CONSTRUCTING FAN CHARTS FROM THE RAGGED EDGE OF SPF FORECASTS

2024

BANCO DE ESPAÑA
Eurosistema

Documentos de Trabajo
N.º 2429

Todd E. Clark, Gergely Ganics and Elmar Mertens

CONSTRUCTING FAN CHARTS FROM THE RAGGED EDGE OF SPF FORECASTS

CONSTRUCTING FAN CHARTS FROM THE RAGGED EDGE OF SPF FORECASTS [*]

Todd E. Clark

FEDERAL RESERVE BANK OF CLEVELAND

Gergely Ganics

BANCO DE ESPAÑA

Elmar Mertens

DEUTSCHE BUNDESBANK

(*) Parts of this paper were circulated earlier under the title "Constructing the Term Structure of Uncertainty from the Ragged Edge of SPF Forecasts" and were completed while G. Ganics was with the Central Bank of Hungary. The authors gratefully acknowledge helpful suggestions and comments received from Cem Çakmakı, Refet Gürkaynak, Ed Knotek, Mike McCracken, James Mitchell, Tom Stark and workshop or conference participants at the NBER Summer Institute 2022, 2022 IAAE conference, 2022 Dolomiti Macro Meeting, 2022 NBER-NSF SBIES conference, 2022 NBER-NSF Time Series conference, 2022 Conference on Real-Time Data Analysis, Methods and Applications, Heidelberg University and Deutsche Bundesbank. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Cleveland, the Federal Reserve System, the Banco de España, the Deutsche Bundesbank or the Eurosystem. Replication files are available at https://github.com/elmarmertens/ClarkGanicsMertensSPFfancharts.

Documentos de Trabajo. N.º 2429

September 2024

https://doi.org/10.53479/37597

ISSN: 1579-8666 (on line)

Abstract

We develop models that take point forecasts from the Survey of Professional Forecasters (SPF) as inputs and produce estimates of survey-consistent term structures of expectations and uncertainty at arbitrary forecast horizons. Our models combine fixed-horizon and fixed-event forecasts, accommodating time-varying horizons and availability of survey data, as well as potential inefficiencies in survey forecasts. The estimated term structures of SPF-consistent expectations are comparable in quality to the published, widely used short-horizon forecasts. Our estimates of time-varying forecast uncertainty reflect historical variations in realised errors of SPF point forecasts and generate fan charts with reliable coverage rates.

Keywords: term structure of expectations, uncertainty, survey forecasts, fan charts.

JEL classification: E37, C53.

Resumen

El presente trabajo desarrolla modelos que utilizan como insumos los pronósticos puntuales de la encuesta Survey of Professional Forecasters (SPF) para hacer estimaciones temporales de las expectativas y la incertidumbre coherentes con la encuesta, para horizontes de proyecciones arbitrarias. Nuestros modelos combinan previsiones de horizonte fijo y evento fijo, acomodando horizontes de tiempo variables y disponibilidad de datos de la encuesta, así como posibles ineficiencias en los pronósticos de esta. Las estructuras temporales estimadas de expectativas congruentes con la SPF son comparables, en calidad, a los pronósticos de corto plazo publicados y ampliamente utilizados. Estas estimaciones de la incertidumbre, de los pronósticos variables en el tiempo, reflejan variaciones históricas de los errores detectados en los pronósticos puntuales de la SPF, y generan *fan charts* con tasas de cobertura fiables.

Palabras clave: estructura temporal de expectativas, incertidumbre, pronósticos de encuestas, *fan charts.*

Códigos JEL: E37, C53.

1 Introduction

Both economic policymaking and forecasting research commonly use the macroeconomic projections of professional forecasters. Such forecasts with long histories include the (US) Survey of Professional Forecasters (SPF), Blue Chip Consensus, and Consensus Economics, as well as Federal Reserve forecasts published in the Tealbook or Greenbook and the Federal Open Market Committee's (FOMC) Summary of Economic Projections (SEP). Typically, the availability of such forecasts is uneven across forecast horizons. For example, the SPF includes both (1) fixed-horizon quarterly point forecasts, at shorter horizons, and (2) fixed-event annual forecasts (i.e., forecasts of a given event, in this case an economic outcome in a specific calendar year, made at different points in time before the event (Nordhaus, 1987)), available at longer horizons. Forecasters, policy analysts, and researchers regularly consider forecast uncertainty; many central banks (e.g., the Federal Reserve and the European Central Bank) publish estimates of forecast uncertainty.

In this paper, we develop models that take published survey point forecasts as inputs and produce estimates of a longer, more complete term structure (across horizons) of survey-consistent point forecasts, along with a term structure of forecast uncertainty. Our applications use the SPF, and our methods can exactly replicate its quarterly forecasts at short horizons while interpolating through the longer-horizon forecasts from the SPF's annual predictions. Our estimates of forecast uncertainty reflect dispersion in past errors from the SPF's point forecasts, with time-varying uncertainty captured through stochastic volatility (SV).[1]

Matching the term structure of SPF forecasts with predictions from a simple time series model of the outcome variable typically requires measurement error, but that is not the case in our models.[2] Specifically, we assume that, beyond a given horizon, forecasts are equal to a drifting mean, while forecasts up to that horizon follow a flexible moving-average structure that can match arbitrary term structures of expectations. Nevertheless, to avoid excessively volatile imputation that can arise from small inconsistencies between quarterly and annual SPF predictions, we model annual forecasts with fat-tailed measurement errors that are a priori close to zero.

Our approach extends Clark et al. (2020) (hereafter, CMM) in a few directions. First, we model an SPF-consistent term structure of quarterly expectations that extends arbitrarily far beyond the

[1] As surveyed in Clark and Mertens (2023), many studies have found that allowing for time-varying conditional variances improves the fit and forecasting performance of time series models.

[2] For examples of matching survey data with relatively coarse time series models, see Aruoba (2020), Crump et al. (2023), Kozicki and Tinsley (2012), and Patton and Timmermann (2011).

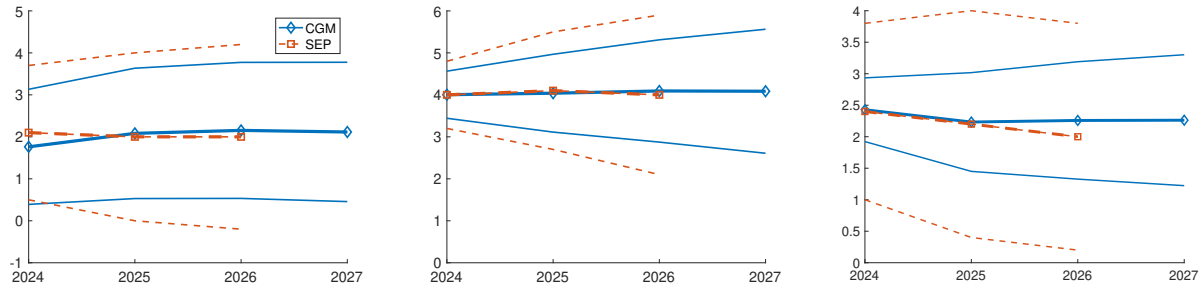
four quarters covered in CMM. Second, our inference on expectations beyond the near term is informed by the SPF's fixed-event (annual) forecasts that extend up to three years ahead. Doing so requires handling the so-called ragged edge of the SPF, whose quarterly publication leads to systematic variations in the fixed-event horizons of annual forecasts. Third, we develop a generalized model that allows for bias (both conditional and unconditional) in the SPF. Time-varying bias in SPF forecasts and other departures from full rationality have been highlighted in studies including Bianchi et al. (2022), Coibion and Gorodnichenko (2015), and Farmer et al. (2024). We provide a rich and flexible model of the joint dynamics between outcomes and expectational data provided by the SPF.

Practically speaking, what might one do with a more complete term structure of expectations and uncertainty? As one application, our approach can be used to produce SPF-consistent fan charts patterned after those from the FOMC's SEP, which report calendar-year forecasts of Q4/Q4 growth rates of GDP and prices and the Q4 level of the unemployment rate. The width of the uncertainty bands in these SEP fan charts is two times the historical root mean square errors (RMSEs) obtained using a few different forecasts computed over 20-year rolling windows, as developed in Reifschneider and Tulip (2019). To illustrate, from a forecast origin of 2024Q1 Figure 1 provides SEP-style fan charts with SPF-consistent calendar-year forecasts of GDP growth, CPI inflation, and the unemployment rate. The fan charts are generated from our estimates of predictive densities that are conditioned on the SPF. For the figure, we then transform draws from these predictive densities into calendar-year predictions for 2024-2027 that match the data conventions of the SEP. The fan charts include a point forecast and a 68 percent forecast uncertainty band. For comparison, the charts also include the corresponding projections (red dotted lines) from the March 2024 SEP. The SPF-consistent point forecasts show GDP growth remaining near 2 percent for the next few years, with unemployment near 4 percent. The SPF forecasts have CPI inflation slowing from 2024 to 2025 and changing little in 2026 and 2027. The SEP's point forecasts are similar, although the SEP includes a further reduction in inflation in 2026. But, in terms of forecast uncertainty, the SPF-consistent and SEP fan charts differ more materially: While the 68 percent forecast interval for the SPF is reasonably wide, the SEP's is notably wider (in keeping with a broader historical pattern that we document later in the paper).

In the body of this paper, we flesh out the approach that underlies this example. We first show that our baseline model yields SPF-consistent quarterly real-time forecasts of GDP growth, unem-

Figure 1: Fan charts and SEP (2024Q1)

(a) Real GDP growth (b) Unemployment rate (c) Inflation



Notes: Model-based annual predictions, with 68% predictive intervals, from our model ("CGM") and the FOMC's SEP per 2024Q1. The predictions are calendar-year forecasts of Q4/Q4 growth rates of GDP and prices and the Q4 level of the unemployment rate. Since the forecast origin is early in the year, the SEP extends only two years out. For inflation, Panel (c) compares model-based densities for CPI against the SEP's forecasts for PCE inflation. The version of the model used to generate these figures treats SPF forecasts as martingales.

ployment, and inflation that match the published fixed-horizon quarterly forecasts from the SPF and interpolate through the annual fixed-event point forecasts. The model's estimates of forecast uncertainty vary over time, temporarily rising around recessions, and generally rise with the forecast horizon. Our second set of results addresses the quality of the extended SPF forecasts obtained from our models (both the baseline specification that treats SPF forecasts as rational and the generalized model that allows departures from full rationality), applying a variety of metrics to out-of-sample forecasts. Our estimates indicate that the generalized model captures reasonably well the empirical extent of non-rationality emphasized by Coibion and Gorodnichenko (2015) and subsequent studies. Nonetheless, the SPF forecasts obtained from our restricted baseline model and the generalized version perform comparably in various metrics of forecast quality. More specifically, our estimated forecasts for the SPF at forecast origin t are efficient predictions of the forecasts published in the SPF at origin $t+1$. Out to forecast horizons as long as 16 quarters ahead, quarterly forecasts from the baseline model and its generalization are similar in point and density accuracy. The two models also perform comparably in unconditional coverage rates, yielding forecast confidence bands that are reasonably accurate, but not perfect. Tests of the uniformity of probability integral transforms indicate that our estimated predictive densities of SPF forecasts are correctly calibrated. Together, we take these results as evidence that the quality of our SPF-consistent estimated forecasts is comparable to the quality of the published short-horizon forecasts, which are widely seen as useful in research and practice.

The paper proceeds as follows. Section 2 provides an overview of the related literature on survey forecasts not covered above or in subsequent sections. Section 3 describes the SPF forecasts and data used. Section 4 presents our models and briefly describes estimation. Section 5 provides results. Section 6 concludes. A supplementary online appendix provides additional results.

2 Related Literature

Over the course of our analysis, we will consider the rationality, calibration, and accuracy of the term structure of SPF-consistent forecasts. A long literature has examined the same properties of published fixed-event and fixed-horizon survey forecasts. More specifically, some studies have examined whether professional forecasts display properties consistent with optimal (typically, under quadratic loss) forecasts and rational expectations. In one example, Patton and Timmermann (2012) develop new rationality tests based on rationality restrictions taking the form of bounds on second moments of the data across forecast horizons and apply them to forecasts from the Federal Reserve's Greenbook. As regards accuracy, focusing on inflation forecasts, Ang et al. (2007) find survey forecasts hard to beat by a battery of forecasting methods.

While we focus on professional forecasts in and of themselves, another long literature has sought to use professional forecasts to improve forecasts from time series models. Faust and Wright (2009) use professional forecasts as jumping-off points for models to improve the accuracy of forecasts from time series models. Wright (2013) shows that the forecast accuracy of a Bayesian vector autoregression (VAR) can be improved by using long-run survey forecasts as priors on the model's steady states. Banbura et al. (2021), Ganics and Odendahl (2021), and Krüger et al. (2017) improve forecasts from Bayesian VARs through entropic tilting toward survey-based nowcasts or forecasts. Frey and Mokinski (2016) instead add survey-based nowcasts as endogenous variables in Bayesian VARs, using priors so that the dynamics of the survey forecasts inform the parameter estimates of the dynamics of the actual data. In a similar vein, Doh and Smith (2022) develop priors that align a VAR's (a priori) forecasts with survey predictions.

While our approach uses SPF forecasts as the only inputs to obtain a term structure of forecasts and uncertainty, a number of other studies have estimated term structures of survey forecasts by using a time series model for y_t that generates forecasts $E_t y_{t+h}$ and assuming that observed survey forecasts are equal to model-implied forecasts plus a measurement error, $F_t y_{t+h} =$

$E_t y_{t+h}$ + noise$_{t+h}$. This approach is employed by studies such as Aruoba (2020), Crump et al. (2023), Grishchenko et al. (2019), Kozicki and Tinsley (2012), and Mertens and Nason (2020). Crucially, in applications such as these, the use of measurement error is necessary, as it serves as fitting error in matching term structures of a tightly parameterized model to the SPF data. In our application, we attach measurement error to annual SPF forecasts out of concern about (small) inconsistencies in the SPF data; while our model can perfectly replicate the SPF without measurement error, its presence avoids excessively volatile imputations.

Potential drawbacks of this approach that features a time series model of the data include the attribution of part of the observed term structure of survey expectations to measurement error and the imposition of a (typically low-order) time series model on the term structure of "true" expectations ($E_t y_{t+h}$). For some applications, such an approach might provide a potentially beneficial form of shrinkage. In addition, the approach can be used to pool surveys from different sources to extract a common set of underlying expectations. However, the measurement error approach comes at the cost of discarding part of the survey respondents' judgment and broader modeling. Instead, as explained below, one of our models aims to fit model-implied forecasts to survey forecasts, $F_t y_{t+h} = E_t y_{t+h}$, and our more general model sees differences between forecasts from the SPF and the model as bias (and thus predictable forecast errors, instead of measurement error).

Complementing our work on a rich term structure of forecast uncertainty several years into the future, many other studies have used the fixed-event forecasts of professionals to examine forecast uncertainty and its variation over time. For example, with data on fixed-event forecasts from Consensus Economics, Patton and Timmermann (2011) use an unobserved components model to examine the predictability of growth and inflation across different forecast horizons and measure average forecast uncertainty by mean square forecast errors.

3 Data

Because the availability of forecasts in the SPF informs aspects of our model, we detail the data in this section before taking up the model in Section 4. We examine quarterly and annual forecasts from the SPF for a basic set of major macroeconomic aggregates commonly included in research on the forecasting performance of models such as VARs or dynamic stochastic general equilibrium models: real GDP growth (RGDP), the unemployment rate (UNRATE), and inflation in the

GDP price index (PGDP) and consumer price index (CPI).[3] The SPF forecasts are widely studied, publicly available, and the longest time series of forecasts for a range of variables.

We obtained the SPF forecasts from the Federal Reserve Bank of Philadelphia's Real-Time Data Research Center. In all cases, we form the point forecasts using the average over all SPF responses. Reflecting the data available, our estimation samples start with 1969Q1 for GDP growth, unemployment, and GDP inflation and 1981Q4 for CPI inflation; the last forecast origin for which we evaluate out-of-sample forecasts is 2023Q4, and we also discuss fan charts conditioned on the 2024Q1 SPF. At each forecast origin, the available fixed-horizon point forecasts typically span five quarters, from the current quarter through the next four quarters. Since 1981Q3, the SPF has included fixed-event point forecasts for the current and next calendar year. In 2005Q3, the forecast horizon for CPI inflation was extended to include annual forecasts for one additional year, i.e., covering up to two calendar years ahead, and in 2009Q2, the forecast horizon was similarly extended for GDP growth and unemployment to extend up to three years ahead.

Given the fixed-event nature of the SPF's calendar-year forecasts, the number of quarters until the end of the longest-horizon forecast varies over the course of a year. For example, the 2021Q4 SPF included fixed-event annual forecasts of growth and unemployment for the current year and the next three, so that the last annual forecast extends 12 quarters ahead (the annual forecast reported for 2024 includes 2024Q4, 12 quarters beyond the 2021Q4 forecast origin). In the 2022Q1 SPF, the last annual forecast for 2025 includes 2025Q4, 15 quarters beyond the forecast origin. These variations in the SPF's effective forecast horizon are also known as the "ragged edge." Our methods allow us to consistently construct fixed-horizon term structures that, at every quarterly forecast origin, extend arbitrarily far beyond the ragged edge.

The SPF uses certain conventions in its forecasts that we incorporate in the measurement specification of our model, as described further in Section 4. In terms of the SPF data we use, predictions for the unemployment rate are expressed directly as quarterly and annual-average levels (depending on the forecast horizon), which can be mapped directly into our model. Similarly, for CPI inflation, the forecasts are reported as percentage changes in quarterly and Q4/Q4 index levels that are used as such in our model. However, for GDP and its price index, the SPF data files provide forecasts in levels, which we convert into growth rates. Specifically, depending on the forecast

[3] We use "GDP" and "GDP price index" to refer to output and price series, even though, in our real-time data, the measures are based on GNP and a fixed-weight deflator for some of the sample.

horizon, point forecasts pertain to quarterly or annual-average levels, which we convert to growth rates based on information included in the survey. For quarterly forecast targets, we use the lagged quarterly level as the basis. To obtain the next-year forecast of annual-average growth, we use the SPF's predictions for the current year as base values (and analogously for the forecasts two and three years ahead). For each forecast horizon, we calculate growth rates for GDP and the GDP price index as differences in log levels, as our model uses log-linearized expressions for quarterly and annual-average growth in these variables that are detailed in Section 4.

To estimate our model, we also need measures of the outcomes of the variables. In the case of GDP growth and GDP inflation, for which data can be substantially revised over time, we obtain real-time measures for quarter $t-1$ data as these data were publicly available in quarter t from the quarterly files of real-time data in the Philadelphia Fed's Real-Time Data Set for Macroeconomists (RTDSM). For forecast evaluation, we measure the outcomes of GDP growth and GDP inflation with the RTDSM vintage published two quarters after the outcome date.[4] Because revisions to quarterly data are relatively small for the unemployment rate and CPI inflation, we simply use the historical time series available in the St. Louis Fed's FRED database to measure the outcomes and corresponding forecast errors for these variables.

4 Model and Estimation

In broad terms, our models can be seen as multivariate unobserved components models. We design the state space specifications to match arbitrary term structures of expectations, with application to the SPF in this paper. In all cases, we specify and estimate the models on a variable-by-variable basis (i.e., separately for GDP growth, unemployment, and each inflation measure).

This section begins by spelling out the general setup of our models, including our underlying assumptions. We then proceed with detailing the transition equations of our state space models, starting with a specification that treats forecast updates as martingale differences and then proceeding to a more general specification that allows for biases and persistence in forecast errors and expectations updates. After laying out transition dynamics, we present the measurement equation included in both models, followed by our specification of the innovation components of the tran-

[4] That is, we use the quarterly vintage in $t+h+2$ to evaluate forecasts for $t+h$ made in t; this is the second estimate available in the RTDSM's vintages.

sition equations. Throughout, we use $\boldsymbol{Y_t}$ to denote the state vector (partially latent) of quarterly forecasts and $\boldsymbol{Z_t}$ to denote the measurement vector of the available SPF point forecasts.

4.1 General setup

Throughout, y_t refers to either (1) the annualized quarterly log growth rate of GDP or its price index, (2) a quarterly average level of the unemployment rate, or (3) the annualized quarterly (simple) percent change in the CPI. We denote SPF-consistent forecasts of the scalar outcome y_{t+h} made in period t by $F_t y_{t+h}$ and rational (or unbiased) forecasts as $E_t y_{t+h}$. "SPF-consistent" means that these forecasts are either directly observed from the SPF or values imputed from our model under the assumption that those values represent the (average) SPF response. The vector $\boldsymbol{Y_t}$ — the "SPF-consistent" term structure of expectations — includes the lagged realized value, y_{t-1}, and forecasts from horizons $h = 0, \ldots, H$:

$$\boldsymbol{Y_t} \equiv \left[y_{t-1},\ F_t y_t,\ F_t y_{t+1}, \ldots, F_t y_{t+h}, \ldots, F_t y_{t+H} \right]' . \tag{1}$$

Note that, with $y_{t-1} = F_t y_{t-1}$, we have $F_t \boldsymbol{Y_t} = \boldsymbol{Y_t}$. To derive the transition equation that captures the dynamics of $\boldsymbol{Y_t}$, we begin with definitional equations for the (lagged) nowcast error, e_{t-1}, as well as SPF forecast updates, $(F_t - F_{t-1})y_{t+h}$:[5]

$$y_{t-1} = F_{t-1} y_{t-1} + e_{t-1}\,, \tag{2}$$

$$F_t y_{t+h} = F_{t-1} y_{t+h} + (F_t - F_{t-1}) y_{t+h}, \quad \forall\ h \geq 0\,. \tag{3}$$

In each of these equations, the term on the left is an element of $\boldsymbol{Y_t}$, whereas the first term on the right is an element of $\boldsymbol{Y_{t-1}}$, which suggests a strategy to derive a recursion for the dynamics of $\boldsymbol{Y_t}$.

We also make the following three assumptions.

