https://doi.org/10.1016/j.ijforecast.2023.06.001.

Lenza, Michele, Inès Moutachaker, and Joan Paredes (2025), "Density forecasts of inflation: A quantile regression forest approach," *European Economic Review*, 178, p. 105079, https://doi.org/10.1016/j.euroecorev.2025.105079.

Meyler, Aidan (2020), "Forecast performance in the ecb spf: ability or chance?" ECB Working Paper Series 2371, European Central Bank.

Montes-Galdon, Carlos, Joan Paredes, and Elias Wolf (2022), "Conditional density forecasting: a tempered importance sampling approach," Working Paper Series 2754, European Central Bank.

Rich, Robert W., Joseph Song, and Joseph Tracy (2012), "The measurement and behavior of uncertainty: evidence from the ecb survey of professional forecasters," Staff Report 588, Federal Reserve Bank of New York, https://doi.org/10.2139/ssrn.2053287.

Robertson, John C., Ellis W. Tallman, and Charles H. Whiteman (2005), "Forecasting using relative entropy," *Journal of Money, Credit and Banking*, 37, 383–401, https://doi.org/10.1353/mcb.2005.0034.

Schennach, Susanne M. (2007), "Point estimation with exponentially tilted empirical likelihood," *The Annals of Statistics*, 35, 634–672.

Stock, James H., and Mark W. Watson (2016), "Core inflation and trend inflation," *The Review of Economics and Statistics*, 98, 770–784, https://doi.org/10.1162/REST_a_00608.

Tallman, Ellis W., and Saeed Zaman (2020), "Combining survey long-run forecasts and nowcasts with BVAR forecasts using relative entropy," *International Journal of Forecasting*, 36, 373–398, https://doi.org/10.1016/j.ijforecast.2019.04.024.

Tallman, Emily, and Mike West (2022), "On entropic tilting and predictive conditioning," Papers 2207.10013, arXiv.org.

Waggoner, Daniel F., and Tao Zha (1999), "Conditional forecasts in dynamic multivariate models," *The Review of Economics and Statistics*, 81, 639–651, https://doi.org/10.1162/003465399558508.

West, Mike (2024), "Perspectives on constrained forecasting," *Bayesian Analysis*, 19, 1013–1039, https://doi.org/10.1214/23-BA1379.

Table 1: Accuracy of tilted and untilted forecasts from BVAR-SVO-t model: 1996-2025

	RMSE				CRPS			
h	Output	Unempl.	Inflation	Int. Rate	Output	Unempl.	Inflation	Int. Rate
0	0.60	0.52	0.94	0.97	0.79	0.63	0.95	0.96
1	1.01	0.71	1.02	0.97	0.93*	0.71	1.01	0.95**
2	0.99	0.76	0.99	0.95**	0.94	0.79	0.99	0.93***
3	0.97	0.78	0.92	0.94**	0.93	0.81	0.96	0.92**
4	0.97	0.79	0.93	0.92**	0.93	0.83	0.93	0.91**
5	0.98	0.81	0.90	0.92**	0.95	0.86	0.91*	0.90**
6	0.98	0.83	0.90	0.91**	0.95*	0.88	0.92*	0.90**
7	1.00	0.86	0.89*	0.91**	0.98	0.89	0.90**	0.90**
8	1.00	0.88	0.88*	0.91**	0.99	0.91	0.90**	0.90**
9	0.99*	0.90	0.90	0.91**	0.98	0.93	0.92**	0.91*
10	0.99	0.91	0.91	0.92**	0.99	0.95	0.92**	0.91*
11	1.00	0.93	0.91	0.92**	1.00	0.97	0.92**	0.91*
12	1.00	0.94	0.94**	0.92**	1.00	0.99	0.93***	0.91*
13	1.00	0.95	0.94*	0.92**	1.01	1.00	0.94**	0.91*
14	1.00	0.97	0.90*	0.92**	1.01	1.01	0.92***	0.91*
15	1.00	0.99	0.93**	0.93**	1.01	1.02	0.94***	0.92*

Note: Forecasts for quarterly outcomes, h quarters ahead. Relative RMSE and CRPS of BVAR-SVO(t)-AR(1) model tilted against SPF histograms for inflation, unemployment and GDP (GDP: only current and next year), with BVAR-SVO(t)-AR(1) in denominator. Evaluation window from 1996Q1 through 2025Q1 (and as far as realized values and tilting targets are available). Reflecting availability of SPF data, the first forecast origin used in this evaluation is 1996Q1. Significance assessed by Diebold-Mariano tests using Newey-West standard errors with $h + 1$ lags. ***, ** and * denote significance at the 1%, 5%, and 10% level, respectively.

Table 2: Accuracy of tilted and untilted forecasts from BVAR-SVO-t model: 2007-2025

	RMSE				CRPS			
h	Output	Unempl.	Inflation	Int. Rate	Output	Unempl.	Inflation	Int. Rate
0	0.56	0.51	0.91	0.97	0.73	0.58	0.92	0.95
1	1.00	0.71	1.02	0.97	0.90**	0.67	1.01	0.94*
2	0.98	0.76	0.97	0.95	0.91	0.77	0.98	0.92**
3	0.96**	0.78	0.88	0.93*	0.89**	0.79	0.94	0.90**
4	0.97**	0.78	0.89	0.91**	0.90**	0.81	0.89	0.89**
5	0.97*	0.80	0.86*	0.89**	0.93	0.84	0.87*	0.87**
6	0.98*	0.83	0.86	0.88**	0.93**	0.85	0.88*	0.87**
7	1.00	0.85	0.85*	0.88**	0.97	0.87	0.86***	0.86**
8	1.00	0.87	0.84*	0.87**	0.98	0.89	0.86**	0.86**
9	0.99**	0.89	0.87	0.87**	0.98*	0.91	0.88**	0.87*
10	0.99*	0.90	0.87	0.88**	0.98	0.94	0.89**	0.87*
11	0.99	0.91	0.87	0.87**	0.99	0.96	0.88**	0.86*
12	1.00	0.93	0.91*	0.87**	1.00	0.98	0.90***	0.86*
13	1.00	0.94	0.92	0.87**	1.01	1.00	0.91**	0.86**
14	1.00	0.96	0.86	0.88**	1.01	1.02	0.88***	0.86**
15	1.00	0.98	0.91*	0.88**	1.01	1.04	0.91***	0.87**