1. The term structure of SPF-consistent forecasts is flat beyond horizon H, with y_t^* denoting the endpoint of the term structure of SPF forecasts:

$$y_t^* \equiv F_t y_{t+H+1} = F_t y_{t+H+j} \quad \forall\ j > 0\,. \tag{4}$$

[5] Some previous studies have also made use of expectational updates, for different purposes. For example, Patton and Timmermann (2012) write a short-horizon forecast as a sum of a long-horizon forecast and forecast revisions, and use it as the basis of an optimal revision regression to test forecast optimality. Coibion and Gorodnichenko (2015) map out the implications of different theories of imperfect information for serial correlation in forecast updates.

2. The endpoint y_t^* is an unbiased (rational) long-run forecast of y_t, corresponding to the Beveridge-Nelson trend in y_t and following a random walk process:

$$y_t^* \equiv \lim_{j\to\infty} E_t y_{t+H+j} = y_{t-1}^* + w_t^* \;\Rightarrow\quad E_{t-1} w_t^* = 0. \tag{5}$$

3. The endpoint of the SPF term structure, y_t^*, is the common trend in all elements of $\boldsymbol{Y_t}$ and deviations from trend, also called "gaps" and denoted by $\tilde{\boldsymbol{Y}}_{\boldsymbol{t}}$, are (mean) stationary:

$$\boldsymbol{Y_t} = \tilde{\boldsymbol{Y}}_{\boldsymbol{t}} + \mathbf{1} y_t^*\,, \quad \text{with} \quad \lim_{h\to\infty} E_t \tilde{\boldsymbol{Y}}_{\boldsymbol{t+h}} = \bar{\boldsymbol{Y}}\,, \tag{6}$$

where the mean vector for the gaps, $\bar{\boldsymbol{Y}}$, captures unconditional bias in the SPF at different horizons (detailed further in the supplementary online appendix), and $\mathbf{1}$ denotes a vector of ones. With this assumption, we are treating SPF forecasts as being at least weakly rational, so that their forecast errors are stationary, which implies that SPF forecasts at different horizons are cointegrated among each other and with the outcome variable (Mertens, 2016).[6]

Together with equation (3), these assumptions imply the following recursive law of motion for the vector of gaps $\tilde{\boldsymbol{Y}}_{\boldsymbol{t}}$ (where $\boldsymbol{I}$ is the identity matrix):

$$\tilde{\boldsymbol{Y}}_{\boldsymbol{t}} = \left(\boldsymbol{I} - \tilde{\boldsymbol{\Psi}}\right)\bar{\boldsymbol{Y}} + \tilde{\boldsymbol{\Psi}}\tilde{\boldsymbol{Y}}_{\boldsymbol{t-1}} + \tilde{\boldsymbol{\eta}}_{\boldsymbol{t}}\,, \tag{7}$$

$$\text{with}\quad \tilde{\boldsymbol{\Psi}} = \begin{bmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & 0 & \ddots & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & \dots & \dots & \dots & \dots & 0 \end{bmatrix}, \qquad \tilde{\boldsymbol{\eta}}_{\boldsymbol{t}} \equiv \begin{bmatrix} e_{t-1} \\ (F_t - F_{t-1})y_t \\ (F_t - F_{t-1})y_{t+1} \\ \vdots \\ (F_t - F_{t-1})y_{t+h} \\ \vdots \\ (F_t - F_{t-1})y_{t+H} \end{bmatrix} - \mathbf{1} w_t^* - \bar{\boldsymbol{\eta}}\,, \tag{8}$$

[6] Other studies that decompose SPF forecasts into (perceptions of) permanent and transitory components include Clements (2022) and Krane (2011).

and $\bar{\boldsymbol{\eta}} \equiv \left(\boldsymbol{I} - \tilde{\boldsymbol{\Psi}}\right) \bar{\boldsymbol{Y}}$. All eigenvalues of $\tilde{\boldsymbol{\Psi}}$ are zeros, so that $\tilde{\boldsymbol{\Psi}}$ is a stable matrix.[7] To close the description of the state equation, we need to model the dynamics of $\tilde{\boldsymbol{\eta}}_{\boldsymbol{t}}$ with particular attention to its persistence, i.e., $E_{t-1}\tilde{\boldsymbol{\eta}}_{\boldsymbol{t}}$, as well as its distribution, which we model via variations in $\mathrm{Var}_{t-1}\, \tilde{\boldsymbol{\eta}}_{\boldsymbol{t}}$. Below we explain the two different state equation specifications considered and the shock process.

The first assumption above also has implications for the setting of H when implementing our model. While the SPF provides forecasts up to three calendar years ahead, its longer-run forecasts are stated in terms of annual predictions that straddle several quarters. Thus, to match the observed SPF, our state vector does not have to extend all the way to the end of the third calendar year ($H = 15$). Instead, we can utilize our assumption that the term structure of SPF-consistent forecasts is flat beyond some horizon H and describe a three-year-ahead forecast with a state space extending only to, say, $H = 12$. For concreteness, consider the case of a forecast, collected in Q1, that should match the average of the four quarters of the third year ahead. With $H = 12$, only the last element of $\boldsymbol{Y}_{\boldsymbol{t}}$ points into the third year ahead, and forecasts for quarters two to four of the third year are set identical to the trend, with zero gaps.

In this spirit, we generally choose H such that when the forecast origin is in Q1, H points to the first quarter of the farthest annual horizon covered by the SPF forecasts. An exception is made for data covering only SPF forecasts up to the next year, where H is set to 5 (instead of 4), since the observed fixed-horizon SPF forecasts already extend through $h = 4$. As part of our out-of-sample forecast analysis, models are (re-)estimated over different sub-samples of data, and we adjust H accordingly. Specifically, we set $H = 5$ for data samples on GDP growth and unemployment prior to 2009Q2, as well as for CPI inflation prior to 2005Q4, and inflation in the GDP price index (entire sample). For estimates covering data that include SPF forecasts up to three years ahead (GDP growth and unemployment since 2009Q2) we set $H = 12$, and for CPI inflation data since 2005Q4, which include SPF forecasts up to two years ahead, we set $H = 8$. Note that, even though the state vector ends with $F_t y_{t+H}$ and H is no larger than 12, our endpoint assumption allows us to simulate forecast densities arbitrarily far ahead, and we report densities up to 16 quarters ahead for all variables throughout.

[7] An earlier working paper version of this paper derived an equivalent recursion for $\boldsymbol{Y}_{\boldsymbol{t}}$; reflecting the common trend in $\boldsymbol{Y}_{\boldsymbol{t}}$, this equivalent representation features a transition matrix with a single unit root and all other eigenvalues equal to zero.

4.2 Transition equation in the martingale case

Our first model assumes SPF forecasts are rational so that $F_t y_{t+h} = E_t y_{t+h} \; \forall \, h$. It follows that forecasts are unbiased, $\bar{\boldsymbol{Y}} = \mathbf{0}$, and that forecast updates are martingale difference sequences (MDS): $E_{t-1}\tilde{\boldsymbol{\eta}}_{\boldsymbol{t}} = 0$ and $E_{t-1} w_t^* = 0$. With these restrictions, the mean dynamics of the gap vector $\tilde{\boldsymbol{Y}}_{\boldsymbol{t}}$ are fully specified by equation (7), with the restriction $\bar{\boldsymbol{Y}} = \mathbf{0}$ imposed.

To describe the density of $\tilde{\boldsymbol{Y}}_{\boldsymbol{t}}$, we use a block-SV structure with fat-tailed shocks to $\tilde{\boldsymbol{\eta}}_{\boldsymbol{t}}$ and w_t^* that will be described further below. Although this MDS-based model is written with a state vector $\boldsymbol{Y}_{\boldsymbol{t}}$ containing SPF forecasts, we should emphasize that, in this specification, we are not actually attributing a specific time series model to the evolution of SPF forecasts; we are taking the observed fixed-horizon and fixed-event forecasts as given and using a time series process to interpolate missing fixed-horizon forecasts out to H steps ahead. With the MDS specification, all we need is the historical evolution of the SPF up to the forecast origin t and previous nowcast errors in order to estimate SPF-consistent point forecasts and uncertainty bands around the forecasts; we do not need a time series model of y_t.

4.3 Transition equation of the VAR model

The MDS model places tight restrictions on the evolution of forecast updates by assuming that SPF forecasts are rational. Our second, more general specification drops that assumption. This model allows for (a) unconditional bias in the form of a non-zero intercept vector $\bar{\boldsymbol{Y}}$ in the transition dynamics of equation (7) and (b) conditional bias by modeling the updates in (detrended) forecasts, $\tilde{\boldsymbol{\eta}}_{\boldsymbol{t}}$, as a VAR(1):[8]

$$\tilde{\boldsymbol{\eta}}_{\boldsymbol{t}} = \tilde{\boldsymbol{\Pi}}\, \tilde{\boldsymbol{\eta}}_{\boldsymbol{t-1}} + \tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{t}}\,, \qquad \text{with} \quad \tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{t}} \sim \mathcal{N}(\mathbf{0}, \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{t}}), \tag{9}$$

and details for $\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{t}}$ to be described further below. In keeping with the common-trend assumption, the transition matrix $\tilde{\boldsymbol{\Pi}}$ is assumed to be stable. It follows that the state equation for the vector of gaps $\tilde{\boldsymbol{Y}}_{\boldsymbol{t}}$ takes a VAR(2) form:

$$\tilde{\boldsymbol{Y}}_{\boldsymbol{t}} = \left(\boldsymbol{I} - \tilde{\boldsymbol{\Psi}}\right)\left(\boldsymbol{I} - \tilde{\boldsymbol{\Pi}}\right)\bar{\boldsymbol{Y}} + \left(\tilde{\boldsymbol{\Psi}} + \tilde{\boldsymbol{\Pi}}\right)\tilde{\boldsymbol{Y}}_{\boldsymbol{t-1}} - \left(\tilde{\boldsymbol{\Psi}}\,\tilde{\boldsymbol{\Pi}}\right)\tilde{\boldsymbol{Y}}_{\boldsymbol{t-2}} + \tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{t}}\,. \tag{10}$$

[8] Our choice of a lag order of 1 is consistent with specifications and related evidence in Coibion and Gorodnichenko (2015) and CMM. In addition, given the limited amount of SPF bias reported in the prior literature (Section 2) , we view a low lag order as appropriate to limit the scope for overfitting.

The MDS model is nested in this specification, under the restrictions $\bar{\boldsymbol{Y}} = \mathbf{0}$ and $\tilde{\boldsymbol{\Pi}} = \mathbf{0}$. Absent these restrictions, the model allows for unconditional and conditional bias in the SPF. With that, the VAR model still captures in $\boldsymbol{Y_t}$ an SPF-consistent term structure of expectations that matches the observed SPF data (exactly for quarterly SPF forecasts, and up to measurement error for the annual forecasts). But predictions for the outcome variables generated by the VAR model will generally differ from $\boldsymbol{Y_t}$, whereas point forecasts generated by the MDS specification will not.

4.4 Measurement equation

To explain the measurement equation of our state space models, we define:

$$\bar{y}_t = 1/4 \cdot \sum_{j=0}^{3} y_{t-j}, \tag{11}$$

$$\text{and} \quad \hat{y}_t \equiv 100 \times \log\left(\frac{I_t + I_{t-1} + I_{t-2} + I_{t-3}}{I_{t-4} + I_{t-5} + I_{t-6} + I_{t-7}}\right), \tag{12}$$

$$\approx 1/16 \cdot \left(y_t + 2 \cdot y_{t-1} + 3 \cdot y_{t-2} + 4 \cdot y_{t-3} + 3 \cdot y_{t-4} + 2 \cdot y_{t-5} + y_{t-6}\right), \tag{13}$$

where I_t denotes the quarterly level of GDP or its price index.

With these definitions of multi-period forecast targets, we can cover the different types of annual forecasts published in the SPF: When t corresponds to a date in Q4, $\bar{y}_t$ represents the annual average level of the unemployment rate as well as Q4/Q4 percent changes in the CPI, and $\hat{y}_t$ captures percent changes in annual average levels of GDP and its price index. As discussed in the supplementary online appendix, the formula for $\hat{y}_t$ in (13) represents a log-linear approximation to the growth rate in average levels of calendar years ending at t and $t-4$ in (12).[9] The measurement equations of our models, detailed below, use the approximation in (13) to relate annual SPF forecasts for GDP growth and GDP price inflation to forecasts (or lagged realizations) of y_t.[10]

We denote survey forecasts collected at forecast origin t for targets y_{t+h}, $\bar{y}_{t+h}$, and $\hat{y}_{t+h}$, by $F_t y_{t+h}$, $F_t \bar{y}_{t+h}$, and $F_t \hat{y}_{t+h}$, respectively. At time t, the SPF provides observations of $F_t y_{t+h}$, $F_t \bar{y}_{t+h}$, and/or $F_t \hat{y}_{t+h}$, for different (but separate) values of $h \geq 0$. The measurement vector $\boldsymbol{Z_t}$ contains the available observations from the SPF as well as a reading of the last realized value,

[9] Mariano and Murasawa (2003) first developed the approximation in a nowcasting setting. Models of multi-period survey forecasts that rely on this (or similar) approximations include Aruoba (2020), Crump et al. (2023), and Patton and Timmermann (2012).

[10] For CPI inflation, the mapping from y_t to $\bar{y}_t$ holds only approximately as well. For CPI, our definition of y_t as the simple percent change follows the SPF convention for its quarterly fixed-horizon forecasts, whereas the annual forecasts are solicited in terms of changes in Q4/Q4 levels.

y_{t-1}, that is available to SPF respondents at time t. In addition to the SPF's quarterly fixed-horizons forecasts, $F_t y_{t+h}$ for $h = 0, 1, \ldots, 4$, we include in $\boldsymbol{Z_t}$ the available readings of fixed-event forecasts for the next year and beyond.[11] For example, with t = 2024Q1, the measurement vector for GDP growth is $\boldsymbol{Z_t} = \left[y_{t-1},\ F_t y_t,\ F_t y_{t+1}, \ldots, F_t y_{t+4}, F_t \hat{y}_{t+7}, F_t \hat{y}_{t+11}, F_t \hat{y}_{t+15} \right]'$. For t = 2024Q2, the annual forecast entries become $F_t \hat{y}_{t+6}$, $F_t \hat{y}_{t+10}$, and $F_t \hat{y}_{t+14}$. The corresponding measurement vectors for the unemployment rate differ only by using annual forecasts expressed in terms of $\bar{y}_t$ instead of $\hat{y}_t$. Further details are described in the supplementary online appendix.

Our framework can match a term structure of expectations to an arbitrary set of observable SPF data. It could do so without a need for measurement noise. However, while not essential, attaching measurement error to at least some SPF observations may be warranted and helpful. It is possible that the reported readings of annual forecasts contain some small discrepancies, due, for example, to the computation of growth rates using GDP forecasts reported in levels with rounding. Consider, for example, the case of SPFs in the third quarter of each year that include an annual forecast for the next-year SPF. Given the handful of quarterly forecasts published in the SPF, to the model the only "free" quarterly forecast for matching the annual forecast is the projection for the fourth quarter of next year. In the applications to GDP growth and inflation in the GDP deflator, that quarterly forecast has a weight of 1/16 in the annual approximation in (13), which in turn means that potential discrepancies get scaled up by a factor of 16 in the imputation of the fourth quarter forecast for next year. As detailed in the supplementary online appendix, these mechanics can (and do) lead to excessively volatile imputations. Accordingly, in our implementation, the measurement equations of our MDS and VAR models add measurement error to the annual forecasts, with separate measurement error processes for each quarter of the year. Formally, we partition the measurement vector into two parts, and assume that only the first part is observed without error:

$$\boldsymbol{Z_t} = \begin{bmatrix} \boldsymbol{Z_{q,t}} \\ \boldsymbol{Z_{a,t}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{C_{q,t}} \\ \boldsymbol{C_{a,t}} \end{bmatrix} \boldsymbol{Y_t} + \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{n_t} \end{bmatrix}, \quad n_{i,t} \sim \mathcal{N}(0, \sigma^2_{i,t}), \tag{14}$$

where $n_{i,t}$ is the ith element of $\boldsymbol{n_t}$ and denotes the measurement error for the ith element of $\boldsymbol{Z_{a,t}}$. The loading matrices $\boldsymbol{C_{q,t}}$ and $\boldsymbol{C_{a,t}}$ vary over time as a function of the available measurements, and contain fixed coefficients of 0, 1, or fractional values as indicated in the measurement definitions

[11] Since the SPF provides fixed-horizon forecasts for up to four quarters ahead, we disregard all current-year (fixed-event) forecasts. In Q4, there is also complete overlap between the SPF's annual forecast for next year and the observed fixed-horizon forecast, and we treat these observations with measurement error.

(11) and (13). The measurement errors are independent across elements of $\boldsymbol{Z}_{a,t}$, modeled as conditionally Gaussian with zero mean and time-varying variance, as detailed below.

4.5 Shock distributions

This section describes our modeling of the shocks in the MDS and VAR specifications, starting with the distribution of $\tilde{\boldsymbol{\varepsilon}}_t$, the shocks to the gaps vector of our state space model. (In the MDS case, $\tilde{\boldsymbol{\eta}}_t = \tilde{\boldsymbol{\varepsilon}}_t$.) Throughout, we model these shocks as conditionally Gaussian: $\tilde{\boldsymbol{\varepsilon}}_t \sim \mathcal{N}(\mathbf{0}, \tilde{\boldsymbol{\Sigma}}_t)$.

We model variations in uncertainty through a time-varying $\tilde{\boldsymbol{\Sigma}}_t$. To do so, we build on the specification of Chan (2020) in which a scalar factor affects the entire Gaussian shock vector, imparting perfectly correlated variations in uncertainty to all shocks. We extend his framework by splitting the shock vector into multiple blocks, each with its own common volatility factor. In our case, the shock vector reflects shocks to forecasts at different horizons, and the block structure allows us to model distinct variations in uncertainty at different forecast horizons. Specifically, we adopt a structure with two blocks, such that $\tilde{\boldsymbol{\varepsilon}}_{\mathbf{1},t}$ consists of shocks to the lagged realized outcome and forecasts up to $h = 3$ and $\tilde{\boldsymbol{\varepsilon}}_{\mathbf{2},t}$ consists of shocks to forecasts for $h = 4, \ldots, H$. The cutoff at $h = 3$ between blocks reflects the desire for the first block to mostly capture forecasts for quarters within the current year, while the second block captures forecasts for quarters in the next year and beyond.[12] Specifically, we adopt the following block-SV structure:

$$\tilde{\boldsymbol{\varepsilon}}_t = \begin{bmatrix} \tilde{\boldsymbol{\varepsilon}}_{\mathbf{1},t} \\ \tilde{\boldsymbol{\varepsilon}}_{\mathbf{2},t} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} & \tilde{\boldsymbol{K}} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\varepsilon}^*_{\mathbf{1},t} \\ \boldsymbol{\varepsilon}^*_{\mathbf{2},t} \end{bmatrix}, \quad \text{with} \begin{bmatrix} \boldsymbol{\varepsilon}^*_{\mathbf{1},t} \\ \boldsymbol{\varepsilon}^*_{\mathbf{2},t} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \lambda_{1,t} \cdot \tilde{\boldsymbol{\Sigma}}_{\mathbf{11}} & \mathbf{0} \\ \mathbf{0} & \lambda_{2,t} \cdot \tilde{\boldsymbol{\Sigma}}_{\mathbf{22}} \end{bmatrix} \right), \quad (15)$$

in which $\tilde{\boldsymbol{K}}$ is a matrix (with dimension $4 \times (H-2)$) of coefficients to be estimated. This SV structure yields the following time-varying variance-covariance matrix of the cyclical shocks:

$$\tilde{\boldsymbol{\Sigma}}_t = \begin{bmatrix} \boldsymbol{I} & \tilde{\boldsymbol{K}} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \lambda_{1,t} \cdot \tilde{\boldsymbol{\Sigma}}_{\mathbf{11}} & \mathbf{0} \\ \mathbf{0} & \lambda_{2,t} \cdot \tilde{\boldsymbol{\Sigma}}_{\mathbf{22}} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \tilde{\boldsymbol{K}} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix}'. \quad (16)$$

This block-SV representation includes a deliberate upper-triangular structure to capture covariance between the shocks to the first and second blocks of shocks. Under this structure, the volatilities of

[12] When the forecast origin is in Q1, $h = 3$ points to the last quarter of the current year. For forecast origins later in the year, $h = 3$ points to the next year.

longer-horizon forecasts — latent states, not directly observed forecasts, for $h \geq 4$ — are driven by a single common, longer-horizon factor and not impacted by volatilities at shorter horizons. Volatilities of shorter-horizon forecasts — which are directly observed — are allowed to be impacted by both the common, shorter-horizon factor and the longer-horizon factor. We think of time-varying volatility as having medium-frequency business cycle drivers as well as some higher frequency drivers; the former impacts forecast volatility at all horizons, whereas the latter only impacts the volatilities of shorter-horizon forecasts.

It is commonly understood that the ordering of variables affects estimates of VARs with SV processes for each variable (see, e.g., Arias et al. (2023)). With our two-block specification, within each block, the common SV specification means that variable ordering does not matter. However, ordering has some impact through the block-triangular structure; changing to a lower-triangular rather than upper-triangular structure would change the model estimates. In this setting, though, for the reasons given above for the upper-triangular structure, this specification choice is less arbitrary than the simple ordering of variables in a conventional VAR with SV.

The scalar factors $\lambda_{1,t}$ and $\lambda_{2,t}$ impart time variation and fat tails to the shock vector $\tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{t}}$. Building on, among others, Carriero et al. (2022), Chan (2020), and Jacquier et al. (2004), we model these factors as the products of iid inverse-gamma draws and persistent stochastic volatility processes:

$$\lambda_{i,t} = \phi_{i,t} \cdot \tilde{\lambda}_{i,t}\,, \qquad \forall\, i = 1, 2\,, \tag{17}$$

$$\text{with} \quad \phi_{i,t} \sim \mathcal{IG}\left(\frac{\nu_i}{2}, \frac{\nu_i}{2}\right), \qquad \log \tilde{\boldsymbol{\lambda}}_{\boldsymbol{t}} \equiv \begin{bmatrix} \log \tilde{\lambda}_{1,t} \\ \log \tilde{\lambda}_{2,t} \end{bmatrix} = \begin{bmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{bmatrix} \log \tilde{\boldsymbol{\lambda}}_{\boldsymbol{t-1}} + \boldsymbol{\epsilon}_{\boldsymbol{t}}^{\boldsymbol{\lambda}}\,, \tag{18}$$

and $\boldsymbol{\epsilon}_{\boldsymbol{t}}^{\boldsymbol{\lambda}} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Phi})$. The iid inverse-gamma draws add fat tails in the form of a multivariate t distribution with ν_i degrees of freedom to each block. The vector SV process $\log \tilde{\boldsymbol{\lambda}}_{\boldsymbol{t}}$ has correlated shocks and is normalized to a mean of zero, obviating the need for normalizing assumptions on the constant-coefficient matrices $\tilde{\boldsymbol{\Sigma}}_{\mathbf{11}}$ and $\tilde{\boldsymbol{\Sigma}}_{\mathbf{22}}$.

To complete the models, we need to specify processes for the trend shock (w_t^*) and measurement error ($n_{i,t}$) for each annual forecast i. We use (independent) horseshoe specifications for these shocks. Carvalho et al. (2010) proposed the horseshoe for modeling sparse regressions, i.e., regressions with many regressors, many of which are expected to be irrelevant, with only a few

attracting substantial mass a posteriori. So while a horseshoe prior places considerable mass on coefficient values of zero, it also has particularly fat tails to generate (few) sizable coefficient estimates. Similarly, our application of the horseshoe to shocks reflects our presumption that most realizations of shocks are close to zero while some can instead be sizable. To allow the size of measurement errors to vary systematically within the year, we apply separate horseshoe processes to forecasts collected in different quarters of the year. The horseshoe has a conditionally Gaussian representation, with conditional mean 0 and a time-varying conditional variance, as indicated in equation (14) for the case of the noise shocks. Similarly, for the trend shocks, we have:

$$w_t^* \sim \mathcal{N}(0, \omega_t^2). \tag{19}$$

In modeling ω_t^2 and $\sigma_{i,t}^2$, the horseshoe representation employs global and local shrinkage (i.e., specific to each variable and each t, respectively) as detailed in the supplementary online appendix.

4.6 Estimation

We estimate the models using Bayesian Markov chain Monte Carlo (MCMC) methods — specifically, a Gibbs sampler. The model estimation is conditioned on joint data for observed realizations and SPF predictions for a given economic variable (like GDP growth), but estimated separately for different economic variables.

In the MDS specification, the objects to be estimated include the ρ_i and $\mathbf{\Phi}$ parameters of the SV processes, the constant innovation covariance matrix $\tilde{\mathbf{\Sigma}}$, the parameters of the horseshoe models for innovations to the long-run forecast and measurement errors, the time series of the latent volatility states in $\tilde{\boldsymbol{\lambda}}_{\boldsymbol{t}}$, and the time series of latent forecast states contained in $\boldsymbol{Y_t}$. We sample the parameters of the SV processes using a conventional Gaussian prior and conditional posterior for ρ_i and an inverse Wishart prior and conditional posterior for $\mathbf{\Phi}$. We sample $\tilde{\mathbf{\Sigma}}$ with an inverse Wishart prior and conditional posterior. We estimate the volatility states in $\tilde{\boldsymbol{\lambda}}_{\boldsymbol{t}}$ with the mixture approach of Kim et al. (1998). Sampling of the horseshoe components in shocks to trend and noise follows the MCMC scheme described in Makalic and Schmidt (2016). Estimation of the VAR requires additional steps in the Gibbs sampler to draw the coefficients $\tilde{\mathbf{\Pi}}$ from their multivariate normal conditional posterior, while accounting for the heteroskedasticity in the shock processes. Further details are provided in the supplementary online appendix.

Sampling the latent forecast states contained in $\boldsymbol{Y}_t$ involves more complexity. Since our model assumes no measurement error in at least some of its observables, we face an ill-defined posterior precision that cannot be directly handled by conventional precision-based samplers. To efficiently draw the latent forecast states in $\boldsymbol{Y}_t$, we build on a new precision-based sampler developed in Mertens (2023) with details described in the supplementary online appendix. We retain 3,000 draws after a burn-in sample of 3,000 initial draws. When simulating the model's predictive density we sample 100 paths of future realizations of stochastic volatility and other state variables for each MCMC draw, resulting in $S = 300,000$ predictive density draws.

5 Results

This section begins with results on the term structures of expectations and forecast uncertainty. We focus on SPF-consistent forecasts from our MDS model specification. We then examine various aspects of the quality of out-of-sample forecasts, from both our MDS and VAR specifications. In all of these results, we examine forecasts starting in 1990Q1. We conclude the section with some comparisons of our SPF-based forecasts against those of the FOMC as reported in the SEP.

5.1 Term structures of expectations and forecast uncertainty

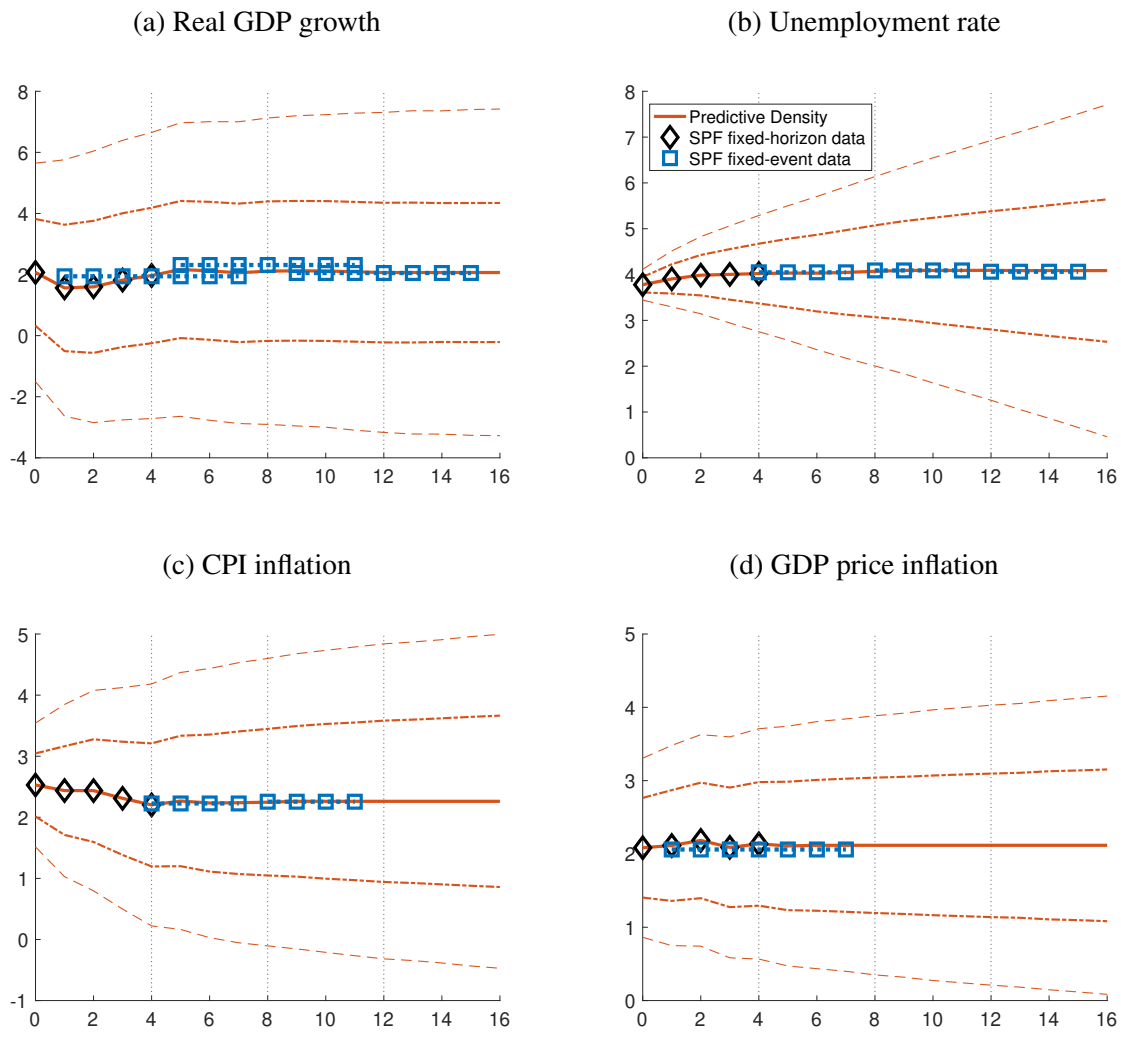
To illustrate the term structures of SPF-consistent expectations produced by our MDS model, Figure 2 presents fan charts of quarterly forecasts from an origin of 2024Q1. As detailed above, to generate these out-of-sample forecasts the model uses as inputs quarterly SPF forecasts up through the four-step-ahead horizon and the available annual fixed-event forecasts (except for the current year), available as of the indicated forecast origin (and not future forecasts or other data). We report the mean forecasts and 68 and 90 percent forecast intervals.

In this example, the SPF-consistent point forecasts show GDP growth remaining around 2 percent, the unemployment rate edging up just a little before settling around 4 percent, and inflation eventually settling at or (in the case of the CPI) a little above 2 percent. As expected, the forecast intervals widen — i.e., forecast uncertainty increases — as the quarterly horizon grows. Below we take up in more detail the term structure of uncertainty and its variation over time.

Regarding the construction of the SPF-consistent quarterly forecasts, by design the MDS model's forecasts for horizons 0 through 4 quarters exactly match the reported SPF projections. Forecasts

at horizons 5 through H interpolate (up to measurement error) through the SPF's published annual forecasts of growth, unemployment, and inflation. That interpolation is not necessarily simply linear; the model is capable of capturing richer dynamics based on the historical comovement among the observed updates in SPF forecasts. At these horizons of 5 and more quarters, there is some uncertainty around the model's estimates of the SPF-consistent quarterly forecast (i.e., uncertainty around the estimate of the point forecast — the state $F_t y_{t+h}$ that is latent for $h \geq 5$). Comparisons

Figure 2: Quarterly fan charts and SPF per 2024Q1



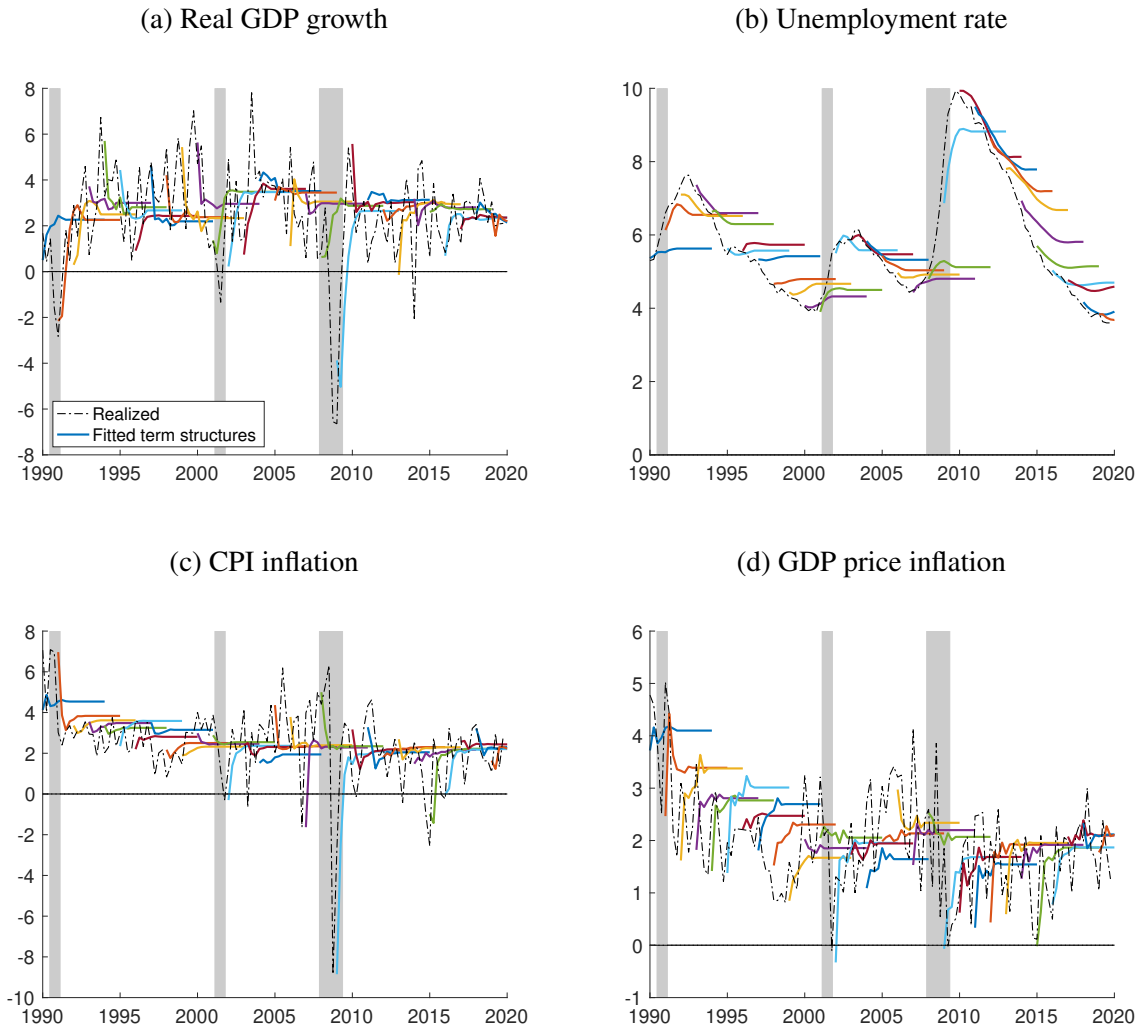
Notes: SPF-consistent predictive densities as generated from our MDS model, showing predictive mean as well as 68% and 90% uncertainty bands. Quarterly forecast horizons on the horizontal axis. Diamonds indicate observed values for SPF fixed-horizon forecasts for $h = 0, 1, 2, 3, 4$. Squares depict observed values for fixed-event calendar-year predictions from the 2024Q1SPF. The fixed-event calendar-year predictions represent linear combinations of (implied) forecasts for multiple quarters, and squares are shown for all quarters involved, with dots connecting the squares associated with a given calendar year. (As explained in the text, annual SPF predictions for GDP growth and GDP price inflation involve quarters in adjacent years.)

omitted in the interest of brevity confirm that, as expected, the uncertainty around the quarterly forecast estimates is reduced by having annual forecasts as measurements. However, the uncertainty around the latent state estimate is a small component of the overall forecast uncertainty reflected in the fan charts. The overall forecast uncertainty is captured and driven by our model's estimates of the time-varying variances of historical forecast errors and updates.

To provide a broader time perspective on the term structure of SPF-consistent expectations, Figure 3 shows the history of term structures of SPF-consistent forecasts (these are all out-of-sample) from 1990 through 2019. For readability, we only report forecasts originating in the first quarter of each year, and the figures stop in 2019 to avoid the pandemic period's volatility. A given colored solid line provides the SPF-consistent quarterly forecasts for a given forecast origin, with different lines for different origins, and dotted lines provide the outcomes of each variable. Across quarters for a given origin, the forecasts often show some fluctuations before settling at the model's take on the long-run mean. Over time, however, these means show some changes; the forecast paths evolve to follow the broad contours of economic outcomes. This pattern is starkest with the decline of actual unemployment and inflation from about 1990 to 2000. Also as expected, from one forecast origin to another, short-horizon forecasts change more abruptly than do the longer-horizon forecasts. The forecasts of growth and unemployment reflect the widely familiar difficulty of forecasting recessions in advance; for example, the SPF-consistent forecasts of growth turn sharply negative after the recessions of 1990-91 and 2007-2009 become evident and, in each case, anticipate recoveries. On the other hand, SPF forecasters were consistently surprised by how quickly unemployment declined after the 2007-2009 recession.

The dynamic behavior of the SPF-consistent forecasts across quarters reflects forecasters' views of underlying data processes. In results detailed in the supplementary online appendix, it is possible to derive univariate processes for the time series y_t implied by estimates of our model of SPF forecasts (conditioning on the entire SPF term structure). Essentially, using the model's steady-state Kalman filter and the Kalman gain, we can obtain moving average (MA) coefficients of the process for y_t. In MDS model-based estimates for GDP growth and the two inflation measures, the coefficient profiles resemble low-order MA processes, with a little more persistence for inflation than growth. Following a surprise increase, these variables steadily return to their longer-run levels. The implied process for the unemployment rate resembles a much more persistent

Figure 3: SPF-consistent term structures of expectations over time (MDS)



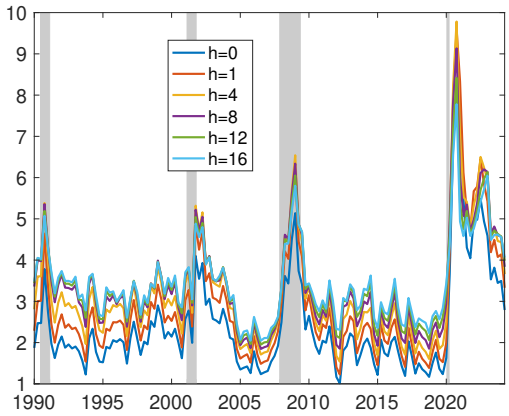
Note: Term structures of expectations for quarterly forecasts, generated out-of-sample from our MDS model at different forecast origins, and realized values. For sake of readability, only forecast origins in Q1 and for years prior to the COVID-19 pandemic are shown. Shaded areas depict NBER recessions.

AR process with hump-shaped dynamics, such that following a surprise increase, unemployment rises further before slowly declining. Estimated processes implied by the VAR specification are qualitatively similar to those from the MDS specification.

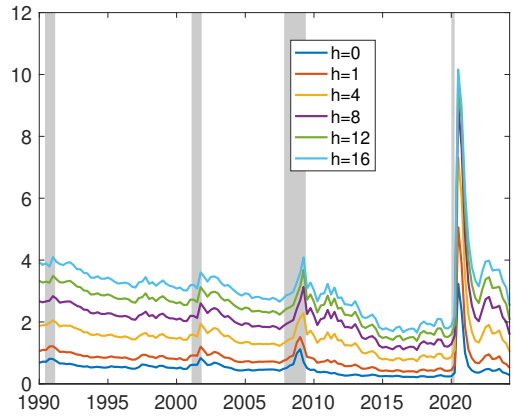
Turning from SPF-consistent point forecasts to the term structure of forecast uncertainty, Figure 4 depicts the term structure of uncertainty around quarterly forecasts, from 1990 to 2023. For constructing the figure, we measure uncertainty by the width of the 68 percent bands of the model's predictive densities estimated in real time (i.e., we report out-of-sample forecast uncertainty). For readability, the chart includes a subset of quarterly horizons, including a few short ones and longer horizons at increments of 4 quarters.

Figure 4: Term structures of uncertainty over time (MDS)

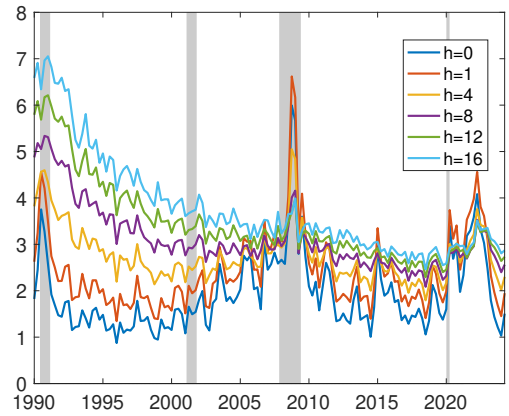
(a) Real GDP growth



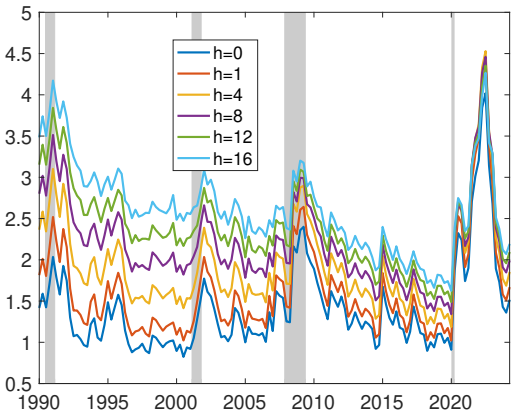
(b) Unemployment rate



(c) CPI inflation



(d) GDP price inflation



Note: Uncertainty measured by the width of 68% predictive intervals, generated out of sample from our MDS model, for selected quarterly forecast horizons. Shaded areas depict NBER recessions.

In these results, uncertainty is noticeably higher at longer horizons (8 or more quarters) than shorter horizons (0 to 4 quarters). While this applies to all variables, short-horizon uncertainty compared to long-horizon uncertainty is relatively high for GDP growth. The uncertainty of out-of-sample forecasts fluctuates significantly over time (a feature we capture by including stochastic volatility in the model). For GDP growth and unemployment, uncertainty rose some following the 2001 recession and increased more notably around the Great Recession and again a few years into the ensuing recovery (while omitted from the chart, our estimates also show uncertainty declining significantly with the Great Moderation that occurred earlier in the sample). Then the outbreak of COVID-19 produced an unprecedented, but temporary, spike in uncertainty in 2020. As of

2023Q4, forecast uncertainty for GDP growth and unemployment remains above its pre-pandemic level. The uncertainty estimates for the inflation measures sharply trended down from 1990 to 2000 (inflation uncertainty also declined significantly with the Great Moderation) and then fluctuated over the remainder of the sample, with some increases around the Great Recession and pandemic.