Note: Forecasts for quarterly outcomes, h quarters ahead. Relative RMSE and CRPS of BVAR-SVO(t)-AR(1) model tilted against SPF histograms for inflation, unemployment and GDP (GDP: only current and next year), with BVAR-SVO(t)-AR(1) in denominator. Evaluation window from 2007Q1 through 2025Q1 (and as far as realized values and tilting targets are available). Reflecting availability of SPF data, the first forecast origin used in this evaluation is 2007Q1. Significance assessed by Diebold-Mariano tests using Newey-West standard errors with $h + 1$ lags. ***, ** and * denote significance at the 1%, 5%, and 10% level, respectively.

Table 3: Forecast accuracy when tilting to different targets (1996-2025)

	RMSE					CRPS				
					Output					
h	Baseline	All 0+1y	GDP	U	Inf	Baseline	All 0+1y	GDP	U	Inf
0	0.60	0.62	0.55	0.69	0.97	0.79	0.80	0.74	0.86	1.00
1	1.01	1.03	1.02	0.97*	1.00	0.93*	0.93	0.93*	0.93**	1.01
2	0.99	0.99	0.98	0.99	1.00	0.94	0.94	0.94	0.96	1.01*
3	0.97	0.97	0.98	0.98*	1.00	0.93	0.94	0.94	0.93*	1.00
7	1.00	1.00	0.98*	0.99	1.00	0.98	1.00	0.94*	0.97	0.98**
11	1.00	1.00	0.99	1.00**	0.99***	1.00	1.01	0.97	0.99	0.98***
15	1.00	1.00	1.00	1.01	1.00***	1.01	1.01	1.00	1.02	0.99***
					Unempl.					
0	0.52	0.51	0.91	0.41	0.99	0.63	0.63	0.92	0.53	0.99
1	0.71	0.72	0.91	0.71	0.99	0.71	0.71	0.89	0.66	1.00
2	0.76	0.78	0.90	0.76	0.99	0.79	0.80	0.87*	0.77	1.00
3	0.78	0.80	0.89	0.76	0.99	0.81	0.82	0.87**	0.78	1.00
7	0.86	0.88	0.91*	0.80	0.99	0.89	0.90	0.91	0.84	1.00
11	0.93	0.94	0.96	0.87	1.00	0.97	0.95	0.96	0.95	1.00
15	0.99	0.99	1.00	0.96	1.00	1.02	1.00	1.02	1.05	1.01
					Inflation					
0	0.94	0.95	0.96	0.92	0.93	0.95	0.96	0.97	0.93	0.94
1	1.02	1.01	1.01	0.99	1.01	1.01	1.01	1.00	0.99	1.00
2	0.99	0.98	0.99	0.94	0.96	0.99	0.99	1.00	0.95	0.97
3	0.92	0.92	0.98	0.91	0.88	0.96	0.96	1.00	0.96	0.94
7	0.89*	0.89*	0.95	0.94	0.85*	0.90**	0.90**	0.96**	0.99	0.86**
11	0.91	0.91*	0.95*	0.98	0.89	0.92**	0.92**	0.97*	1.02	0.89**
15	0.93**	0.93*	0.97*	0.98	0.93	0.94***	0.94***	0.98	1.02	0.93**
					Int. Rate					
0	0.97	0.97	0.96	0.98	1.03**	0.96	0.96*	0.96*	0.96	1.02*
1	0.97	0.96*	0.96**	0.99	1.04**	0.95**	0.94**	0.94***	0.95**	1.05**
2	0.95**	0.95**	0.94***	0.98	1.04**	0.93***	0.92***	0.92***	0.93*	1.06**
3	0.94**	0.93**	0.93***	0.96	1.05**	0.92**	0.91***	0.91***	0.91*	1.08***
7	0.91**	0.91**	0.91***	0.87*	1.05**	0.90**	0.91**	0.90***	0.83*	1.09***
11	0.92**	0.95**	0.93**	0.80**	1.05**	0.91*	0.95*	0.92**	0.76**	1.09***
15	0.93**	0.97**	0.95*	0.79*	1.04	0.92*	0.97*	0.94	0.75**	1.07*

Note: Forecasts for quarterly outcomes, h quarters ahead. Relative RMSE and CRPS of SVO-t AR(1) model tilted against different SPF histogram targets (untilted SVO-t AR(1) in denominator). Sample: "full sample". Tilting targets as indicated in column headers. Significance assessed by Diebold-Mariano tests using Newey-West standard errors with $h+1$ lags. ***, ** and * denote significance at the 1%, 5%, and 10% level, respectively.

Table 4: Accuracy of forecasts tilted to selected SPF percentiles: 1996-2025

	RMSE				CRPS			
h	Output	Unempl.	Inflation	Int. Rate	Output	Unempl.	Inflation	Int. Rate
Tilting to 50th percentiles of SPF histograms								
0	1.55*	0.72	0.91	1.10**	1.19***	0.86	0.97	1.04
1	1.15	0.87	1.10	1.12*	1.00	0.96	1.04	1.04
2	1.07	1.02	1.00	1.12	1.02	1.00	1.01	1.01
3	1.00	1.03	0.92	1.10	0.99	1.00	0.96	0.99
4	0.99	1.01	0.97	1.09	0.97	0.97	0.95	0.98
5	0.99	0.99	0.93	1.07	0.99	0.96	0.95	0.97
6	0.98*	1.01	0.94	1.05	0.99	0.94	0.95	0.96
7	1.00	1.00	0.91	1.03	1.01	0.94	0.95	0.95
11	1.00	0.97	0.94	0.99	1.02	0.97	0.97	0.95
15	1.01**	1.01	0.94*	0.98	1.02**	1.02	0.99	0.96
Tilting to 15th, 50th, and 85th percentiles of SPF histograms								
0	0.65	0.66	0.92	1.02	0.93	0.77	0.96	1.00
1	1.09	0.75	1.02	1.01	0.96	0.82	1.00	0.97
2	1.00	0.88	0.98	1.01	0.96	0.86	0.97	0.96
3	0.98	0.85	0.90	0.99	0.95	0.84	0.93	0.94*
4	0.97*	0.84	0.93	0.98	0.93*	0.84	0.92*	0.93*
5	0.98	0.85	0.90*	0.97	0.96	0.85	0.90**	0.92**
6	0.99	0.87	0.90	0.96	0.97	0.86	0.91*	0.91**
7	0.99	0.88	0.90	0.95	0.98	0.87	0.91**	0.91**
11	1.01	0.93	0.90*	0.94*	1.01	0.96	0.92**	0.92*
15	1.01*	0.98	0.93**	0.94**	1.02	1.00	0.95**	0.93**

Note: Forecasts for quarterly outcomes, h quarters ahead. Relative RMSE and CRPS of SVO-t AR(1) model tilted against Inflation and Unemployment and GDP Nowcast (medians) (untilted SVO-t AR(1) in denominator). Evaluation window from 1996Q1 through 2025Q1 (and as far as realized values and tilting targets are available). Reflecting availability of SPF data, the first forecast origin used in this evaluation is 1996Q1. Significance assessed by Diebold-Mariano tests using Newey-West standard errors with $h + 1$ lags. ***, ** and * denote significance at the 1%, 5%, and 10% level, respectively.