5.2 Quality of MDS and VAR forecasts

This section examines the quality of our SPF-consistent forecasts along various dimensions. This quality assessment includes out-of-sample forecasts from both the MDS and VAR models. Unless otherwise noted, our evaluations use forecasts through 2023Q4. (The supplementary online appendix contains corresponding results for a sample of forecasts ending before the pandemic, which yield the same qualitative findings.) After presenting the results from the MDS and VAR specifications, we discuss interpretations and implications of their relative performance.

Regarding departures from forecast rationality, before taking up forecast quality we assess the extent to which the general VAR specification reflects them. Building on the framework of Nordhaus (1987) and subsequent empirical work, as well as work on information rigidities in macroeconomics, Coibion and Gorodnichenko (2015) propose quantifying departures from rationality with the coefficients of regressions of forecast errors on forecast updates:

$$y_{t+h} - F_t y_{t+h} = \alpha_h + \beta_h \left(F_t - F_{t-1}\right) y_{t+h} + \text{error}_{t+h} \,. \tag{20}$$

Many subsequent studies — often pooling estimates across forecast horizons — have provided related evidence using this regression for various forecasts. We apply this metric to estimates of our VAR specification. From standard regression calculus, the β_h coefficient of (20) is a function of the moments of forecast errors and updates. As detailed in the supplementary online appendix, our VAR specification of SPF forecasts can be used to compute corresponding moment values using the estimated coefficients of the VAR and in turn to compute the implied β_h coefficient. We use this approach to obtain pooled estimates of β for the horizons for which observed fixed-horizon forecasts are available from the SPF.[13] As reported in the supplementary online appendix, these

[13] The SPF generally provides fixed-horizon forecasts for quarters $h = 0$ through 4. Since the right-hand side of equation (20) requires observations for an $(h+1)$-step-ahead forecast, it is typical to pool forecasts for $h = 0$ through 3, which we do as well.

estimates are generally in line with the literature that finds positive coefficients of small-to-modest magnitudes, indicating some departures from full rationality in professional forecasts.

One metric of the quality of our estimated term structure of SPF (point) forecasts is whether they are good predictions of the SPF published in the subsequent quarter — in particular, whether our estimates of forecasts not directly observed at forecast origin t are good predictions of the forecasts (at the same horizon) directly observed in the SPF published at forecast origin $t+1$. We rely on estimates of regressions styled after the efficiency test of Mincer and Zarnowitz (1969):

$$F_{t+1}y_{t+h} = \alpha_h + \beta_h F_t y_{t+h} + \text{error}_{t+h}\,, \tag{21}$$

where we only use horizons such that the $t+1$ forecast on the left side is directly observed and the t forecast on the right relies — at longer horizons — on latent quarterly forecasts estimated from our model. (While we simplify notation here to just refer to quarterly forecasts, our regressions also cover annual forecasts at longer horizons.) If the predictions are good, the slope coefficient β_h will be close to 1, with an intercept α_h of 0. Table 1 reports these regression estimates for a sample of 1990-2023, separately by model, variable, and horizon.

Generally, these tests of the predictability of SPF point forecasts indicate that our models fare reasonably well in producing forecasts that are efficient or nearly efficient forecasts of the next and directly observed SPF. For example, for GDP growth, the slope coefficients obtained for the MDS forecasts are all close to 1; only at a horizon of $h = 4$ does the estimate differ significantly from unity. With MDS forecasts, the same pattern applies for UNRATE, PGDP, and CPI: most slope coefficients are close to 1, with one or two horizons departing significantly from a coefficient of unity. By this metric, the quality of the forecasts at longer horizons appears to be comparable to that at shorter horizons. The results are broadly similar for SPF forecasts obtained with the VAR generalization of the model, albeit with a few more rejections of slope coefficients of unity (most sharply for PGDP). Of course, given the count of tests reflected in the table's results for different variables and horizons, some rejections of unity slope coefficients are to be expected given Type II error. Broadly, we take these results as evidence that our interpolation of quarterly forecasts at longer horizons is yielding forecasts comparable in quality to the shorter-horizon forecasts.

As another measure of forecast quality, we conduct an out-of-sample evaluation of real-time forecasts generated by the MDS and VAR versions of the model. Table 2 compares the accuracy

Table 1: Predictability of SPF point forecasts

	RGDP		UNRATE		PGDP		CPI	
Forecast	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR
h = 0	1.41	1.32	0.86	0.83	0.99	0.98	1.15	1.19
	(0.24)	(0.25)	(0.08)	(0.10)	(0.05)	(0.06)	(0.16)	(0.15)
h = 1	1.02	1.02	0.91	0.86	1.02	0.99	1.01	1.00
	(0.09)	(0.09)	(0.06)	(0.08)	(0.04)	(0.04)	(0.07)	(0.06)
h = 2	1.02	0.97	0.94	0.90	**0.94**	**0.85**	0.94	0.95
	(0.08)	(0.08)	(0.06)	(0.07)	(0.03)	(0.03)	(0.04)	(0.05)
h = 3	0.94	**0.80**	0.96	0.91	**0.92**	**0.86**	**0.92**	**0.92**
	(0.09)	(0.09)	(0.05)	(0.06)	(0.04)	(0.03)	(0.04)	(0.04)
h = 4	**0.87**	**0.57**	0.97	0.93	**0.91**	**0.90**	0.94	**0.89**
	(0.06)	(0.09)	(0.05)	(0.06)	(0.03)	(0.04)	(0.04)	(0.03)
y = 1	0.94	0.91	0.96	0.93	**0.93**	**0.91**	0.98	0.96
	(0.09)	(0.07)	(0.05)	(0.06)	(0.03)	(0.03)	(0.05)	(0.05)
y = 2	0.94	0.96	0.92	0.95	—	—	0.85	**0.74**
	(0.09)	(0.10)	(0.07)	(0.04)	—	—	(0.11)	(0.09)
y = 3	0.95	0.59	**0.76**	0.95	—	—	—	—
	(0.06)	(0.24)	(0.08)	(0.04)	—	—	—	—

Notes: Estimated slope coefficients of Mincer-Zarnowitz regressions for model-based predictions of next-quarter's published values for SPF forecasts at different forecast horizons. Heteroskedasticity-consistent standard errors in parentheses. Bold font distinguishes coefficient estimates significantly different from 1 with a 10% confidence level. Evaluation window from 1990Q1 to 2023Q4 (and as far as data for SPF forecasts at the different horizons are available).

of point forecasts, evaluated in terms of RMSE, as well as density predictions, evaluated by the continuous ranked probability score (CRPS). From 1990Q1 onward, out-of-sample forecasts are generated for quarterly horizons from $h = 0$ through 16, by re-estimating the model at each forecast origin (using all available data since 1968Q4), and simulating its predictive density. The results are reported as ratios with MDS results in the denominator, so that entries below (above) 1 indicate that the VAR forecasts are more (less) accurate than the MDS forecasts.

These results indicate that the point and density forecasts (again, these are out-of-sample) from the MDS and VAR specifications have broadly similar accuracy over the full sample. For RGDP, the RMSE and CRPS ratios are all very close to 1, at both shorter horizons (for which SPF forecasts are directly observed) and longer horizons (for which quarterly forecasts must be interpolated by the models). For the inflation measures, the RMSE and CRPS ratios are typically just a bit above

Table 2: Relative forecast accuracy of MDS vs. VAR model

	RMSE				CRPS			
h	RGDP	UNRATE	PGDP	CPI	RGDP	UNRATE	PGDP	CPI
0	1.01	1.12	0.99	0.91**	1.02	0.95	1.00	0.93***
1	1.04	1.05	1.02	1.01	1.01	1.00	1.00	1.01
2	0.99	1.07	1.01	1.01	0.99	1.01	0.99	1.01
3	1.00	1.04	1.02	1.01	1.01	1.00	1.00	1.01
4	1.00	1.02	1.03	1.01	1.00	1.00	1.00	1.01
5	1.00	1.02	1.04	1.01	1.00	1.01	1.01	1.01
6	1.00	1.02	1.04	1.01	1.01	1.02	1.02	1.01
7	1.00	1.02	1.04	1.01	0.99	1.03	1.02	1.01
8	1.00	1.02*	1.04	1.01	1.00	1.03	1.02	1.01
9	1.00	1.03*	1.04	1.01	1.00	1.04	1.03	1.01
10	1.00	1.03	1.04	1.01	1.00	1.04	1.03	1.02
11	1.00	1.02	1.04	1.01	1.00	1.03	1.03	1.02
12	1.00	1.01	1.04	1.01	1.00	1.02	1.03	1.01
13	1.00	1.00	1.04	1.01	1.00	1.00	1.03	1.01
14	1.00	0.99	1.04	1.01	1.01	0.99	1.04	1.01
15	1.00	1.00	1.04	1.01	1.01	1.00	1.03	1.01
16	1.00	1.00	1.04	1.00	1.01	0.99	1.03	1.01

Note: Relative RMSE and CRPS of VAR model (with MDS in denominator). Quarterly forecast horizons, h. Evaluation window from 1990Q1 through 2023Q4 (and as far as realized values are available). 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.

1.00, giving a very slight advantage (not enough to be considered material) to the MDS model.[14] The MDS model's advantage over the VAR is modestly greater for UNRATE.

As another measure of forecast quality, Table 3 reports unconditional coverage rates — the frequency with which actual quarterly outcomes fall within 68 and 90 percent forecast intervals. (In the 90 percent case, there are relatively few observations available for evaluating accuracy in the tails of the distributions.) A frequency of more (less) than 68/90 percent means that, on average over a given sample, the estimated forecast density is too wide (narrow). We judge the significance of the results using p-values of t-statistics for the null hypothesis that the empirical coverage rate equals the nominal rate. To save space, the table reports results for just the VAR specification; the MDS model yields similar coverage rates, reported in the supplementary online appendix.

[14] Instances of relative RMSE that are worse for the VAR model compared to the MDS case are consistent with results reported in Bianchi et al. (2022), who find generally poor out-of-sample predictability of SPF forecasts based on past forecasts alone.

Overall, these results indicate that our models applied to the SPF yield forecasts and densities that are reasonable, but not perfect, from the perspective of unconditional coverage. The coverage rates are best for CPI and PGDP inflation, with empirical coverage rates comparable to nominal rates and relatively few rejections of correct coverage. Coverage performance is more mixed for GDP growth. At longer horizons, coverage rates are close to nominal rates (with few significant departures), although empirical rates run a bit below nominal rates, indicating the intervals are a little too narrow. At shorter horizons, the empirical coverage rates are more notably below nominal rates. Unconditional coverage performance is also mixed for unemployment forecasts — for 68 percent coverage and less so for 90 percent coverage. For unemployment, 68 percent predictive intervals tend to be too wide at shorter horizons and too narrow at longer horizons.

Of course, unconditional coverage rates may be seen as a limited window into the calibration of the predictive densities. Rossi and Sekhposyan (2019) develop a broader approach to assessing

Table 3: Forecast coverage rates (VAR model)

	RGDP		UNRATE		PGDP		CPI	
h	68%	90%	68%	90%	68%	90%	68%	90%
0	50.00***	81.62***	76.47*	94.85**	55.88***	84.56*	66.18	91.18
1	55.56***	79.26***	77.78**	93.33	62.22	82.22**	60.74	84.44
2	58.21**	82.84**	77.61*	94.03	65.67	85.82	63.43	84.33
3	53.38***	82.71*	78.20*	91.73	64.66	86.47	64.66	86.47
4	56.06**	85.61	68.94	90.91	62.88	86.36	64.39	84.09
5	59.54	86.26	66.41	89.31	64.89	87.79	68.70	87.02
6	58.46**	86.15	63.85	88.46	64.62	87.69	66.15	88.46
7	64.34	86.05	62.02	88.37	64.34	89.92	65.89	87.60
8	62.50	84.38	60.16	87.50	65.62	88.28	68.75	86.72
9	62.20	83.46	56.69*	86.61	66.93	86.61	66.14	88.19
10	62.70	84.92	56.35*	85.71	65.87	89.68	67.46	87.30
11	66.40	83.20	52.00**	83.20	69.60	88.80	72.00	88.00
12	63.71	84.68	50.00**	82.26	69.35	89.52	70.16	87.90
13	69.11	84.55	48.78**	82.93	67.48	90.24	70.73	87.80
14	66.39	86.89	49.18**	82.79	69.67	89.34	69.67	88.52
15	64.46	86.78	48.76**	82.64	69.42	88.43	69.42	89.26
16	66.67	87.50	48.33*	81.67	71.67	90.00	71.67	89.17

Note: Coverage rates for uncertainty bands with nominal levels of 68% and 90% for out-of-sample forecasts at quarterly forecast horizons, h. Evaluation window from 1990Q1 through 2023Q4 (and as far as realized values are available). 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.

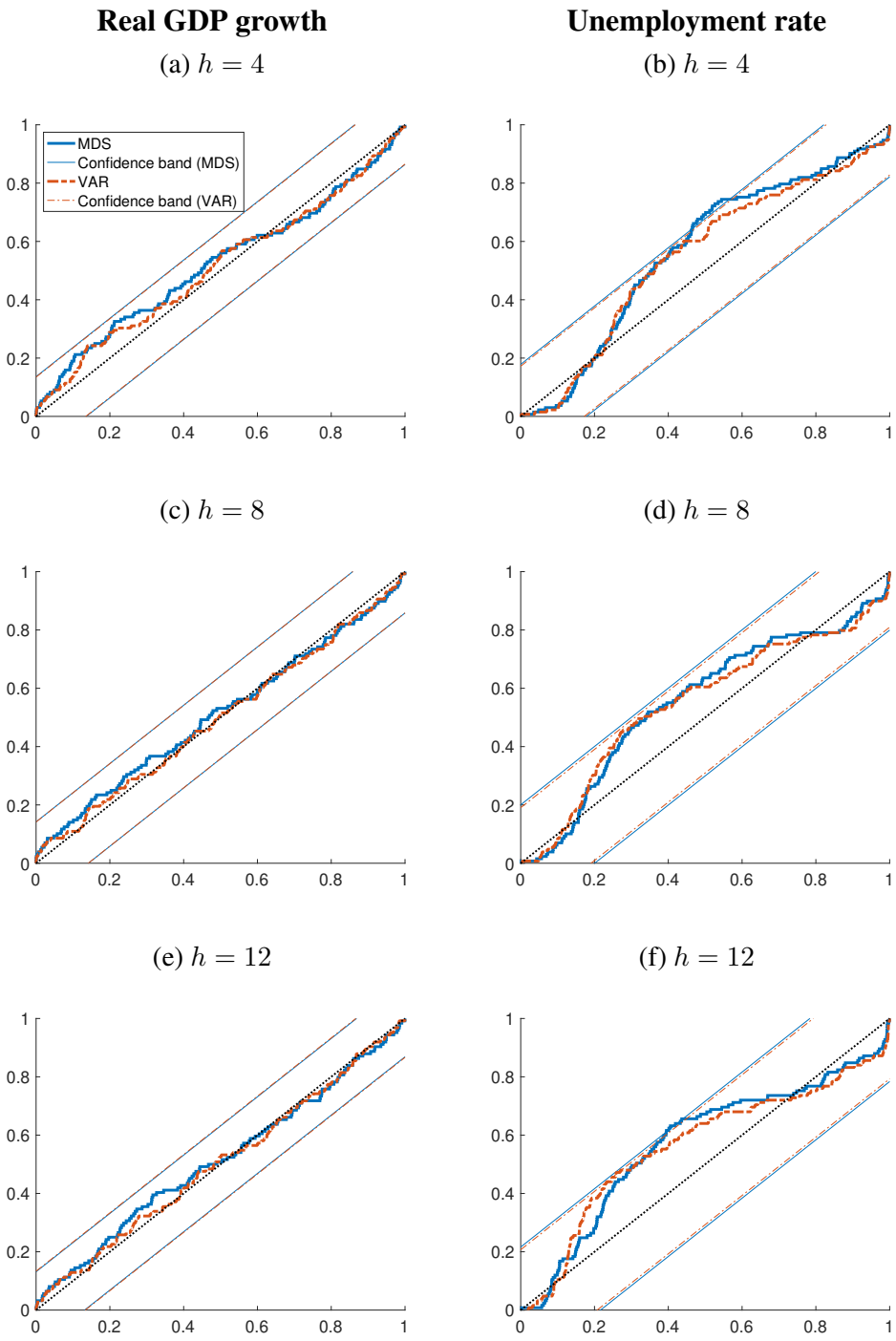
whether predictive densities are correctly specified: testing the uniformity of the empirical cumulative distribution functions (CDFs) of probability integral transforms (PITs) of forecasts. Their approach compares empirical CDFs to the 45-degree line (representing a uniform CDF) with appropriate confidence intervals for departures from uniformity. In the interest of brevity, Figure 5 provides these results for MDS and VAR forecasts of GDP growth and unemployment, for selected horizons (the supplementary online appendix provides results for inflation).

The PITs comparisons indicate that the predictive densities are correctly calibrated for GDP growth, with empirical CDFs of the PITs within the confidence intervals. This applies to both the MDS and VAR forecasts. Importantly, the forecasts are correctly specified at both the shorter horizons for which quarterly SPF forecasts are directly observed and the longer horizons for which quarterly forecasts must be estimated with our models. The departures from uniformity are somewhat greater for the PITs of unemployment forecasts, with the empirical CDFs further away from the 45-degree line. Graphically, while these departures come close to or hit the confidence interval bounds, they generally do not fall outside the bounds, except in the case of $h = 4$ forecasts from the MDS model. In this instance, the VAR performs a little better than the MDS model by allowing some bias and rationality departures in the forecasts.

While we omit details in the interest of brevity (estimates are available in the supplementary online appendix), there is one metric by which the MDS specification has a clear advantage: Marginal data densities (the conventional Bayesian measure of model fit) estimated with one-step-ahead predictive likelihoods indicate that the MDS specification fits the data significantly better than the more general VAR model. Nonetheless, we have shown that the VAR model implies departures from full rationality in keeping with those documented in Coibion and Gorodnichenko (2015) and subsequent studies, in which forecast updates of professional forecasters show statistically significant serial correlation. Yet the battery of additional metrics of forecast quality reported in this section indicate that the more restricted MDS specification is on par with the more general VAR; the models are hard to distinguish in those comparisons. These results suggest little benefit — to forecast quality — to generalizing our baseline model to allow departures from our MDS assumption. But they also show that doing so imposes little cost on forecast quality.

In our assessment, the ambiguity around the best model can be seen as consistent with the broader literature on survey forecasts. The strong evidence of bias and other departures from rationality in survey forecasts comes from in-sample analysis of the forecasts. Various studies

Figure 5: PITs for real GDP growth and the unemployment rate



Notes: Empirical cumulative distributions of probability integral transforms (PITs) for GDP growth and unemployment at selected quarterly forecast horizons. All forecasts are generated out of sample by our MDS and VAR models, and evaluated over an evaluation window from 1990Q1 through 2023Q4 (and as far as realized values are available). 95% confidence bands for tests of correct calibration from Rossi and Sekhposyan (2019); computed separately for each model, but with nearly identical plot lines.

have found these departures more challenging to exploit on an out-of-sample basis. Croushore (2010) documents that deviations of the SPF (and the Livingston Survey) forecasts from rationality are typically short-lived (episodic) and hard to exploit in real time; more recently, Foerster and Matthes (2022) and Hajdini and Kurmann (2024) have found similar results. Similarly, Eva and Winkler (2023) argue that the departures from rationality in Coibion and Gorodnichenko (2015) and subsequent studies do not apply on an entirely out-of-sample basis and cannot be used (in linear models) to improve on the SPF forecasts in real time. Biases being short-lived, contributing at most only a small fraction to overall MSE, is consistent with notions of SPF forecasts providing at least boundedly optimal, if not highly rational, forecasts, as discussed in, for example, Mertens and Nason (2020). On the other hand, Bianchi et al. (2022) find that machine learning algorithms — fed large amounts of information — can be used to improve forecast accuracy in survey forecasts subject to time-varying bias. Ultimately, what model would we recommend? Our own inclination is for the simpler MDS model that yields forecasts that are fully SPF-consistent. But researchers with a stronger preference to allow for biases and other departures from rationality in SPF forecasts can comfortably proceed with the VAR specification.

5.3 SPF forecasts as compared to SEP projections

In the introduction, we illustrated the use of our approach to produce SEP-like fan charts for SPF-consistent forecasts. Of course, this raises a question as to how our SPF forecasts compare to SEP forecasts over a longer time period. As context, note that Reifschneider and Tulip (2019) showed that survey forecasts and Federal Reserve forecasts are very similar in terms of RMSE accuracy over long periods of time. Their comparison relied on the fixed-event annual forecasts published by various professional sources, including surveys; we will instead use our term structure of quarterly SPF forecasts to construct the fixed-event forecasts defined as in the SEP, which are Q4/Q4 growth and inflation rates and the Q4 unemployment rate. Since SEP forecasts are not available before late 2007, the sample for comparison is less than 20 years; so we rely on informal comparisons rather than engaging in formal inference that would likely be imprecise in the small sample available.[15]

[15]The available time series for the SEP begins in October 2007, and the charts lack SEP observations for 2020 because, in March 2020, in the aftermath of the pandemic's outbreak, the FOMC did not publish an SEP. For inflation, we compare SEP forecasts for PCE inflation with our CPI forecasts, since the SPF provides forecasts for PCE inflation only since 2007.

To compare SEP forecasts with our forecasts constructed from the SPF using the MDS model, Figure 6 reports the errors in real-time forecasts (errors measured as realizations less forecasts) from the SEP and SPF, along with 68 percent uncertainty bands. The figure provides results for forecasts of the current year and the following two years. Note that, for each year from 2008 through 2023, we report estimates for SPF and SEP forecasts published every quarter; within each year, the (fixed-event annual) forecast horizon shrinks before resetting the next year, resulting in sawtooth patterns in the charts. In keeping with the SEP fan charts, in the SEP results in our figure the errors are based on point predictions measured as the median forecasts across participants, and the uncertainty bands are $+/-$ one times the historical RMSEs published in the SEP's Table 2. In the SPF results in the figure, the forecast errors are based on the posterior mean forecast obtained from our MDS model, and the uncertainty bands are based on the 16th and 84th percentiles of the predictive distribution.

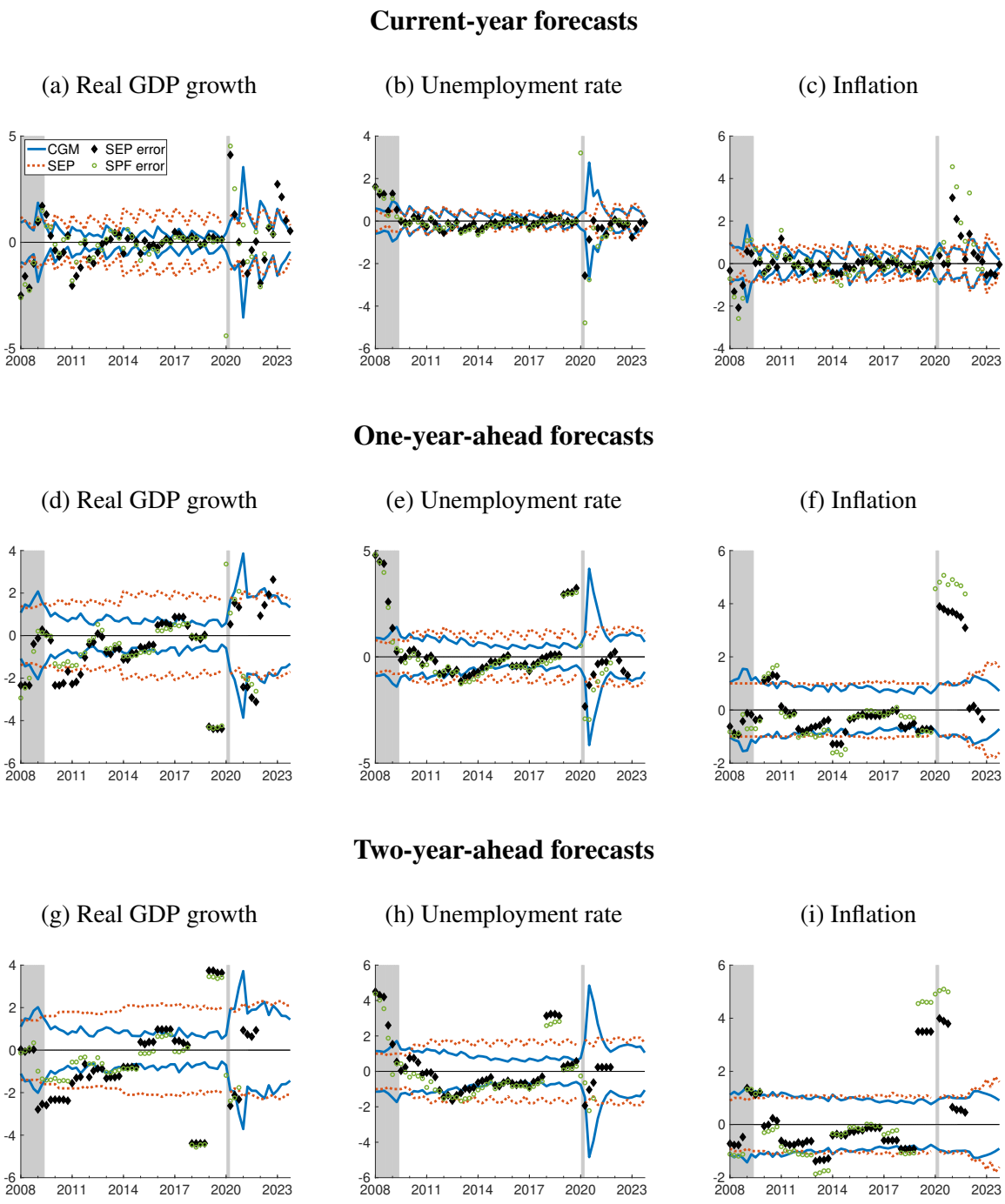
The forecast errors reported in dots in the charts are broadly similar for the SEP and SPF, indicating that the forecasts were also broadly similar. Both the SEP and SPF-consistent projections of GDP growth, unemployment, and inflation made relatively large errors around the Great Recession and the pandemic. In some instances, such as with inflation following the pandemic, SPF forecasts were less accurate than SEP projections, whereas the reverse is true in other cases.

The 68 percent forecast uncertainty bands show more differences across the SEP and SPF than do the point forecasts. The SPF-consistent uncertainty bands from our MDS model are noticeably narrower than the SEP bands based on a simple 20-year rolling window of historical RMSEs of various professional forecasts. These differences are starker for GDP growth and unemployment than for inflation. The errors in SPF-consistent forecasts fall outside the SPF bands with some regularity, whereas the errors in SEP forecasts are rarely outside the SEP bands.[16] (Ideally, for correct coverage, about 32 percent of errors would be outside the bands.) The SEP appears to instead have bands wide enough to imply empirical coverage rates well above the nominal 68 percent rate. In this context, while the SPF-consistent bands are imperfect, in practice they could be a useful point of comparison to bands published in SEP fan charts.

Our SPF-consistent forecast uncertainty bands also differ from the SEP bands in their time variation. As might be expected, our estimates show more variation over time than do the SEP

[16] In results omitted in the interest of brevity, we have verified that the patterns in results shown above for coverage rates of quarterly forecasts carry over to SEP-styled fixed-event forecasts.

Figure 6: Error bands of annual forecasts: MDS model vs. SEP



Note: Forecast errors measured as realizations less forecasts. Model-based uncertainty bands correspond to 68% predictive intervals. FOMC's SEP bands reflect the historical RMSEs of professional forecasts over the previous 20 years as described by Reifschneider and Tulip (2019). SPF-consistent densities and (ex-post) errors for calendar-year definitions of the SEP generated from our MDS model. For inflation, we compare model-based densities for CPI against the SEP's forecasts for PCE inflation.

measures, because our models include stochastic volatility, whereas the SEP measures treat forecast error variances as being constant over 20-year windows. Following the Great Recession, our estimates of uncertainty rise notably; the SEP measures show much less change. In the aftermath of the recession, the SEP bands eventually widen a little, but they do so too late and too little in comparison to our model's estimates. In the ensuing economic recovery, our estimates fall quickly. From 2015 to 2019, both our estimates of uncertainty and the associated forecast error bands and the SEP estimates of uncertainty and forecast bands are relatively stable, with SEP bands wider than the SPF bands. Following the volatility induced by the outbreak of the pandemic in 2020, for GDP growth and unemployment our SV-based estimates of uncertainty rose sharply, to exceed the SEP estimates, resulting in much wider 68 percent forecast error bands, but only temporarily. For CPI inflation, our SPF-based estimates of forecast uncertainty have contours similar to those of the SEP bands until the pandemic. The pandemic's forecast errors initially caused uncertainty bands from our model to widen and then narrow, whereas the SEP bands trended upward with some delay.

6 Conclusion

This paper develops models that take published survey point forecasts — in our applications, from the SPF — as inputs and produce estimates of a longer, more complete term structure (across horizons) of survey-consistent point forecasts, along with a term structure of forecast uncertainty. Our estimates of forecast uncertainty reflect dispersion in past errors from the SPF's point forecasts, with time-varying uncertainty captured through stochastic volatility. Our methods can exactly replicate the SPF's quarterly forecasts at short horizons while interpolating through the available annual forecasts at longer horizons. To avoid excessively volatile imputation that can arise from small inconsistencies between quarterly and annual SPF predictions, we model annual forecasts with fat-tailed measurement errors that are a priori close to zero. Extending previous work, our models (1) provide SPF-consistent quarterly term structures of expectations and uncertainty that extend arbitrarily far ahead, while (2) handling the ragged edge of the SPF's fixed-event forecasts at longer horizons, and (3) offering a way to capture potential inefficiencies in SPF forecasts.

After illustrating the use of our approach to produce fan charts of annual forecasts (of Q4/Q4 percent changes or Q4 levels) directly analogous to those published by the FOMC, we show that

our baseline model yields SPF-consistent quarterly real-time forecasts of GDP growth, unemployment, and inflation that match the published fixed-horizon quarterly forecasts from the SPF and interpolate through the annual fixed-event point forecasts. The model's estimates of forecast uncertainty vary over time, temporarily rising around recessions, and generally rise with the forecast horizon. The SPF forecasts obtained from our models perform comparably — and reasonably well — in various metrics of forecast quality, including efficiency as forecasts of the future SPF, unconditional coverage rates, and overall density calibration. The quality of our SPF-consistent estimated forecasts is comparable to the quality of the published short-horizon forecasts that have been widely used in research and practice. Our models generate fan charts with time-varying forecast uncertainty and reliable coverage rates.

References

Ang, Andrew, Geert Bekaert and MinWei. (2007). "Do macro variables, asset markets, or surveys forecast inflation better?" *Journal of Monetary Economics*, 54(4), pp. 1163–1212. https://doi.org/10.1016/j.jmoneco.2006.04.006

Arias, Jonas E., Juan F. Rubio-Ramirez and Minchul Shin. (2023). "Macroeconomic forecasting and variable ordering in multivariate stochastic volatility models". *Journal of Econometrics*, 235(2), pp. 1054–1086. https://doi.org/10.1016/j.jeconom.2022.04.013

Aruoba, S. Boragan. (2020). "Term structures of inflation expectations and real interest rates". *Journal of Business & Economic Statistics*, 38(3), pp. 542–553. https://doi.org/10.1080/07350015.2018.1529599

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. https://doi.org/10.2139/ssrn.3838719

Bianchi, Francesco, Sydney C. Ludvigson and Sai Ma. (2022). "Belief distortions and macroeconomic fluctuations". *American Economic Review*, 112(7), pp. 2269–2315. https://doi.org/10.1257/aer.20201713

Carriero, Andrea, Todd E. Clark, Massimiliano Marcellino and Elmar Mertens. (2022). "Addressing COVID-19 outliers in BVARs with stochastic volatility". *Review of Economics and Statistics*, forthcoming. https://doi.org/10.1162/rest_a_01213

Carvalho, Carlos M., Nicholas G. Polson and James G. Scott. (2010). "The horseshoe estimator for sparse signals". *Biometrika*, 97(2), pp. 465–480. https://doi.org/10.1093/biomet/asq017

Chan, Joshua C. C. (2020). "Large Bayesian VARs: A flexible Kronecker error covariance structure". *Journal of Business & Economic Statistics*, 38(1), pp. 68–79. https://doi.org/10.1080/07350015.2018.1451336

Clark, Todd E., Michael W. McCracken and Elmar Mertens. (2020). "Modeling time-varying uncertainty of multiple-horizon forecast errors". *The Review of Economics and Statistics*, 102(1), pp. 17–33. https://doi.org/10.1162/rest_a_00809

Clark, Todd E., and Elmar Mertens. (2023). "Stochastic volatility in Bayesian vector autoregressions". In *Oxford Research Encyclopedia in Economics and Finance*. Oxford University Press. https://doi.org/10.1093/acrefore/9780190625979.013.919

Clements, Michael P. (2022). "Individual forecaster perceptions of the persistence of shocks to GDP". *Journal of Applied Econometrics*, 37(3), pp. 640–656. https://doi.org/10.1002/jae.2884

Coibion, Olivier, and Yuriy Gorodnichenko. (2015). "Information rigidity and the expectations formation process: A simple framework and new facts". *American Economic Review*, 105(8), pp. 2644–2678. https://doi.org/10.1257/aer.20110306

Croushore, Dean. (2010). "An evaluation of inflation forecasts from surveys using real-time data". *The B.E. Journal of Macroeconomics*, 10(1), pp. 1–32. https://doi.org/10.2202/1935-1690.1677

Crump, Richard K., Stefano Eusepi, Emanuel Moench and Bruce Preston. (2023). "The term structure of expectations". In Bachmann, Rüdiger, Giorgio Topa and Wilbert van der Klaauw (eds.), *Handbook of Economic Expectations*, chap. 17. Academic Press, pp. 507–540. https://doi.org/10.1016/B978-0-12-822927-9.00025-2

Diebold, Francis X., and Roberto S. Mariano. (1995). "Comparing predictive accuracy". *Journal of Business and Economic Statistics*, 13(3), pp. 253–263. https://doi.org/10.2307/1392185

Doh, Taeyoung, and A. Lee Smith. (2022). "A new approach to integrating expectations into VAR models". *Journal of Monetary Economics*, 132, pp. 24–43. https://doi.org/10.1016/j.jmoneco.2022.08.001

Eva, Kenneth, and Fabian Winkler. (2023). "A comprehensive empirical evaluation of biases in expectation formation". Finance and Economics Discussion Series, 2023-042, Board of Governors of the Federal Reserve System (U.S.). https://doi.org/10.17016/FEDS.2023.042

Farmer, Leland, Emi Nakamura and Jon Steinsson. (2024). "Learning about the long run". *Journal of Political Economy*, forthcoming. https://doi.org/10.1086/730207

Faust, Jon, and Jonathan H. Wright. (2009). "Comparing Greenbook and reduced form forecasts using a large realtime dataset". *Journal of Business & Economic Statistics*, 27(4), pp. 468–479. https://doi.org/10.1198/jbes.2009.07214

Foerster, Andrew, and Christian Matthes. (2022). "Learning about regime change". *International Economic Review*, 63(4), pp. 1829–1859. https://doi.org/10.1111/iere.12585

Frey, Christoph, and Frieder Mokinski. (2016). "Forecasting with Bayesian vector autoregressions estimated using professional forecasts". *Journal of Applied Econometrics*, 31(6), pp. 1083–1099. https://doi.org/10.1002/jae.2483

Ganics, Gergely, and Florens Odendahl. (2021). "Bayesian VAR forecasts, survey information, and structural change in the euro area". *International Journal of Forecasting*, 37(2), pp. 971– 999. https://doi.org/10.1016/j.ijforecast.2020.11.001

Grishchenko, Olesya, Sarah Mouabbi and Jean-Paul Renne. (2019). "Measuring inflation anchoring and uncertainty: A U.S. and euro area comparison". *Journal of Money, Credit, and Banking*, 51(5), pp. 1053–1096. https://doi.org/10.1111/jmcb.12622

Hajdini, Ina, and Andr´e Kurmann. (2024). "Predictable forecast errors in full-information rational expectations models with regime shifts". Working Paper, 24-08, Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202408

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(1), pp. 185–212. https://doi.org/10.1016/j.jeconom.2003.09.001

Kim, Sangjoon, Neil Shephard and Siddhartha Chib. (1998). "Stochastic volatility: Likelihood inference and comparison with ARCH models". *The Review of Economic Studies*, 65(3), pp. 361–393. https://doi.org/10.1111/1467-937X.00050

Kozicki, Sharon, and P.A. Tinsley. (2012). "Effective use of survey information in estimating the evolution of expected inflation". *Journal of Money, Credit, and Banking*, 44(1), pp. 145–169. https://doi.org/10.1111/j.1538-4616.2011.00471.x

Krane, Spencer D. (2011). "Professional forecasters' view of permanent and transitory shocks to GDP". *American Economic Journal: Macroeconomics*, 3(1), pp. 184–211. https://doi.org/10.1257/mac.3.1.184

Krüger, Fabian, Todd E. Clark and Francesco Ravazzolo. (2017). "Using entropic tilting to combine BVAR forecasts with external nowcasts". *Journal of Business & Economic Statistics*, 35(3), pp. 470–485. https://doi.org/10.1080/07350015.2015.1087856

Makalic, Enes, and Daniel F. Schmidt. (2016). "A simple sampler for the horseshoe estimator". *IEEE Signal Processing Letters*, 23(1), pp. 179–182. https://doi.org/10.1109/LSP.2015.2503725

Mariano, Roberto S., and Yasutomo Murasawa. (2003). "A new coincident index of business cycles based on monthly and quarterly series". *Journal of Applied Econometrics*, 18(4), pp. 427–443. https://doi.org/10.1002/jae.695

Mertens, Elmar. (2016). "Measuring the level and uncertainty of trend inflation". *The Review of Economics and Statistics*, 98(5), pp. 950–967. https://doi.org/10.1162/REST_a_00549

Mertens, Elmar. (2023). "Precision-based sampling for state space models that have no measurement error". *Journal of Economic Dynamics and Control*, 154, p. 104720. https://doi.org/10.1016/j.jedc.2023.104720

Mertens, Elmar, and James M. Nason. (2020). "Inflation and professional forecast dynamics: An evaluation of stickiness, persistence, and volatility". *Quantitative Economics*, 11(4), pp. 1485–1520. https://doi.org/10.3982/QE980

Mincer, Jacob A., and Victor Zarnowitz. (1969). "The evaluation of economic forecasts". In Mincer, Jacob A. (ed.), *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, chap. 1. NBER, pp. 3–46.

Newey, W.K., and K.D. West. (1987). "A simple positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix". *Econometrica*, 55, pp. 703–708. https://doi.org/10.2307/1913610

Nordhaus, William D. (1987). "Forecasting efficiency: Concepts and applications". *The Review of Economics and Statistics*, 69(4), pp. 667–674. https://doi.org/10.2307/1935962

Patton, Andrew J., and Allan Timmermann. (2011). "Predictability of output growth and inflation: A multi-horizon survey approach". *Journal of Business & Economic Statistics*, 29(3), pp. 397–410. https://doi.org/10.1198/jbes.2010.08347

Patton, Andrew J., and Allan Timmermann. (2012). "Forecast rationality tests based on multihorizon bounds". *Journal of Business & Economic Statistics*, 30(1), pp. 1–17. https://doi.org/10.1080/07350015.2012.634337

Reifschneider, David, and Peter Tulip. (2019). "Gauging the uncertainty of the economic outlook using historical forecasting errors: The Federal Reserve's approach". *International Journal of Forecasting*, 35(4), pp. 1564–1582. https://doi.org/10.1016/j.ijforecast.2018.07.016

Rossi, Barbara, and Tatevik Sekhposyan. (2019). "Alternative tests for correct specification of conditional predictive densities". *Journal of Econometrics*, 208(2), pp. 638–657. https://doi.org/10.1016/j.jeconom.2018.07.008

Wright, Jonathan H. (2013). "Evaluating real-time VAR forecasts with an informative democratic prior". *Journal of Applied Econometrics*, 28(5), pp. 762–776. https://doi.org/10.1002/jae.2268

SUPPLEMENTARY RESULTS

Constructing Fan Charts from the Ragged Edge of SPF Forecasts*

Todd E. Clark,[1] Gergely Ganics,[2] and Elmar Mertens[3]

[1]*Federal Reserve Bank of Cleveland,* [2]*Banco de España,* [3]*Deutsche Bundesbank.*

July 20, 2024

Abstract

This online appendix provides results and descriptions that supplement our paper.

Contents

List of Tables

*The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Cleveland, the Federal Reserve System, the Banco de España, the Deutsche Bundesbank, or the Eurosystem. Replication files are available at `https://github.com/elmarmertens/ClarkGanicsMertensSPFfancharts`.

List of Figures

I SPF data and measurements

This appendix provides further information about the survey data used in our paper and obtained from the Survey of Professional Forecasters (SPF) for the US. In addition, this appendix describes additional details about the mapping between the survey data and the model. Finally, this appendix assesses potential inconsistencies between observed SPF forecasts for quarterly and annual forecast targets and the potential for such inconsistencies to generate excessively volatile imputations in our model's term structure of expectation, when assuming that both quarterly and annual forecasts are observed without error.

I(a) Data availability

As described in the paper, we use SPF forecasts for four variables: growth in real GDP, the unemployment rate, and inflation in CPI and GDP prices. For these variables, the SPF provides fairly long samples of data, albeit with differing availabilities of forecasts at different forecast horizons. The availability of forecasts at different horizons for each of the four variables is listed in Table A.1. All forecast data listed in the table is used for our analysis (of course, for our out-of-sample analysis, the data is used only subject to real-time availability).

Table A.1: Availability of SPF point forecasts

		Fixed-horizon	Fixed-event calendar years		
Variable	Mnemonic	Quarters 0 – 4	next	2-year	3-year
Real GDP	RGDP	1968Q4	1981Q3	2009Q2	2009Q2
Unemployment rate	UNRATE	1968Q4	1981Q3	2009Q2	2009Q2
GDP price index	PGDP	1968Q4	1981Q3	NA	NA
CPI inflation	CPI	1981Q3	1981Q3	2005Q3	NA

Note: The table reports the first quarters in which SPF predictions become available in our data set for the stated variables and horizons. NA stands for not available. Prior to 1992, RGDP corresponds to real GNP, while PGDP corresponds to the GNP implicit deflator. The SPF's published data files include point forecasts for RGDP and PGDP in levels, which we convert to continuously compounded growth rates. The SPF also provides current year predictions that are, however, disregarded in our analysis due to overlap with the quarterly fixed-horizon predictions.

Table A.2: Choice of maximal H in term-structure vector $\boldsymbol{Y_t}$

Variable	Samples	H
RGDP	prior 2009Q2	5
	since 2009Q2	12
UNRATE	prior 2009Q2	5
	since 2009Q2	12
PGDP	entire sample	5
CPI	prior 2005Q4	5
	since 2005Q4	8

In principle, to track forecasts for up to three calendar years ahead, our model's state vector should need to track (latent) quarterly forecasts, $F_t y_{t+h}$, for up to $h = 15$ quarters ahead. However, as described in Section 4 of the paper, the paucity of quarterly forecast data at longer horizons allows us to discipline our models by assuming that forecasts converge to a common trend level sooner than for 15 quarters ahead. We denote the maximal horizon for which we track deviations from trend by H. We generally choose H such that when the forecast origin is in Q1, H points to the first quarter of the farthest annual horizon covered by the SPF forecasts. An exception is made for data covering only SPF forecasts up to the next year, where H is set to 5 (instead of 4), since the observed fixed-horizon SPF forecasts already extend through $h = 4$. As part of our out-of-sample forecast analysis, models are (re-)estimated over different sub-samples of data, and we adjust H accordingly. Reflecting the availability of long-horizon SPF data for different variables, we set H as listed in Table A.2. Note that, even though the state vector ends with $F_t y_{t+H}$ and H is no larger than 12, our endpoint assumption indicated in the paper allows us to simulate forecast densities arbitrarily far ahead, and we report densities up to 16 quarters ahead for all variables throughout.