Figure 1: SPF histogram for GDP growth per 2025Q3

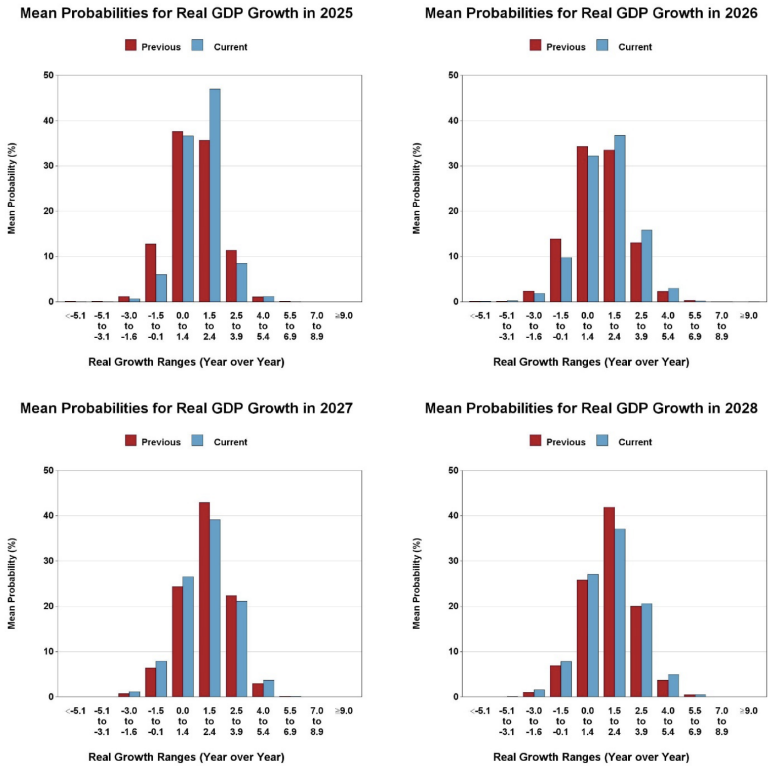


Figure 2: SPF histogram for unemployment rate per 2025Q3

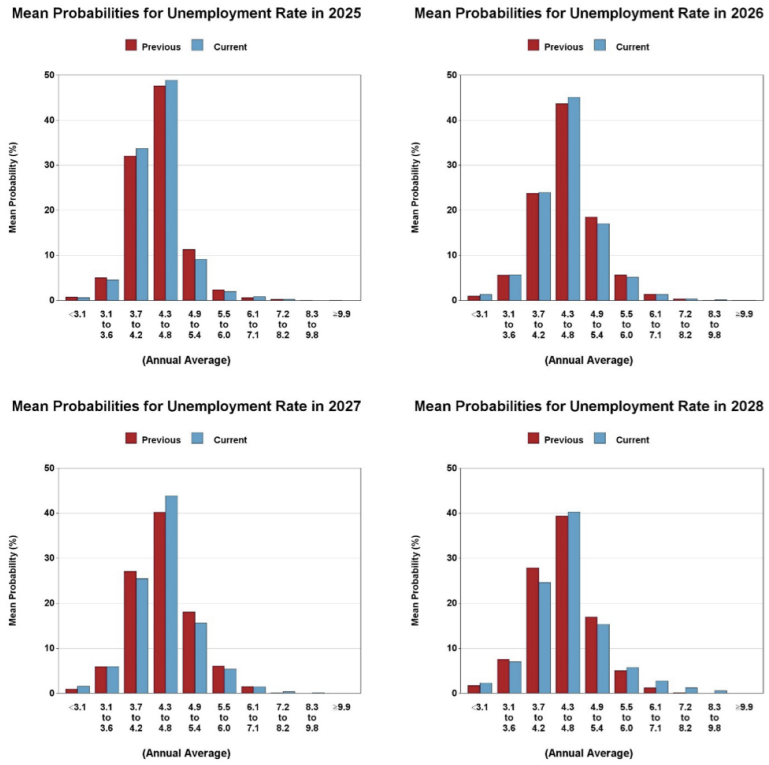


Figure 3: SPF histogram for core PCE inflation per 2025Q3

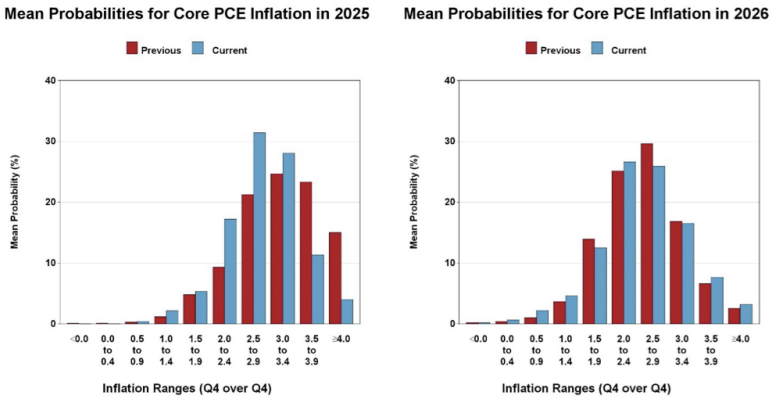
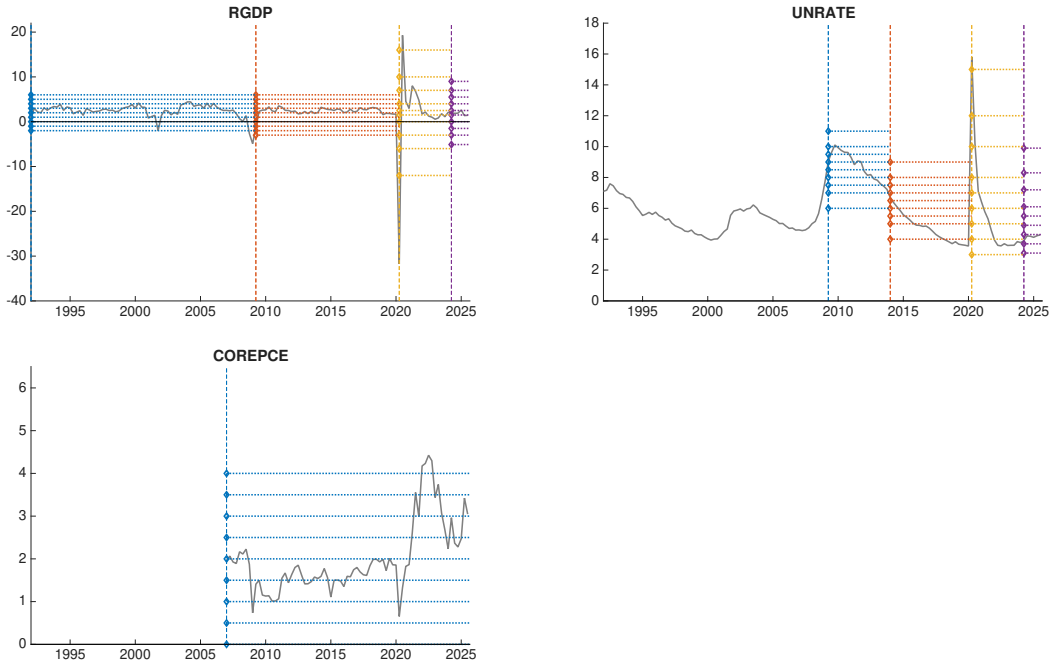
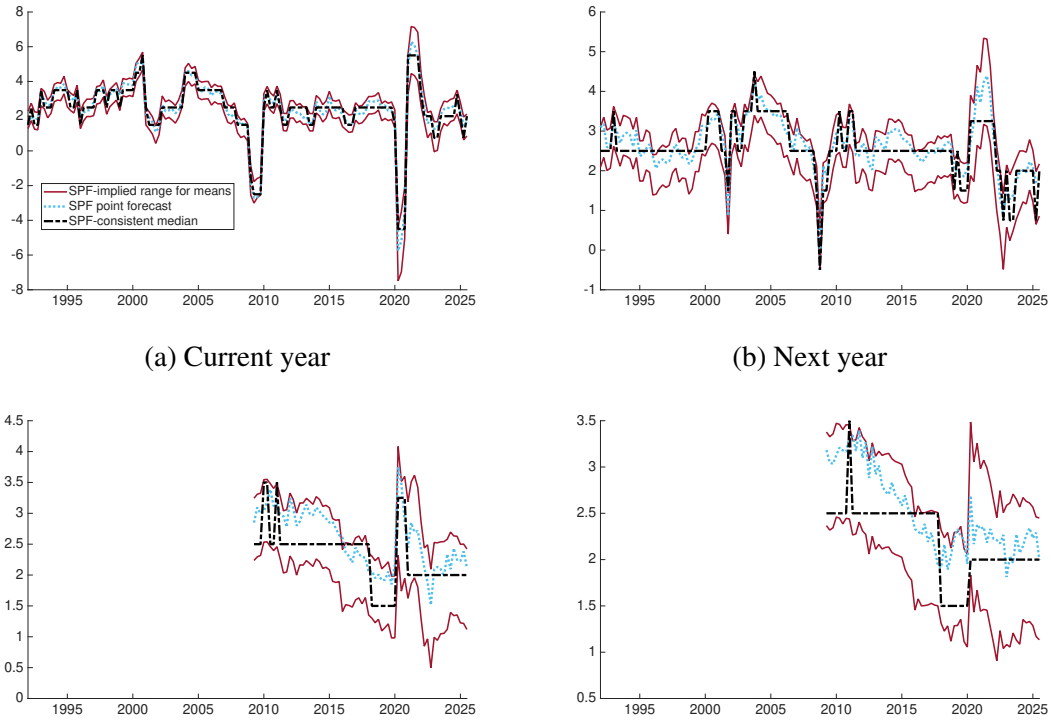


Figure 4: SPF histogram bins over time (since 1992)



Note: Solid grey lines provide historical data on growth and unemployment (current vintage).