I(b) Measurement equations

This subsection provides examples of the measurement vector and the loading matrix introduced in Section 4.4 of the main paper, defined as

$$\boldsymbol{Z_t} = \begin{bmatrix} \boldsymbol{Z_{q,t}} \\ \boldsymbol{Z_{a,t}} \end{bmatrix}, \tag{A.1}$$

$$\boldsymbol{C_t} = \begin{bmatrix} \boldsymbol{C_{q,t}} \\ \boldsymbol{C_{a,t}} \end{bmatrix}, \tag{A.2}$$

respectively, where the subscripts (q and a) reflect the partition of the arrays according to the quarterly or annual horizons, and the time index t highlights that these objects vary over time as a function of the available measurements. In the following examples, we take $t = 2024Q1$ as the first forecast origin, and then illustrate how the corresponding elements change when we move to the next quarter, $t = 2024Q2$. Before illustrating the measurement equations with examples for each variable and at different forecasting origins, we review the data definitions used throughout the paper for outcome variables and SPF data.

I(b.1) Data definitions

As indicated in the paper, the (quarterly) outcome variable y_t refers to the following data definitions for each variable:

RGDP: We measure GDP growth by the annualized quarterly growth rate (400 times the log change) of real GDP.

UNRATE: The quarterly average level of the unemployment rate.

PGDP: We measure inflation in GDP prices by the annualized quarterly growth rate of the GDP deflator (as with RGDP, calculated as 400 times the quarterly log change).

CPI: For CPI inflation, we take the annualized simple growth rate of quarterly CPI levels.

For each variable, we denote quarterly outcomes as defined above by y_t.

Considering their treatment in the SPF, the variables listed above can broadly be categorized into two groups: Variables from the national income and product accounts (NIPA) and non-NIPA variables. The NIPA variables are GDP growth and GDP price inflation, while the non-NIPA variables are the unemployment rate and CPI inflation.

For the non-NIPA variables, the SPF provides forecasts for the data definitions stated above. In addition, for the non-NIPA variables, annual forecasts can be represented as four-quarter averages of y_t over the calendar year, whose realization we denote by $\bar{y}_t$, as defined in equation (11) of the paper:

$$\bar{y}_t = \frac{1}{4} \cdot \sum_{j=0}^{3} y_{t-j} \,. \tag{11}$$

For the unemployment rate, the annual forecast targets are defined as annual average levels. For CPI inflation the annual forecasts target Q4/Q4 growth, which we approximate by the arithmetic four-quarter mean of quarterly growth rates (so as to avoid further transformations of the SPF quarterly forecasts).

In contrast, for the NIPA variables, the SPF provides (in the publicly available data files) forecasts in levels. We convert these to log growth rates.[1] Moreover, SPF calendar-year forecasts for NIPA variables reflect annual-average levels. Denoting the quarterly level of a NIPA variable by I_t, growth in the annual-average level for the calendar year ending at quarter t is measured in equation (12) of our paper as follows:

$$\hat{y}_t \equiv 100 \times \log\left(\frac{I_t + I_{t-1} + I_{t-2} + I_{t-3}}{I_{t-4} + I_{t-5} + I_{t-6} + I_{t-7}}\right), \tag{12}$$

Of course, the definition of $\hat{y}_t$ is non-linear, and to capture annual forecasts for GDP growth (and inflation in GDP prices) we employ a log-linear approximation involving 7 quarterly growth rates. The approximation itself has been popularized by the work of Mariano and Murasawa (2003) on nowcasting, and its accuracy has been favorably evaluated in the context of SPF forecasts by (amongst others) Patton and Timmermann (2011). Moreover, the approximation is commonly used in the related literature on SPF forecasts, with notable examples provided by Aruoba (2020), Crump et al. (2023), and Patton and Timmermann (2011). Specifically, this is a log-linear approximation, around a steady-state of zero growth, $I_t/I_{t-j} = 1$ for all j, and with $y_t \equiv 400 \cdot \log(I_t/I_{t-1})$, and is equation (13) of the paper we get:

$$\begin{aligned}\hat{y}_t &\approx \frac{100}{4} \cdot \left(\log\frac{I_t}{I_{t-4}} + \log\frac{I_{t-1}}{I_{t-5}} + \log\frac{I_{t-2}}{I_{t-6}} + \log\frac{I_{t-3}}{I_{t-7}}\right), \\ &= 1/16 \cdot (y_t + 2 \cdot y_{t-1} + 3 \cdot y_{t-2} + 4 \cdot y_{t-3} + 3 \cdot y_{t-4} + 2 \cdot y_{t-5} + y_{t-6}) \,. \end{aligned} \tag{13}$$

The measurement equations of our model, described further below, use the approximation in (13) to relate the annual SPF forecasts measured as in (12) to forecasts (or lagged realizations) of y_t.

To recap, the SPF provides forecasts for targets y_{t+h}, $\bar{y}_{t+h}$ (annual forecasts of non-NIPA variables), and $\hat{y}_{t+h}$ (annual forecasts of NIPA variables). To match forecasts of annual average levels and their growth rates, $t + h$ should, of course, correspond to a date in Q4, so that $\bar{y}_{t+h}$ and $\hat{y}_{t+h}$

[1] Similar transformations are also applied, for example, by Aruoba (2020) and others.

denote outcomes that are realized at the end of a calendar year. We denote the corresponding SPF forecasts collected at forecast origin t by $F_t y_{t+h}$, $F_t \bar{y}_{t+h}$, and $F_t \hat{y}_{t+h}$, respectively. At time t, we match observed forecasts from the SPF with measurement equations for $F_t y_{t+h}$, $F_t \bar{y}_{t+h}$, and/or $F_t \hat{y}_{t+h}$, for different (but separate) values of $h \geq 0$. We treat the calendar-year forecasts of the unemployment rate and CPI inflation as readings of $F_t \bar{y}_{t+h}$, while treating annual forecasts for growth in real GDP and GDP prices as data on $F_t \hat{y}_{t+h}$. The remainder of this appendix describes the details of this matching with specific examples for each variable.

I(b.2) Real GDP growth

In $\boldsymbol{Z}_{\boldsymbol{q},2024Q1}$, we have y_{2023Q4} (real GDP growth of 2023Q4) and quarterly SPF point forecasts (as of 2024Q1) targeting 2024Q1, 2024Q2, 2024Q3, 2024Q4 and 2025Q1, formally:

$$\boldsymbol{Z}_{\boldsymbol{q},2024Q1} = \left[y_{2023Q4}, F_{2024Q1} y_{2024Q1}, F_{2024Q1} y_{2024Q2}, F_{2024Q1} y_{2024Q3}, F_{2024Q1} y_{2024Q4}, F_{2024Q1} y_{2025Q1} \right]'. \tag{A.3}$$

In $\boldsymbol{Z}_{\boldsymbol{a},2024Q1}$, we have annual real GDP forecasts targeting 2025, 2026 and 2027 (note that annual forecasts are associated with the fourth quarter of the target year):

$$\boldsymbol{Z}_{\boldsymbol{a},2024Q1} = \left[F_{2024Q1} \hat{y}_{2025Q4}, F_{2024Q1} \hat{y}_{2026Q4}, F_{2024Q1} \hat{y}_{2027Q4} \right]'. \tag{A.4}$$

The loading matrix takes the following form:

$$\boldsymbol{C}_{2024\text{Q1}} = \left[\begin{array}{c} \boldsymbol{C}_{\boldsymbol{q},2024\text{Q1}} \\ \hline \boldsymbol{C}_{\boldsymbol{a},2024\text{Q1}} \end{array}\right] \tag{A.5}$$

$$= \left[\begin{array}{cccccccccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & \frac{1}{16} & \frac{2}{16} & \frac{3}{16} & \frac{4}{16} & \frac{3}{16} & \frac{2}{16} & \frac{1}{16} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{16} & \frac{2}{16} & \frac{3}{16} & \frac{4}{16} & \frac{3}{16} & \frac{2}{16} & \frac{1}{16} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{16} & \frac{2}{16} & \frac{3}{16} & \frac{4}{16} \end{array}\right], \tag{A.6}$$

where the horizontal line marks the distinction between the mapping into quarterly (upper part) and annual average (lower part) growth rates. Also, note that the loading matrix reflects the assumption that the term structure of SPF-consistent forecasts is flat beyond $H = 12$. In other words, the gaps $\tilde{Y}_{t+H+j}$ are assumed to be zero for all $j > 0$ (thus $\tilde{Y}_{2027\text{Q2}} = \tilde{Y}_{2027\text{Q3}} = \cdots = 0$), and at those horizons only the trend loading matters, hence the forecasts are set identical to the trend, with zero gaps.

When we move to the next quarter, $t = 2024\text{Q2}$, in $\boldsymbol{Z}_{\boldsymbol{q},2024\text{Q2}}$ we have $y_{2024\text{Q1}}$ (real GDP growth of 2024Q1) and quarterly SPF point forecasts (as of 2024Q2) targeting 2024Q2, 2024Q3, 2024Q4, 2025Q1 and 2025Q2, formally:

$$\boldsymbol{Z}_{\boldsymbol{q},2024\text{Q2}} = \left[y_{2024\text{Q1}}, F_{2024\text{Q2}}y_{2024\text{Q2}}, F_{2024\text{Q2}}y_{2024\text{Q3}}, F_{2024\text{Q2}}y_{2024\text{Q4}}, F_{2024\text{Q2}}y_{2025\text{Q1}}, F_{2024\text{Q2}}y_{2025\text{Q2}}\right]', \tag{A.7}$$

hence compared to equation (A.3) all quarterly forecast targets (and the lagged realization) move ahead by one quarter.

In $\boldsymbol{Z}_{\boldsymbol{a},2024\text{Q2}}$, we have annual real GDP forecasts targeting 2025, 2026 and 2027, as before, but as annual forecasts are associated with the fourth quarter of the target year, the forecast horizons shrink by one quarter relative to equation (A.4):

$$\boldsymbol{Z}_{\boldsymbol{a},2024Q2} = \left[F_{2024Q2}\hat{y}_{2025Q4}, F_{2024Q2}\hat{y}_{2026Q4}, F_{2024Q2}\hat{y}_{2027Q4} \right]' . \tag{A.8}$$

The loading matrix, $\boldsymbol{C}_{2024Q2}$, has the following components:

$$\boldsymbol{C}_{\boldsymbol{q},2024Q2} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{A.9}$$

and

$$\boldsymbol{C}_{\boldsymbol{a},2024Q2} = \begin{bmatrix} 0 & \frac{1}{16} & \frac{2}{16} & \frac{3}{16} & \frac{4}{16} & \frac{3}{16} & \frac{2}{16} & \frac{1}{16} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{16} & \frac{2}{16} & \frac{3}{16} & \frac{4}{16} & \frac{3}{16} & \frac{2}{16} & \frac{1}{16} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{16} & \frac{2}{16} & \frac{3}{16} & \frac{4}{16} & \frac{3}{16} \end{bmatrix} . \tag{A.10}$$

with

$$\boldsymbol{C}_{2024Q2} = \begin{bmatrix} \boldsymbol{C}_{\boldsymbol{q},2024Q2} \\ \hline \boldsymbol{C}_{\boldsymbol{a},2024Q2} \end{bmatrix} . \tag{A.11}$$

Compared to $\boldsymbol{C}_{2024Q1}$ in equation (A.6), in equation (A.9) we see that $\boldsymbol{C}_{\boldsymbol{q},2024Q2} = \boldsymbol{C}_{\boldsymbol{q},2024Q1}$ (i.e., the loadings associated with the lagged realization and the fixed-horizon forecasts remain unchanged), while the non-zero elements of the first two rows of $\boldsymbol{C}_{\boldsymbol{a},2024Q1}$ in equation (A.10) shift to the left, and so does the third row, with the weight $\frac{3}{16}$ appearing as the last (bottom right) entry (i.e., the loadings associated with the fixed-event forecasts change, in line with shrinking forecast horizons).

In case of a forecast origin in Q3, the measurement vector and loadings can be constructed analogously to the previous examples. However, when the forecast origin is in Q4, we need to adapt the procedures, on account of the next-year growth rate being modeled as a linear combination of growth rates in several quarters that also comprise the 2nd quarter of the current year, i.e., y_{t-2},

whereas the state vector comprises only y_{t-1}, and $F_t y_{t+h}$ for $h \geq 0$. In principle, a measure of y_{t-2} is contained in the lagged state vector. However, in our measurements, the lagged realization of GDP growth contained in $\boldsymbol{Y_t}$ is informed by the first release data available at the time t round of the SPF. Instead, when t is in Q4, the reading of y_{t-2} relevant for constructing $F_t\hat{y}_{t+4}$ should be provided by the time t vintage of GDP data. Thus, short of tracking data revisions in our state space, we adjust the data construction and measurement loading for $F_t\hat{y}_{t+4}$, when t is in Q4. Specifically, we define:

$$F^{Q4}_{t+4}\hat{y}_{t+4} \equiv \frac{16}{15} \cdot \left(F_{t+4}\hat{y}_{t+4} - \frac{y_{t-2}}{16}\right) \tag{A.12}$$

$$= \frac{1}{15}\left(y_{t+4} + 2 \cdot y_{t+3} + 3 \cdot y_{t+2} + 4 \cdot y_{t+1} + 3 \cdot y_t + 2 \cdot y_{t-1}\right). \tag{A.13}$$

Consider the example of $t =$ 2023Q4.[2] The measurement vector and associated loadings for the SPF's quarterly fixed-horizon forecasts remains unchanged relative to our earlier examples for forecast origins in Q1 and Q2. However, considering the annual fixed-event forecasts, the measurement vector becomes

$$\boldsymbol{Z_{a,}}_{\text{2023Q4}} = \left[F^{Q4}_{\text{2023Q4}}\hat{y}_{\text{2024Q4}}, F_{\text{2023Q4}}\hat{y}_{\text{2025Q4}}, F_{\text{2023Q4}}\hat{y}_{\text{2026Q4}}\right]', \tag{A.14}$$

and the associated measurement loadings are as follows:

$$\boldsymbol{C_{a,}}_{\text{2023Q4}} = \begin{bmatrix} \frac{2}{15} & \frac{3}{15} & \frac{4}{15} & \frac{3}{15} & \frac{2}{15} & \frac{1}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{16} & \frac{2}{16} & \frac{3}{16} & \frac{4}{16} & \frac{3}{16} & \frac{2}{16} & \frac{1}{16} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{16} & \frac{2}{16} & \frac{3}{16} & \frac{4}{16} & \frac{3}{16} & \frac{2}{16} & \frac{1}{16} \end{bmatrix}. \tag{A.15}$$

I(b.3) Unemployment rate

In $\boldsymbol{Z_{q,}}_{\text{2024Q1}}$, we have y_{2023Q4} (unemployment rate in 2023Q4) and quarterly SPF point forecasts (as of 2024Q1) targeting 2024Q1, 2024Q2, 2024Q3, 2024Q4 and 2025Q1, formally:

$$\boldsymbol{Z_{q,}}_{\text{2024Q1}} = \left[y_{\text{2023Q4}}, F_{\text{2024Q1}}y_{\text{2024Q1}}, F_{\text{2024Q1}}y_{\text{2024Q2}}, F_{\text{2024Q1}}y_{\text{2024Q3}}, F_{\text{2024Q1}}y_{\text{2024Q4}}, F_{\text{2024Q1}}y_{\text{2025Q1}}\right]'. \tag{A.16}$$

[2] At the time of writing, the SPF has released only forecast data through 2024Q2. However, we would expect the same definitions to continue to apply also for the 2024Q4 release.

In $\boldsymbol{Z_{a,}}_{\text{2024Q1}}$, we have annual average unemployment rate forecasts targeting 2025, 2026 and 2027 (note that annual forecasts are associated with the fourth quarter of the target year):

$$\boldsymbol{Z_{a,}}_{\text{2024Q1}} = \left[F_{\text{2024Q1}}\bar{y}_{\text{2025Q4}}, F_{\text{2024Q1}}\bar{y}_{\text{2026Q4}}, F_{\text{2024Q1}}\bar{y}_{\text{2027Q4}} \right]' . \tag{A.17}$$

The loading matrix takes the following form:

$$\boldsymbol{C}_{\text{2024Q1}} = \left[\frac{\boldsymbol{C_{q,}}_{\text{2024Q1}}}{\boldsymbol{C_{a,}}_{\text{2024Q1}}} \right] = \left[\begin{array}{cccccccccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} \end{array} \right], \tag{A.18}$$

where the horizontal line marks the distinction between the mapping into quarterly (upper part) and annual average (lower part) unemployment rates. Also, note that the loading matrix reflects the assumption that the term structure of SPF-consistent forecasts is flat beyond $H = 12$. In other words, the gaps $\tilde{Y}_{t+H+j}$ are assumed to be zero for all $j > 0$ (thus $\tilde{Y}_{\text{2027Q2}} = \tilde{Y}_{\text{2027Q3}} = \cdots = 0$), and at those horizons only the trend loading matters, hence the forecasts are set identical to the trend, with zero gaps.

When the forecast origin moves to the next quarter, $t =$ 2024Q2, in $\boldsymbol{Z_{q,}}_{\text{2024Q2}}$ we have y_{2024Q1} (unemployment rate in 2024Q1) and quarterly SPF point forecasts (as of 2024Q2) targeting 2024Q2, 2024Q3, 2024Q4, 2025Q1 and 2025Q2, formally:

$$\boldsymbol{Z_{q,}}_{\text{2024Q2}} = \left[y_{\text{2024Q1}}, F_{\text{2024Q2}}y_{\text{2024Q2}}, F_{\text{2024Q2}}y_{\text{2024Q3}}, F_{\text{2024Q2}}y_{\text{2024Q4}}, F_{\text{2024Q2}}y_{\text{2025Q1}}, F_{\text{2024Q2}}y_{\text{2025Q2}} \right]' , \tag{A.19}$$

hence compared to equation (A.16) all quarterly forecast targets (and the lagged realization) move ahead by one quarter. In $\boldsymbol{Z_{a,}}_{\text{2024Q2}}$, we have annual average unemployment rate forecasts targeting

2025, 2026 and 2027 as before, but as annual forecasts are associated with the fourth quarter of the target year, the forecast horizons shrink by one quarter relative to equation (A.17):

$$\boldsymbol{Z}_{\boldsymbol{a},2024Q2} = \left[F_{2024Q2}\bar{y}_{2025Q4}, F_{2024Q2}\bar{y}_{2026Q4}, F_{2024Q2}\bar{y}_{2027Q4} \right]' . \tag{A.20}$$

The loading matrix takes the following form:

$$\boldsymbol{C}_{2024Q2} = \left[\frac{\boldsymbol{C}_{\boldsymbol{q},2024Q2}}{\boldsymbol{C}_{\boldsymbol{a},2024Q2}} \right] = \left[\begin{array}{cccccccccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} \end{array} \right] . \tag{A.21}$$

Compared to $\boldsymbol{C}_{2024Q1}$ in equation (A.18), in equation (A.21) we see that $\boldsymbol{C}_{\boldsymbol{q},2024Q2} = \boldsymbol{C}_{\boldsymbol{q},2024Q1}$ (i.e., the loadings associated with the lagged realization and the fixed-horizon forecasts remain unchanged), while the non-zero elements of the first two rows of $\boldsymbol{C}_{\boldsymbol{a},2024Q1}$ shift to the left, and so does the third row, with the weight $\frac{1}{4}$ appearing as the last (bottom right) entry (i.e., the loadings associated with the fixed-event forecasts change, in line with shrinking forecast horizons).

I(b.4) GDP price inflation

In $\boldsymbol{Z}_{\boldsymbol{q},2024Q1}$, we have y_{2023Q4} (GDP price inflation of 2023Q4) and quarterly SPF point forecasts (as of 2024Q1) targeting 2024Q1, 2024Q2, 2024Q3, 2024Q4 and 2025Q1, formally:

$$\boldsymbol{Z}_{\boldsymbol{q},2024Q1} = \left[y_{2023Q4}, F_{2024Q1}y_{2024Q1}, F_{2024Q1}y_{2024Q2}, F_{2024Q1}y_{2024Q3}, F_{2024Q1}y_{2024Q4}, F_{2024Q1}y_{2025Q1} \right]' . \tag{A.22}$$

In $\boldsymbol{Z}_{\boldsymbol{a},2024Q1}$, we have annual GDP price inflation forecasts targeting 2025 (note that the annual forecast is associated with the fourth quarter of the target year):

$$\boldsymbol{Z}_{\boldsymbol{a},\text{2024Q1}} = \left[F_{\text{2024Q1}} \hat{y}_{\text{2025Q4}} \right]' . \tag{A.23}$$

The loading matrix takes the following form:

$$\boldsymbol{C}_{\text{2024Q1}} = \left[\frac{\boldsymbol{C}_{\boldsymbol{q},\text{2024Q1}}}{\boldsymbol{C}_{\boldsymbol{a},\text{2024Q1}}} \right] = \left[\begin{array}{ccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & \frac{1}{16} & \frac{2}{16} & \frac{3}{16} & \frac{4}{16} & \frac{3}{16} \end{array} \right], \tag{A.24}$$

where the horizontal line marks the distinction between the mapping into quarterly (upper part) and annual average (lower part) GDP price inflation rates. Also, note that the loading matrix reflects the assumption that the term structure of SPF-consistent forecasts is flat beyond $H = 5$. In other words, the gaps $\tilde{Y}_{t+H+j}$ are assumed to be zero for all $j > 0$ (thus $\tilde{Y}_{\text{2025Q3}} = \tilde{Y}_{\text{2025Q4}} = \cdots = 0$), and at those horizons only the trend loading matters, hence the forecasts are set identical to the trend, with zero gaps.

When the forecast origin moves to the next quarter, $t = $ 2024Q2, in $\boldsymbol{Z}_{\boldsymbol{q},\text{2024Q2}}$ we have y_{2024Q1} (GDP price inflation of 2024Q1) and quarterly SPF point forecasts (as of 2024Q2) targeting 2024Q2, 2024Q3, 2024Q4, 2025Q1 and 2025Q2, formally:

$$\boldsymbol{Z}_{\boldsymbol{q},\text{2024Q2}} = \left[y_{\text{2024Q1}}, F_{\text{2024Q2}} y_{\text{2024Q2}}, F_{\text{2024Q2}} y_{\text{2024Q3}}, F_{\text{2024Q2}} y_{\text{2024Q4}}, F_{\text{2024Q2}} y_{\text{2025Q1}}, F_{\text{2024Q2}} y_{\text{2025Q2}} \right]', \tag{A.25}$$

hence compared to equation (A.22) all quarterly forecast targets (and the lagged realization) move ahead by one quarter.

In $\boldsymbol{Z}_{\boldsymbol{a},\text{2024Q2}}$, we have annual GDP price inflation forecast targeting 2025 as before, but as annual forecasts are associated with the fourth quarter of the target year, the forecast horizon shrinks by one quarter relative to equation (A.23):

$$\boldsymbol{Z}_{\boldsymbol{a},\text{2024Q2}} = \left[F_{\text{2024Q2}} \hat{y}_{\text{2025Q4}} \right]' . \tag{A.26}$$

The loading matrix takes the following form:

$$\boldsymbol{C}_{2024\text{Q2}} = \left[\frac{\boldsymbol{C}_{\boldsymbol{q},2024\text{Q2}}}{\boldsymbol{C}_{\boldsymbol{a},2024\text{Q2}}}\right] = \left[\begin{array}{ccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & \frac{1}{16} & \frac{2}{16} & \frac{3}{16} & \frac{4}{16} & \frac{3}{16} & \frac{2}{16} \end{array}\right]. \tag{A.27}$$

Compared to $\boldsymbol{C}_{2024\text{Q1}}$ in equation (A.24), in equation (A.27) we see that $\boldsymbol{C}_{\boldsymbol{q},2024\text{Q2}} = \boldsymbol{C}_{\boldsymbol{q},2024\text{Q1}}$ (i.e., the loadings associated with the lagged realization and the fixed-horizon forecasts remain unchanged), while the non-zero elements of $\boldsymbol{C}_{\boldsymbol{a},2024\text{Q1}}$ shift to the left, with the weight $\frac{2}{16}$ appearing as the last (bottom right) entry (i.e., the loadings associated with the fixed-event forecast change, in line with shrinking forecast horizon).

In case of a forecast origin in Q3, the measurement vector and loadings can be constructed analogously to the previous examples. However, when the forecast origin is in Q4, such as $t =$ 2023Q4, we need to amend the data definition and measurement loadings as described above in the case of RGDP, and the measurement loadings for the (sole) annual forecast becomes:

$$\boldsymbol{C}_{\boldsymbol{a},2023\text{Q4}} = \left[\begin{array}{ccccccc} \frac{2}{15} & \frac{3}{15} & \frac{4}{15} & \frac{3}{15} & \frac{2}{15} & \frac{1}{15} & 0 \end{array}\right]. \tag{A.28}$$

I(b.5) CPI inflation

In $\boldsymbol{Z}_{\boldsymbol{q},2024\text{Q1}}$, we have $y_{2023\text{Q4}}$ (CPI inflation of 2023Q4) and quarterly SPF point forecasts (as of 2024Q1) targeting 2024Q1, 2024Q2, 2024Q3, 2024Q4 and 2025Q1, formally:

$$\boldsymbol{Z}_{\boldsymbol{q},2024\text{Q1}} = \left[y_{2023\text{Q4}}, F_{2024\text{Q1}}y_{2024\text{Q1}}, F_{2024\text{Q1}}y_{2024\text{Q2}}, F_{2024\text{Q1}}y_{2024\text{Q3}}, F_{2024\text{Q1}}y_{2024\text{Q4}}, F_{2024\text{Q1}}y_{2025\text{Q1}}\right]'. \tag{A.29}$$

In $\boldsymbol{Z}_{\boldsymbol{a},2024\text{Q1}}$, we have annual Q4/Q4 CPI inflation forecasts targeting 2025 and 2026:

$$\boldsymbol{Z}_{\boldsymbol{a},2024\text{Q1}} = \left[F_{2024\text{Q1}}\bar{y}_{2025\text{Q4}}, F_{2024\text{Q1}}\bar{y}_{2026\text{Q4}}\right]'. \tag{A.30}$$

The loading matrix takes the following form:

$$\boldsymbol{C}_{2024Q1} = \left[\begin{array}{c} \boldsymbol{C}_{\boldsymbol{q},2024Q1} \\ \hline \boldsymbol{C}_{\boldsymbol{a},2024Q1} \end{array}\right] = \left[\begin{array}{cccccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} \end{array}\right], \tag{A.31}$$

where the horizontal line marks the distinction between the mapping into quarterly (upper part) and Q4/Q4 (lower part) CPI inflation rates. Also, note that the loading matrix reflects the assumption that the term structure of SPF-consistent forecasts is flat beyond $H = 8$. In other words, the gaps $\tilde{Y}_{t+H+j}$ are assumed to be zero for all $j > 0$ (thus $\tilde{Y}_{2026Q2} = \tilde{Y}_{2026Q3} = \cdots = 0$), and at those horizons only the trend loading matters, hence the forecasts are set identical to the trend, with zero gaps.

When the forecast origin moves to the next quarter, $t = 2024Q2$, in $\boldsymbol{Z}_{\boldsymbol{q},2024Q2}$ we have y_{2024Q1} (CPI inflation of 2024Q1) and quarterly SPF point forecasts (as of 2024Q2) targeting 2024Q2, 2024Q3, 2024Q4, 2025Q1 and 2025Q2, formally:

$$\boldsymbol{Z}_{\boldsymbol{q},2024Q2} = \left[y_{2024Q1}, F_{2024Q2}y_{2024Q2}, F_{2024Q2}y_{2024Q3}, F_{2024Q2}y_{2024Q4}, F_{2024Q2}y_{2025Q1}, F_{2024Q2}y_{2025Q2}\right]', \tag{A.32}$$

hence compared to equation (A.29) all quarterly forecast targets (and the lagged realization) move ahead by one quarter.

In $\boldsymbol{Z}_{\boldsymbol{a},2024Q2}$, we have annual Q4/Q4 CPI inflation forecasts targeting 2025 and 2026 as before, but the forecast horizon shrinks by one quarter relative to equation (A.30):

$$\boldsymbol{Z}_{\boldsymbol{a},2024Q2} = \left[F_{2024Q2}\bar{y}_{2025Q4}, F_{2024Q2}\bar{y}_{2026Q4}\right]'. \tag{A.33}$$

The loading matrix, $\boldsymbol{C}_{2024Q2}$, has the following components:

$$\boldsymbol{C}_{\boldsymbol{q},\text{2024Q2}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{A.34}$$

and

$$\boldsymbol{C}_{\boldsymbol{a},\text{2024Q2}} = \begin{bmatrix} 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} \end{bmatrix}, \tag{A.35}$$

with

$$\boldsymbol{C}_{\text{2024Q2}} = \begin{bmatrix} \boldsymbol{C}_{\boldsymbol{q},\text{2024Q2}} \\ \hline \boldsymbol{C}_{\boldsymbol{a},\text{2024Q2}} \end{bmatrix}. \tag{A.36}$$

Compared to $\boldsymbol{C}_{\text{2024Q1}}$ in equation (A.31), in equation (A.34) we see that $\boldsymbol{C}_{\boldsymbol{q},\text{2024Q2}} = \boldsymbol{C}_{\boldsymbol{q},\text{2024Q1}}$ (i.e., the loadings associated with the lagged realization and the fixed-horizon forecasts remain unchanged), while the non-zero elements of the first row of $\boldsymbol{C}_{\boldsymbol{a},\text{2024Q1}}$ in equation (A.34) shift to the left, and so does the second row, with the weight $\frac{1}{4}$ appearing as the last (bottom right) entry (i.e., the loadings associated with the fixed-event forecasts change, in line with shrinking forecast horizons).

I(c) Measurement error in annual forecasts

This appendix serves to motivate our choice for modeling data on annual SPF forecasts with measurement error. To preview the main arguments of this appendix:

- In general, the goal of this paper is to treat observed SPF data largely as is, and we maintain the assumption that treats quarterly fixed-horizon forecasts from the SPF as observed without error.

- However, when jointly using data on quarterly and annual forecasts from the SPF, overlap in their respective forecast targets raises the question of whether both types of readings are perfectly consistent with each other. In case of next-year forecasts collected in Q4, there is perfect overlap with fixed-horizon forecasts collected for $h = 1, 2, 3, 4$, and it is straightforward to check whether both readings are a perfect match. As illustrated below, quarterly forecasts collected in Q4 are indeed often close to the corresponding next-year forecast reported by the SPF. However, deviations have not been infrequent either, and at times even quite sizable.

- Moreover, as argued below, even small inconsistencies can lead to outsized effects on imputed quarterly forecasts at longer horizons, as quarterly forecasts for the near term are taken as given from the SPF.

I(c.1) Observed inconsistencies in next-year forecasts collected in Q4

In Q4, the SPF reports a set of quarterly forecasts for $h = 0, 1, 2, 3, 4$ as well as an annual forecast for the next year. Evidently, the quarterly forecasts perfectly overlap with the next-year forecast. Here we assess how well the reported next-year forecasts match what is implied by the observed quarterly SPF forecasts (and lagged data if necessary).

To match quarterly SPF forecasts with the next year forecast, recall that, in case of the unemployment rate the annual forecast is directly defined as the arithmetic average of unemployment over the four quarters of the year. In case of CPI, the same holds approximately (since the annual forecast is to reflect Q4/Q4 inflation, measured as a simple growth rate). As discussed above in Appendix I(b), for the NIPA variables (GDP growth and inflation in GDP prices), the annual forecast reflects an (approximate) linear combination of growth in the four quarters of the targeted year, as well as the last three quarters of the previous year. Thus, in the case of NIPA variables, when matching the next-year forecasts with quarterly forecasts collected in Q4, we also need to utilize the nowcast, as well as two lags of data.

Figure A.1 plots the difference between next year forecasts collected in Q4, and what is implied for the same forecast target by the SPF's quarterly fixed-horizon forecast that were collected jointly with the annual forecast. All data definitions follow the procedures employed in our model estimates, as described in Appendix I(b) above. Formally, according to the data definitions em-

ployed in our paper, annual forecast targets are linear combinations of quarterly values, and we can generically write:

$$\hat{y}_t = \sum_{j=0}^{J-1} w_j y_{t-j} \,, \tag{A.37}$$

with $J = 7$ and $w_0 = 1/16$, $w_1 = 2/16$, $w_2 = 3/16$, $w_3 = 4/16$, $w_4 = 3/16$, $w_5 = 2/16$, and $w_6 = 1/16$ in case of GDP growth and GDP price inflation, and $J = 4$ and $w_j = 1/4$ in case of the unemployment rate and CPI inflation.[3] Figure A.1 then plots the difference between the actual next-year forecast, denoted $F_t^o \hat{y}_{t+4}$, and the implied value constructed from the quarterly forecasts, $F_t \hat{y}_{t+4} \equiv \sum_{j=0}^{J-1} w_j F_t^o y_{t+4-j}$. The superscript "'o" in F_t^o denotes observed SPF data (or, if needed, lagged realized data known by the SPF respondents in real time).[4]

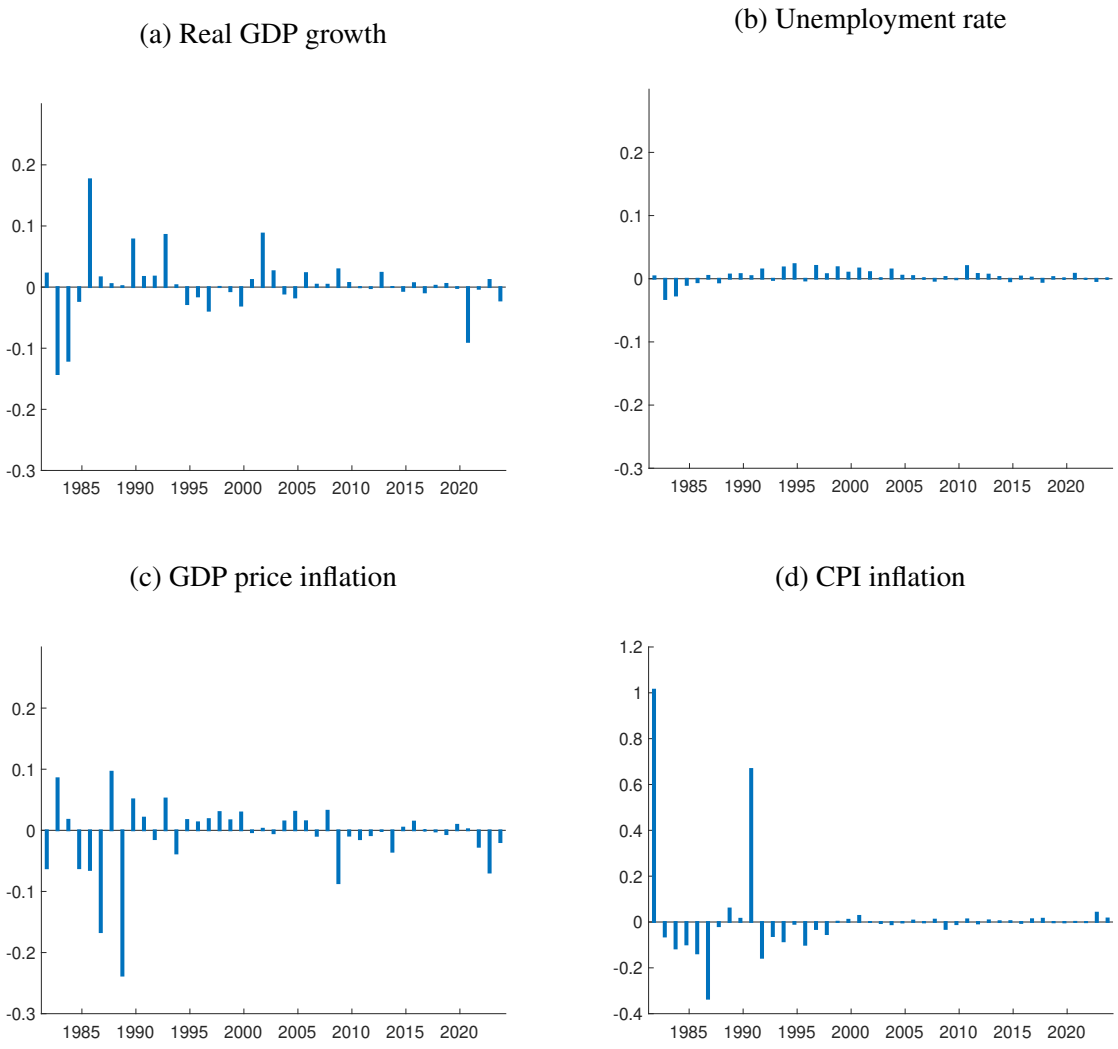
Overall, Figure A.1 shows that there are indeed some differences between the actual next-year forecast from the SPF and what is implied by the corresponding quarterly SPF predictions. For many observations of all variables, the differences are quite small (less than 10 basis points (bp) in absolute value), but there are also cases when the differences are much larger than that. The occurrence of sizable differences varies across variables. For the unemployment rate, differences are always well below 5bp. For CPI inflation, there are notable differences in the 1980s and 1990s, with two particularly large outliers in 1981 and 1990, of about 100 and 70 basis points, respectively. For GDP growth, and GDP price inflation, differences do not exceed 30bp in absolute value, but there are still some observations with differences of 20bp or more, in particular prior to 1990. All in all, this suggests that there are some measurement issues and that these are quite unevenly distributed across time and variables. As we will argue next, even small inconsistencies can lead to outsized effects on imputed quarterly forecasts at longer horizons, as quarterly forecasts for the near term are taken as given from the SPF.

Consider the case of a forecast origin in Q3, henceforth denoted t. In this case, there is large overlap between the observed SPF forecast for next year, $F_t^o \hat{y}_{t+5}$, and what is covered by the equally observed quarterly forecasts, $F_t^o y_{t+h}$ for $h = 0, 1, 2, 3, 4$. In fact, quarterly forecasts for $h \leq 4$ (and lagged data) account for every component of $F_t \hat{y}_{t+5}$ except for the forecast of Q4 next

[3] The notation adopted here generalizes a little the more specific notation for $\hat{y}_t$ and $\bar{y}_t$ as adopted in the paper. At slight abuse of our earlier notation, we denote here annual forecasts for all variables simply by $\hat{y}_t$, without distinguishing further between $\hat{y}_t$ and $\bar{y}_t$.

[4] Lagged realized data is needed to construct $F_t \hat{y}_{t+4}$ in case of GDP growth and GDP price inflation, for which we have $J = 7$. Real time vintages that correspond to SPF rounds are obtained from the Philadelphia Fed.

Figure A.1: Observed inconsistencies between quarterly and next-year forecasts collected in Q4



Notes: Difference between actual and implied SPF annual forecasts for next year collected in Q4. Implied SPF forecasts constructed from available quarterly forecast for one to four-quarters ahead, as well as lagged data (as needed for GDP growth and GDP price inflation). Only Q4 observations since 1981Q4. (The SPF began reporting next-year forecasts in 1981Q3.)

year, $F_t y_{t+5}$. In this case, a measurement equation of the form in (A.37) already determines the imputed value for $F_t \hat{y}_{t+5}$ without any further need for modeling the stochastic evolution of data and SPF:

$$\hat{F}_t y_{t+5} = \frac{1}{w_0} \cdot \left(F_t \hat{y}^o_{t+5} - \sum_{j=1}^{J-1} w_j F^o_t y_{t+5-j} \right) , \qquad \text{when } t \text{ in Q3} , \tag{A.38}$$

where the "hat" in $\hat{F}_t$ denotes an imputed value. If the assumption is correct that annual SPF forecasts map into the SPF's quarterly forecasts without error, an imputation as in (A.38) should

hold exactly. However, as we have seen in Figure A.1 for the case when t is in Q4, this is not always the case, casting doubt over the applicability of (A.38) when t is in Q3.

Instead, consider the case where the observed annual forecast is measured only with some (potentially) small error, while we maintain the assumption that quarterly forecasts are observed without error (so that $F_t y_{t+j} = F_t^o y_{t+j}$ for $j \leq 4$). In this case, we can write the measurement equation for the annual forecast as follows:

$$F_t^o \hat{y}_{t+5} = \sum_{j=0}^{J-1} w_j F_t y_{t+5-j} + n_t. \tag{A.39}$$

But, application of an imputation, as in (A.38), that assumes the absence of measurement error distorts the imputed value by the measurement error:

$$\Rightarrow \quad \hat{F}_t y_{t+5} = \frac{1}{w_0} \cdot \left(F_t \hat{y}_{t+5}^o - \sum_{j=1}^{J-1} w_j F_t^o y_{t+5-j} \right), \qquad \text{when } t \text{ in Q3}, \tag{A.40}$$

$$= F_t y_{t+5} + \frac{1}{w_0} n_t \,. \tag{A.41}$$

With $1/w_0 = 16$ (for GDP growth and GDP inflation) or $1/w_0 = 4$ (for the unemployment rate and CPI inflation), even small measurement errors in the annual forecast can lead to large distortions in the imputed value for the forecast made in Q3 this year (t) for the quarterly outcome in Q4 next year.

By a similar logic, observations in Q1 and Q2 for the SPF's next-year and quarterly forecasts directly restrict a weighted average of imputed values for the quarterly forecasts for the next year. When the imputations are made based on an error-free measurement equation as in (A.37), while the actual data is affected by measurement error, these restrictions on imputed values are again distorted by measurement error:

$$\text{When } t \text{ in Q2:} \quad \hat{\alpha}_{t,5:6} \equiv \frac{1}{w_0} \cdot \left(F_t \hat{y}_{t+6}^o - \sum_{j=2}^{J-1} w_j F_t^o y_{t+6-j} \right) \tag{A.42}$$

$$= \frac{w_0 F_t y_{t+6} + w_1 F_t y_{t+5}}{w_0 + w_1} + \frac{n_t}{w_0 + w_1} \,. \tag{A.43}$$

$$\text{And, when } t \text{ in Q1:} \quad \hat{\alpha}_{t,5:7} = \frac{1}{w_0} \cdot \left(F_t \hat{y}_{t+7}^o - \sum_{j=3}^{J-1} w_j F_t^o y_{t+7-j} \right) \tag{A.44}$$

$$= \frac{w_0 F_t y_{t+7} + w_1 F_t y_{t+6} + w_2 F_t y_{t+5}}{w_0 + w_1 + w_2} + \frac{n_t}{w_0 + w_1 + w_2} \,. \tag{A.45}$$

Notably, the weight on the measurement error in these calculations is increasing for imputations made later in a given calendar year.

For simplicity, we call $\hat{F}_t y_{t+5}$, $\hat{\alpha}_{t,5:6}$, and $\hat{\alpha}_{t,5:7}$ "naive" imputations. Figure A.2 plots these naive imputations against data for realized outcomes and the observed next-year SPF. Overall, the naive imputations appear to track (or predict) the data fairly well. Of note, $\hat{\alpha}_{t,5:6}$, and $\hat{\alpha}_{t,5:7}$ should reflect (weighted) averages of quarterly outcomes, which could be expected to be less volatile than quarterly outcomes or forecasts thereof. Indeed, imputations made in Q3 (yellow diamonds) tend to stand out more often than those made in Q1 and Q2 (orange squares and blue circles, respectively). More importantly, Figure A.2 shows some patterns that are reminiscent of what is shown in Figure A.1 for inconsistencies between observed and implied values for the next-year forecast at Q4 origins (which are a direct reflection of measurement error): While the naive imputations track the data particularly well for the unemployment rate and CPI inflation, imputations made in Q3 for GDP growth and GDP inflation show notable outliers, which could be indicative of the measurement error term in (A.41), and with decreasing effect for values constructed in Q2 and Q1, as predicted by equations (A.43) and (A.45).