Figure 5: SPF-implied mean and median forecasts for real GDP growth



(a) Current year

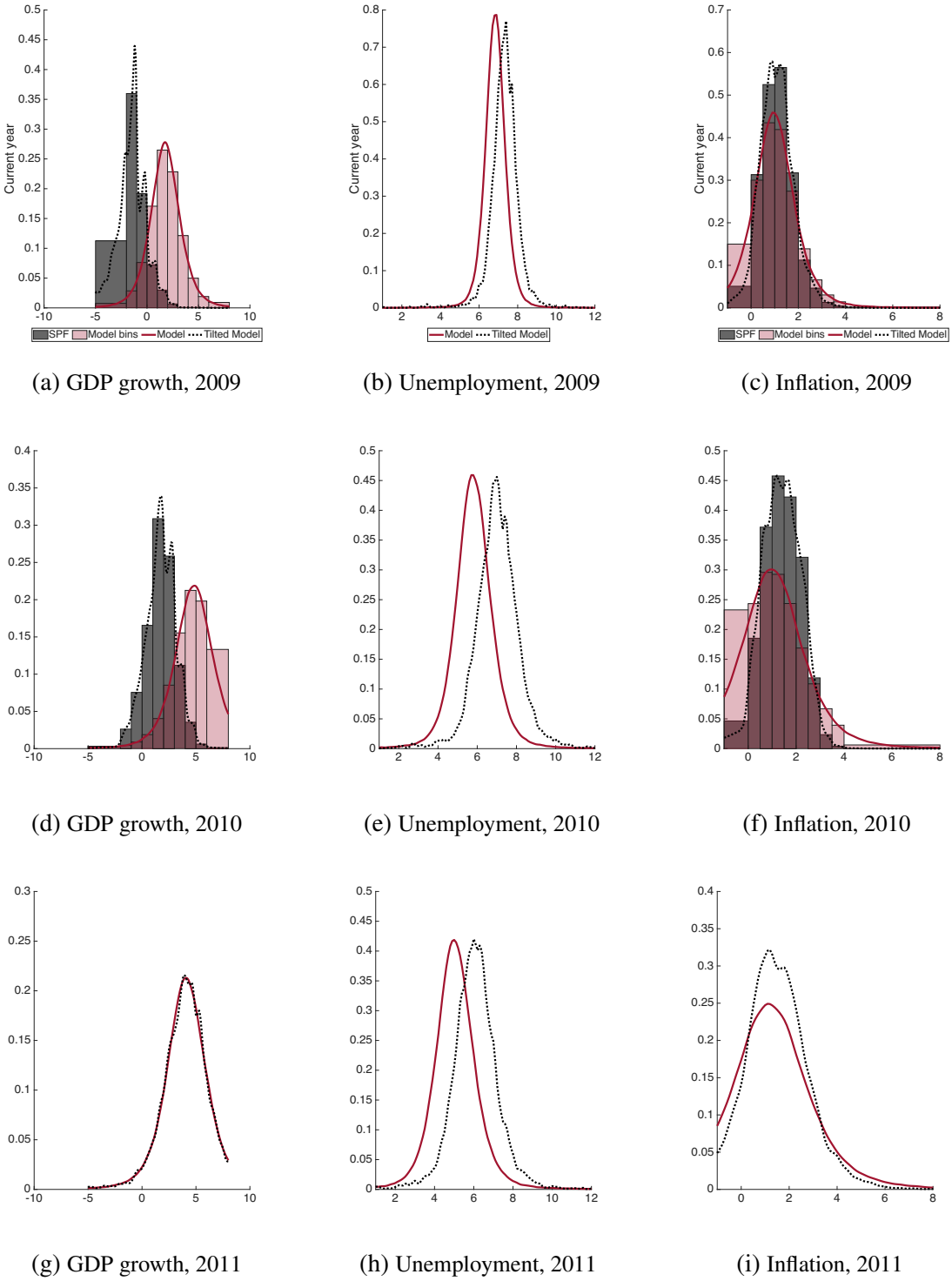
(b) Next year

(c) Two years ahead

(d) Three years ahead

Note: SPF-consistent median is the midpoint of the bin containing the median. Upper and lower bounds for SPF-consistent means are computed from lower and upper bounds of the bins. The point forecasts are the average of individual forecasters' point forecasts. Data is shown for available SPF histograms since 1992.

Figure 6: Examples of tilted and untilted predictive densities: 2009Q1



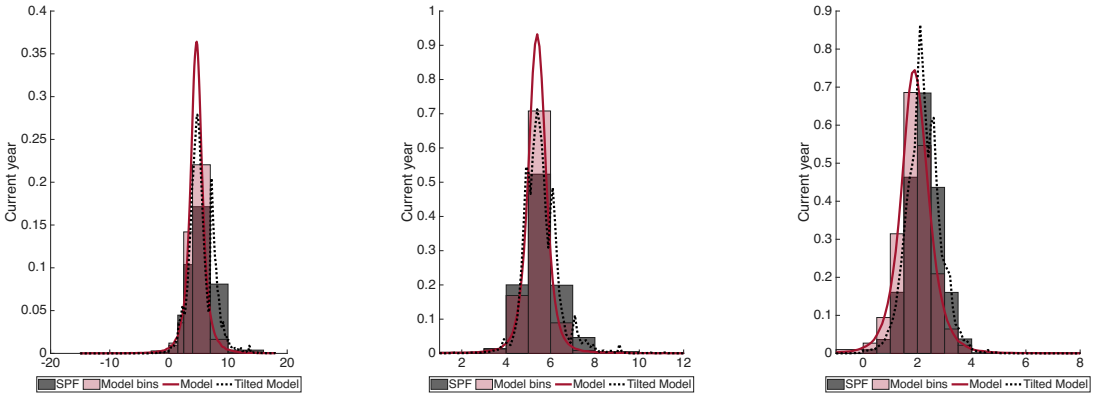
(a) GDP growth, 2009 (b) Unemployment, 2009 (c) Inflation, 2009

(d) GDP growth, 2010 (e) Unemployment, 2010 (f) Inflation, 2010

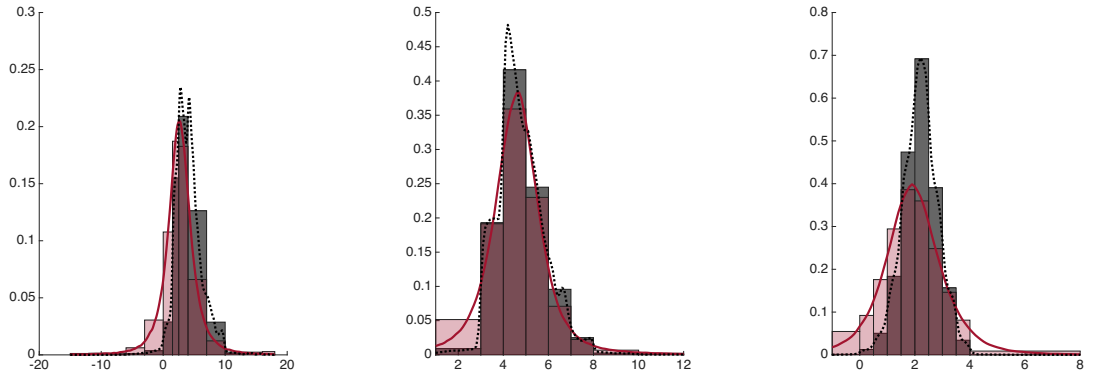
(g) GDP growth, 2011 (h) Unemployment, 2011 (i) Inflation, 2011

Note: Model-based densities in red, tilted densities and SPF histograms in black.

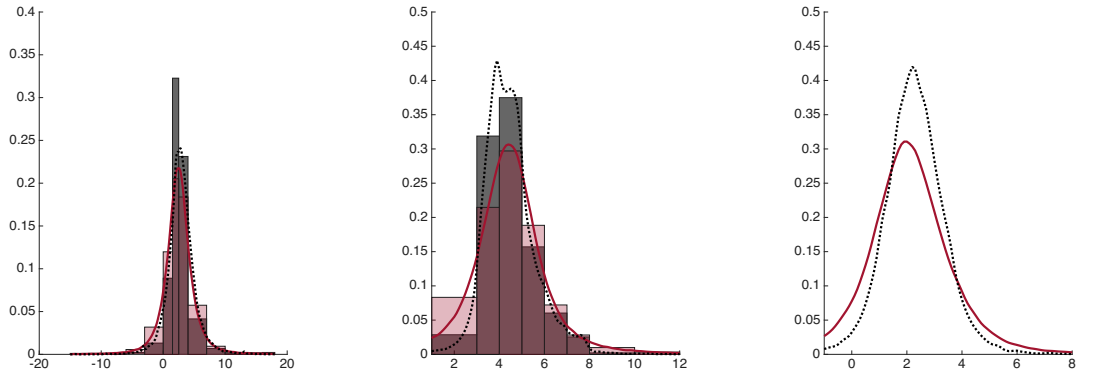
Figure 7: Examples of tilted and untilted predictive densities: 2021Q2



(a) GDP growth, 2021 (b) Unemployment, 2021 (c) Inflation, 2021



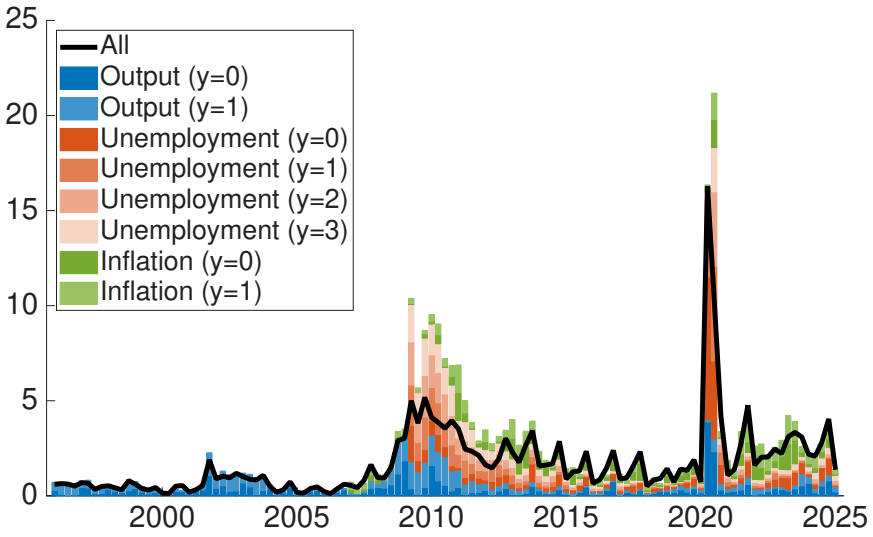
(d) GDP growth, 2022 (e) Unemployment, 2022 (f) Inflation, 2022



(g) GDP growth, 2023* (h) Unemployment, 2023 (i) Inflation, 2023

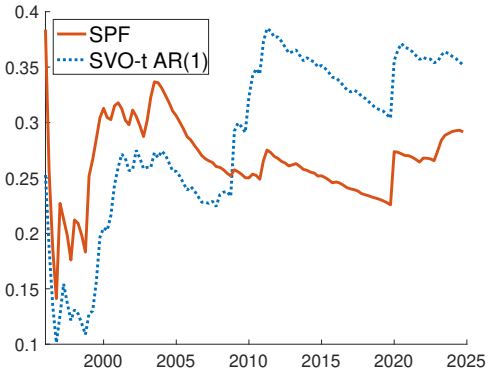
Note: Asterisk denotes non-targeted SPF histogram. Model-based densities in red, tilted densities and SPF histograms in black.

Figure 8: KLIC decomposition over time

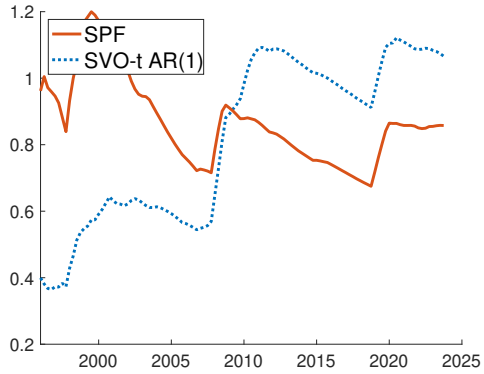


Note: KLICs from entropic tilting at different forecast origins. 'All' refers to our baseline set of targets, and the colored bars indicate the KLICs obtained from targeting a given individual histogram.

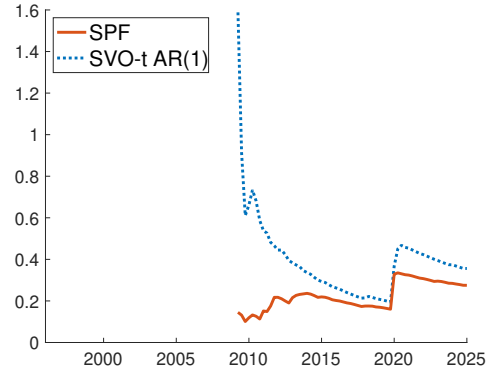
Figure 9: DRPS for forecasts from BVAR-SVO-t model and SPF



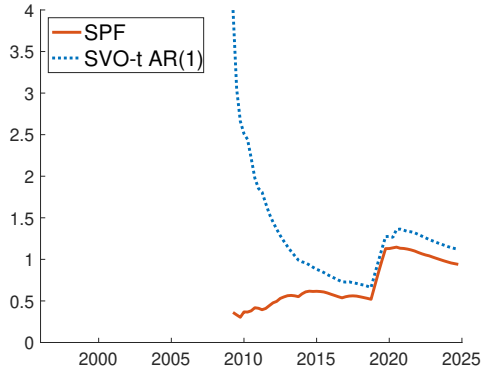
(a) GDP growth, current year



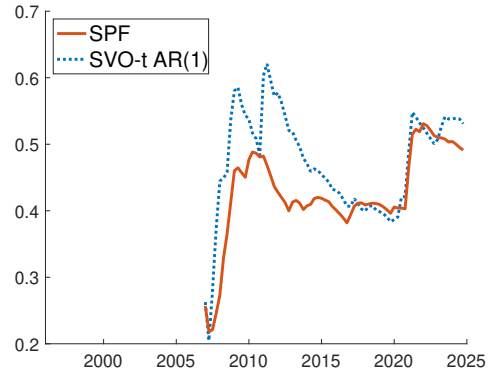
(b) GDP growth, next year



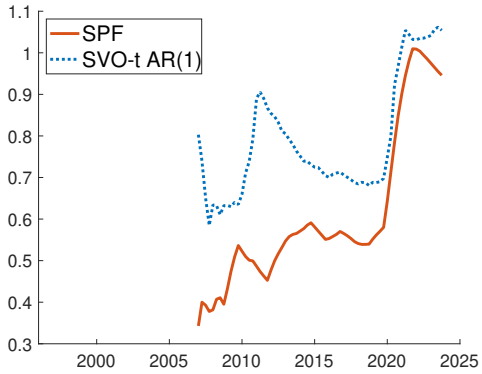
(c) Unemployment rate, current year



(d) Unemployment rate, next year



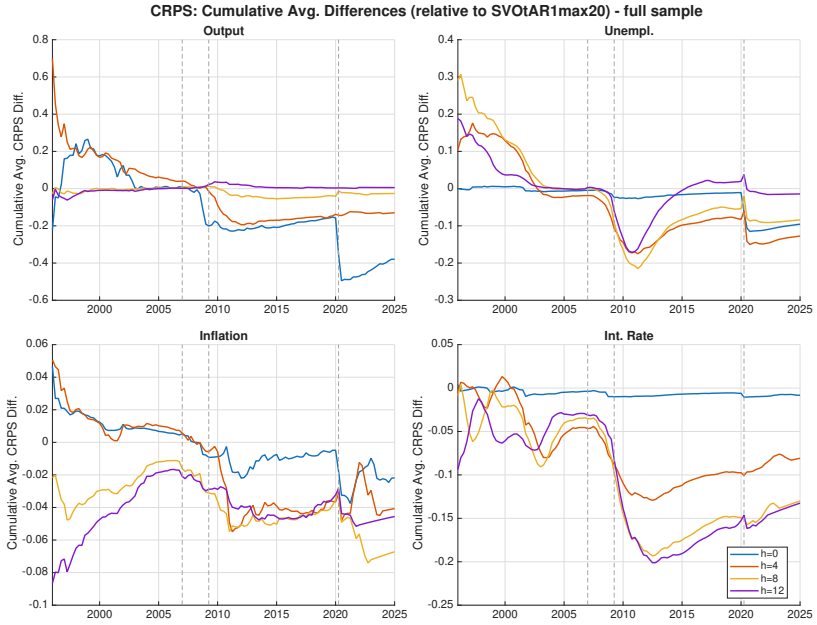
(e) Core PCE inflation, current year



(f) Core PCE inflation, next year

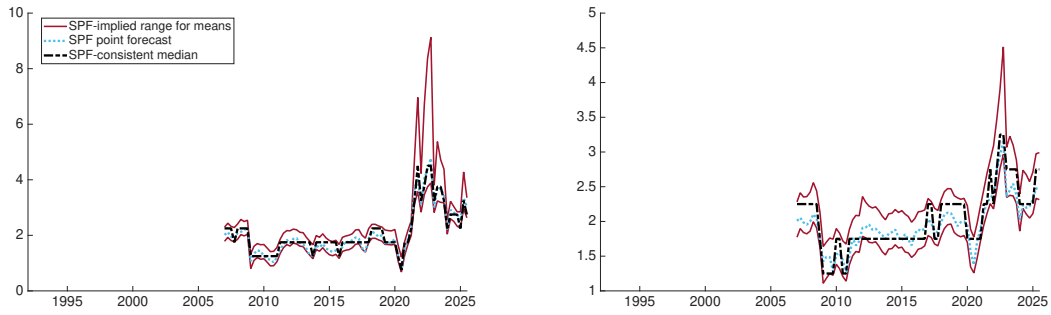
Note: The figure shows time-averaged discrete rank probability score (DRPS) computed over growing windows. DRPS are shown for the BVAR-SVO-t model and SPF histogram forecasts. The DRPS is calculated for current year and one-year-ahead forecasts, respectively. Lower values indicate better density-forecast performance. The x-axis represents the forecast horizon, while the y-axis indicates the cumulative average of the relative probability score.

Figure 10: Accuracy of tilted to untilted density forecasts over time



Note: Cumulative average CRPS differences between tilted and untilted densities generated from BVAR-SVO-t. Negative values indicate better CRPS of tilted densities.

Figure 11: SPF-implied mean and median forecasts for core PCE inflation

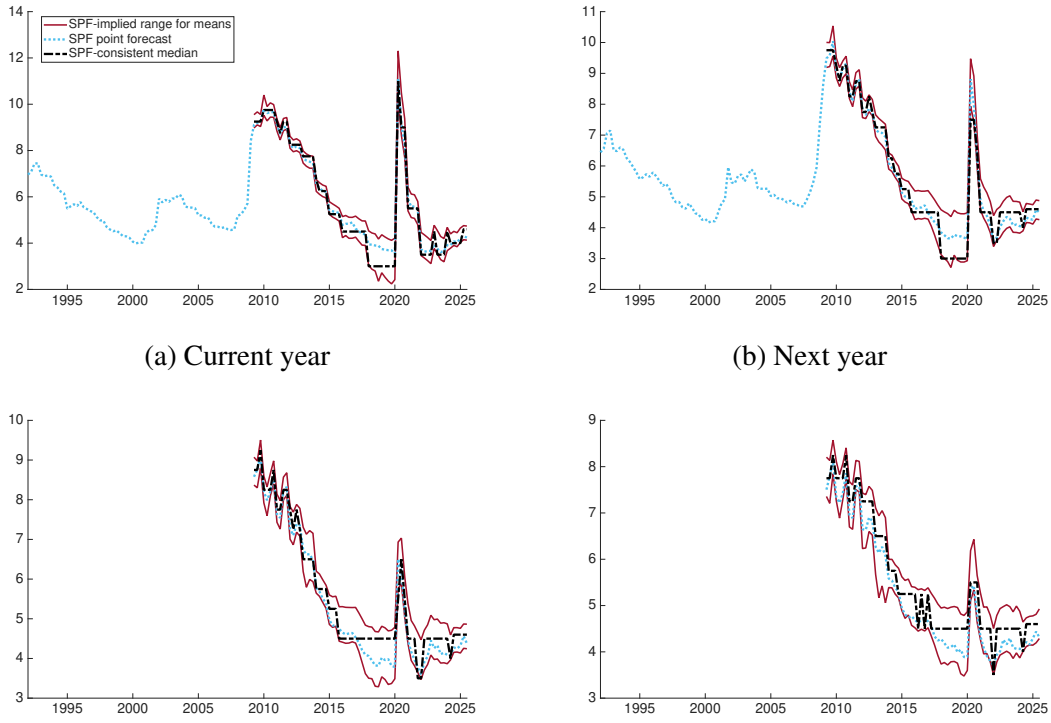


(a) Current year

(b) Next year

Note: SPF-consistent median is the midpoint of the bin containing the median. Upper and lower bounds for SPF-consistent means are computed from lower and upper bounds of the bins. The point forecasts are the average of individual forecasters' point forecasts. Core PCE inflation histograms are available from the SPF since 2007.

Figure 12: SPF-implied mean and median forecasts for unemployment rate



(a) Current year

(b) Next year

(c) Two years ahead

(d) Three years ahead

Note: SPF-consistent median is the midpoint of the bin containing the median. Upper and lower bounds for SPF-consistent means are computed from lower and upper bounds of the bins. The point forecasts are the average of individual forecasters' point forecasts. Unemployment rate histograms are available from the SPF since 2009.