I(c.2) Excessively volatile imputations when treating annual forecasts without error

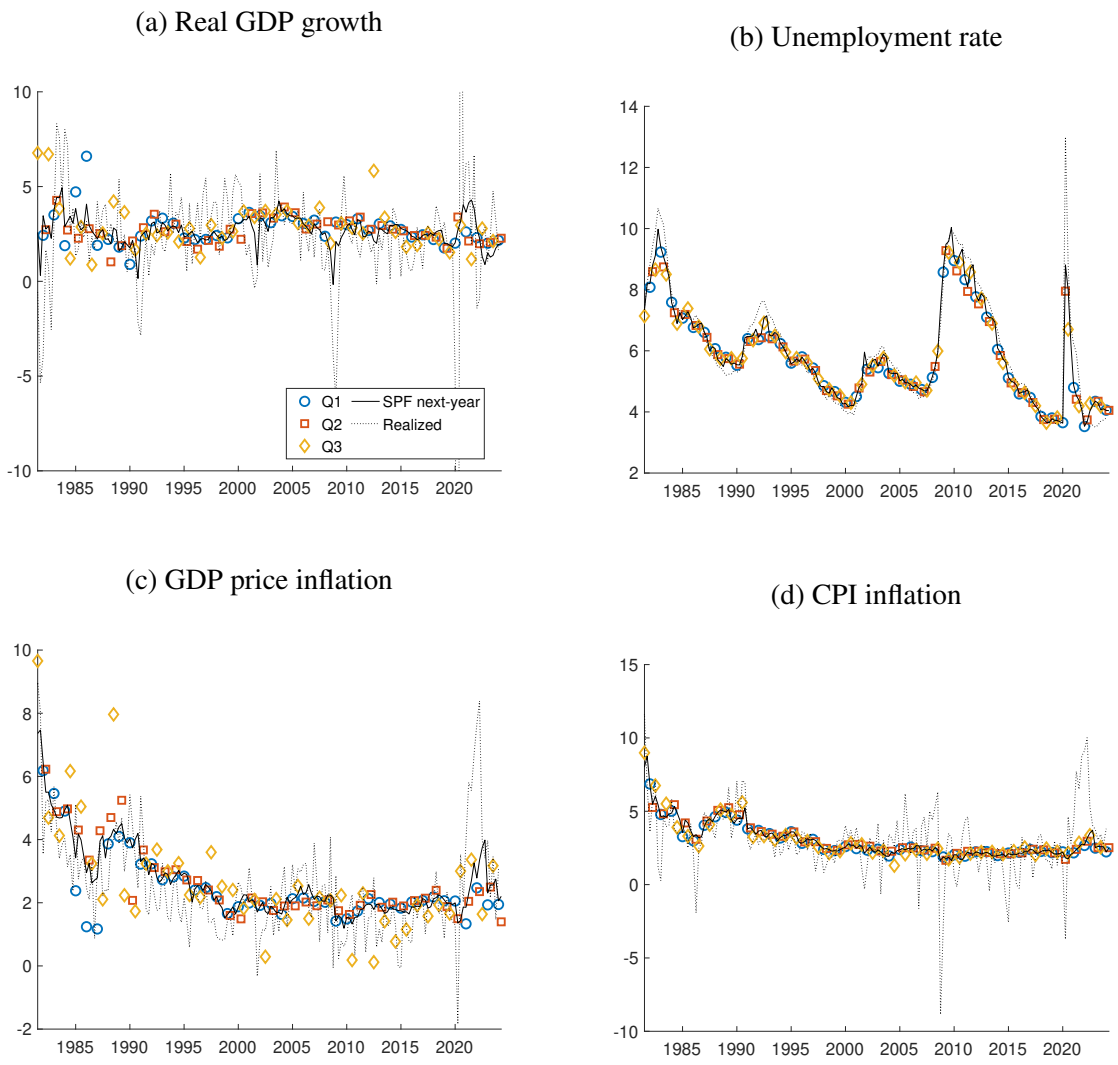
Of course, the arguments presented so far can only highlight the particular risk of distortions in imputed values when annual forecasts are not perfectly consistent with quarterly forecasts. And, the results shown in Figures A.1 and A.2 are at best indicative of the extent to which such inconsistencies may be relevant in the data. At least for SPF rounds in Q4, the data suggests that inconsistencies cannot always be neglected.

As such, we evaluated models that allow for measurement error in annual forecasts, as described in Section 4 of the paper, as well as versions that assume all SPF data are observed without error. The latter case was also the basis for earlier versions of our manuscript. In light of Figure A.1, we drop Q4 observations from the estimation data for models that assume annual forecasts are observed without error.[5]

Strikingly, term structures of SPF-consistent forecasts that we imputed from models without measurement error are notably more volatile than those imputed from models that allow for mea-

[5] In results not shown, we also evaluated models that drop Q4 observations from the estimation data for models that allow for measurement error in annual forecasts. The results were largely similar to those presented in the paper.

Figure A.2: Naive imputations of quarterly forecasts at longer horizons



Notes: "Naive" imputations, $\hat{F}_t y_{t+5}$, $\hat{\alpha}_{t,5:6}$, and $\hat{\alpha}_{t,5:7}$, of quarterly forecasts at longer horizons as defined in equations (A.41), (A.43), and (A.45), respectively. Imputations made in Q1, Q2, and Q3 as indicated. Observations since 1981Q3, which is when the SPF reported its first next-year forecast. For sake of readability of SPF information, COVID-19 outliers in realized data are not shown.

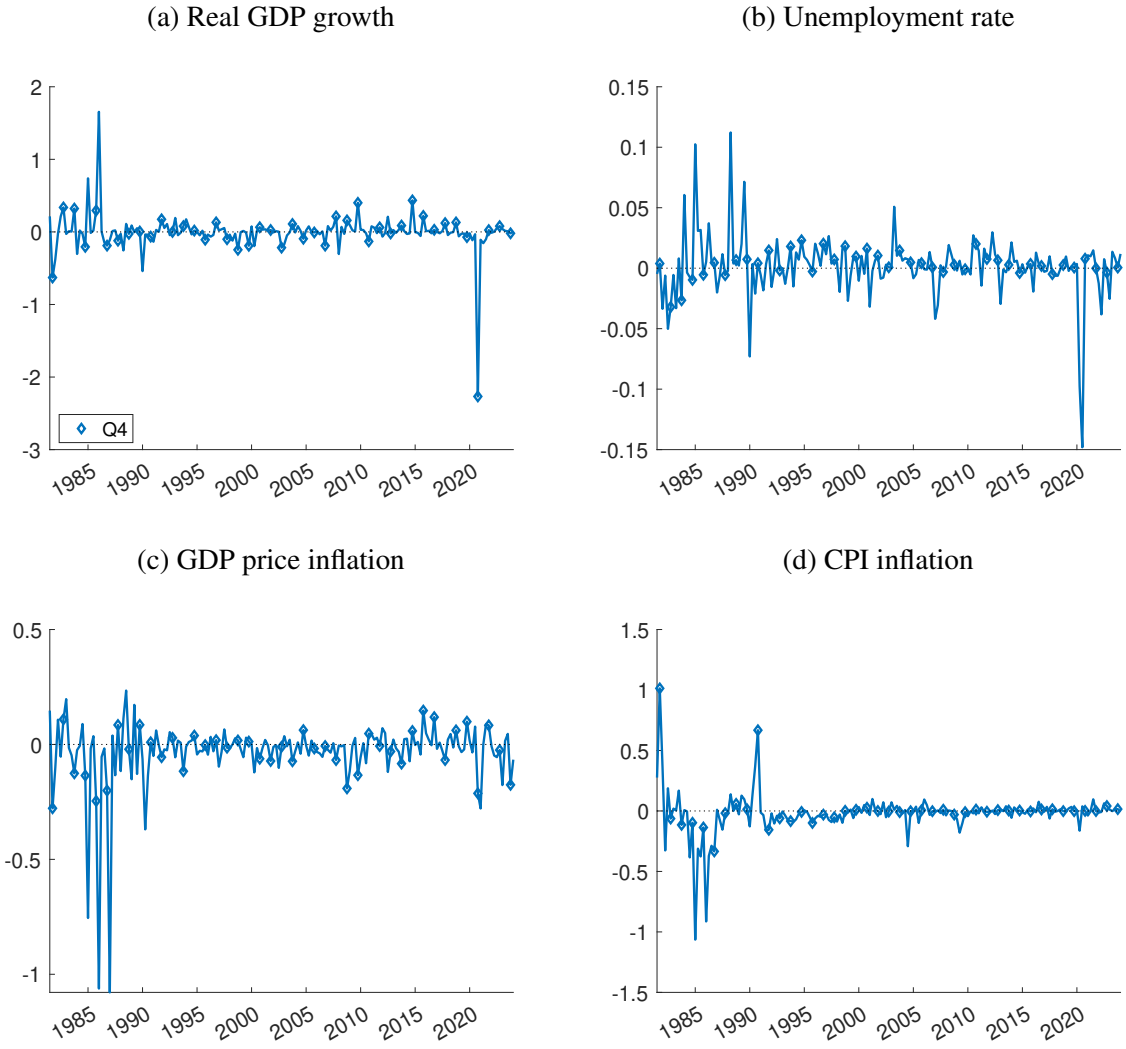
surement error in annual forecasts. This is particularly true for GDP growth and GDP price inflation, where the volatility of imputed forecasts at longer horizons is much higher in models that assume no measurement error in annual forecasts. Of course, these findings support the argument that even small inconsistencies between quarterly and annual forecasts can lead to outsized effects on imputed quarterly forecasts at longer horizons, as quarterly forecasts for the near term are taken as given from the SPF. Moreover, in feedback received from forecasting practitioners, the more "wiggly" imputations obtained from models that omit measurement error were considered to be less credible. In a similar vein, Table A.3 shows that one-step ahead forecasts of SPF data

generated by a version of our model that assumes no measurement error in annual forecasts are dramatically worse in predicting future survey data than our baseline model which does assume measurement error (with results shown in Table 1 of the paper and restated as Table A.4 below).

All told, these results lead us to prefer — and adopt in the revised paper — models that allow for measurement error in observed annual SPF forecasts. Given the occasional nature of inconsistencies that are directly detectable (Figure A.1) and related patterns in naive imputations made in different quarters of the year (Figure A.2), we chose (1) to specify separate measurement error processes for data observed in different quarters of the year, and (2) to adopt fat-tailed specifications for the measurement errors, which place much mass a priori on errors being zero, while retaining the flexibility to fit occasionally sizable occurrences of errors. Details of the measurement error specifications are described in Section 4 of the paper and Appendix II below. Throughout, we maintain the specification that quarterly SPF forecasts are observed without error, in keeping with the goal of this paper to treat observed SPF data largely as is.

Figures A.3 through A.6 report time series of estimates of the measurement noise in annual forecasts — specifically, posterior medians in observed SPF forecasts for the next calendar year ahead up through three years ahead (for variables with forecasts at these horizons), for both the MDS and VAR specifications. These estimates show that, in keeping with our choice of a horseshoe specification of the shocks' distributions, the noise shocks are usually small, but very occasionally large. This pattern is especially stark for the next-year forecasts. The relatively very large noise shocks in Figure A.3 tend to occur in the instances of large inconsistencies between quarterly and next-year forecasts indicated in Figure A.1, which are primarily early in the sample and to a lesser extent around the time of the outbreak of the COVID-19 pandemic. However, in keeping with the logic described above with imputation issues, the size of the noise shocks tends to be generally larger than the size of the inconsistencies; this naturally stems from the large weights that some quarters of forecasts can get due to the weights of the Mariano-Murasawa approximation of annual GDP growth and GDP price inflation. Another evident pattern in the estimates is that, except in the case of the unemployment rate (for which noise shocks are generally small) the noise shocks tend to be larger (in absolute value) at the year-ahead horizon than longer horizons. To the extent that the measurement error is linked to inconsistencies in annual and observed quarterly forecasts, this is to be expected, given that quarterly forecasts are only observed at shorter horizons and therefore

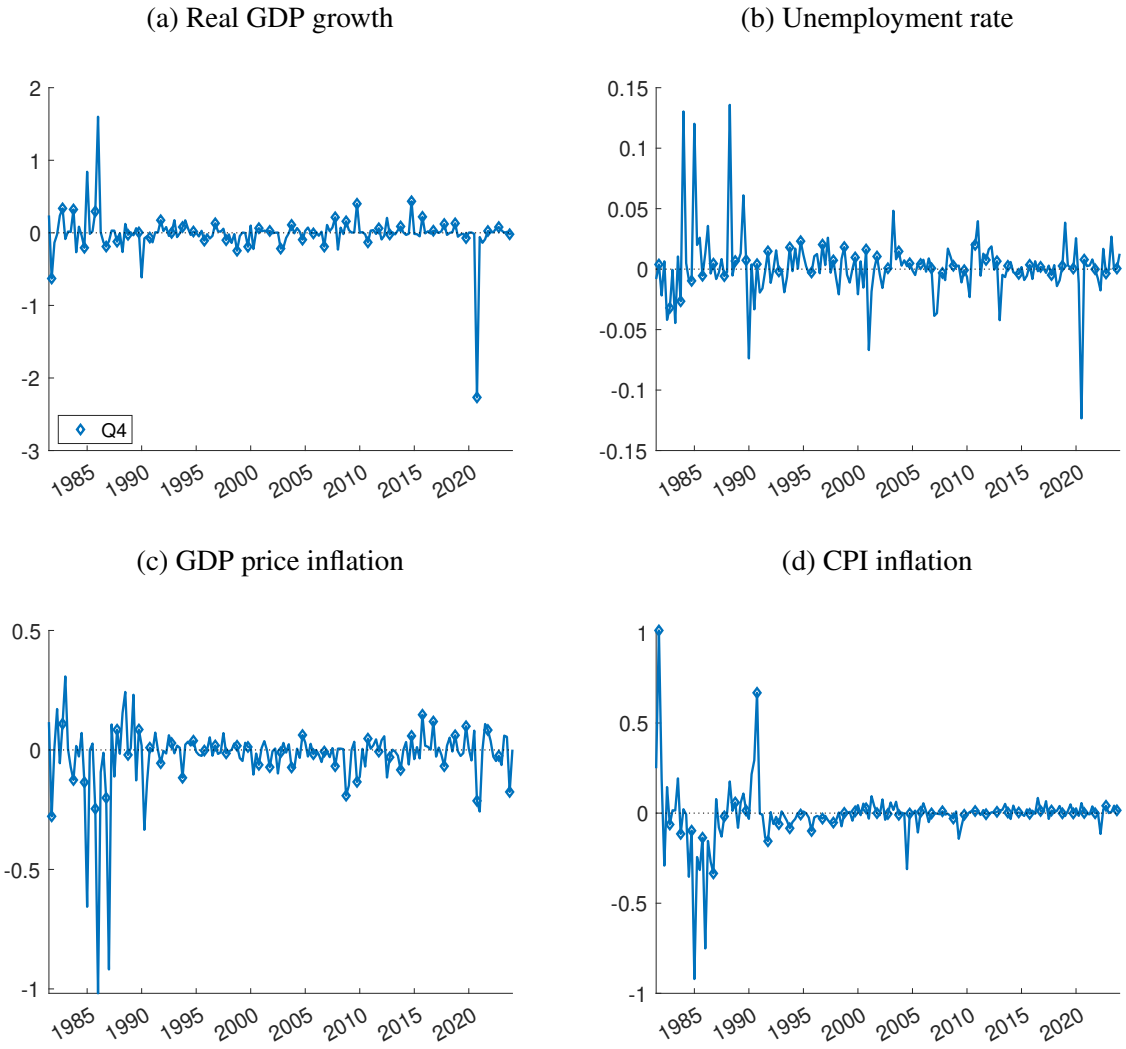
Figure A.3: Estimated noise in SPF next-year forecasts (MDS, 2024Q1)



Notes: Posterior medians of noise levels in observed SPF forecasts for the next calendar year ahead. Q4 observations marked by a diamond. Estimates from the MDS model using data through 2024Q1.

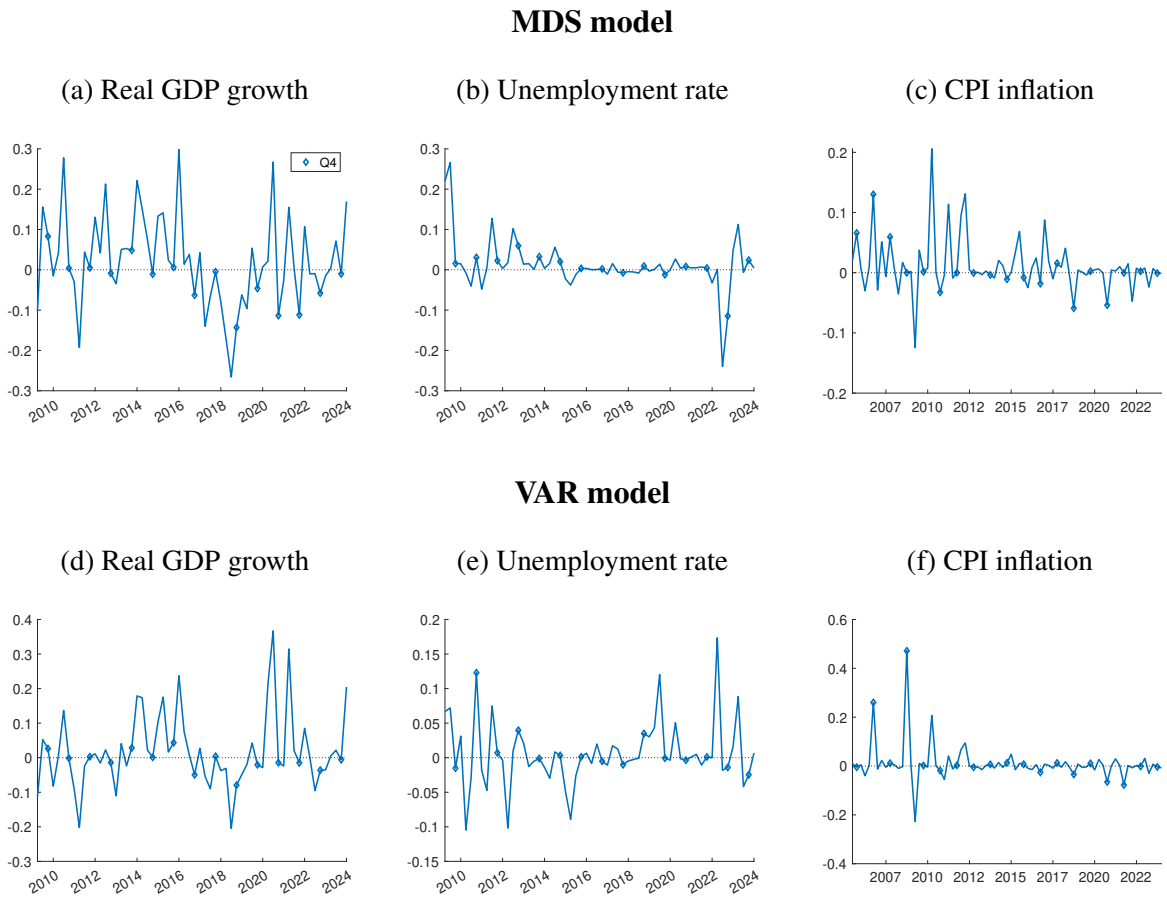
consistency is only an issue for the year-ahead forecasts. Finally, the estimated shocks are similar for the MDS and VAR specifications, most clearly and strongly for the year-ahead horizon.

Figure A.4: Estimated noise in SPF next-year forecasts (VAR, 2024Q1)



Notes: Posterior medians of noise levels in observed SPF forecasts for the next calendar year ahead. Q4 observations marked by a diamond. Estimates from the VAR model using data through 2024Q1.

Figure A.5: Estimated noise in SPF two-years ahead forecasts (2024Q1)

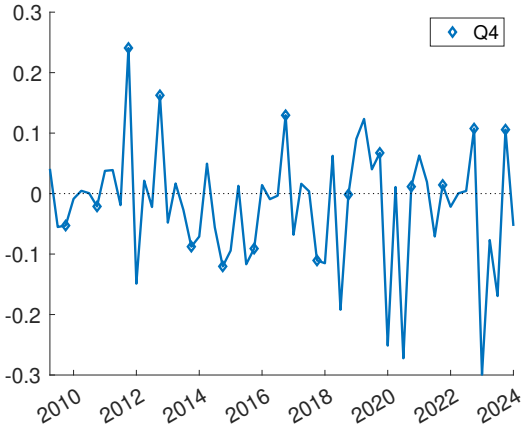


Notes: Posterior medians of noise levels in observed SPF forecasts for the next calendar year ahead. Q4 observations marked by a diamond. Estimates based on data through 2024Q1.

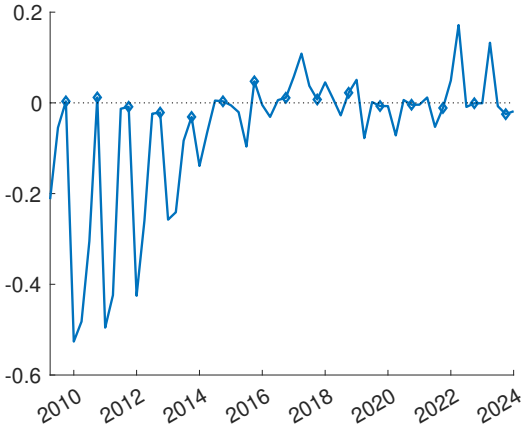
Figure A.6: Estimated noise in SPF three-years ahead forecasts (2024Q1)

MDS model

(a) Real GDP growth

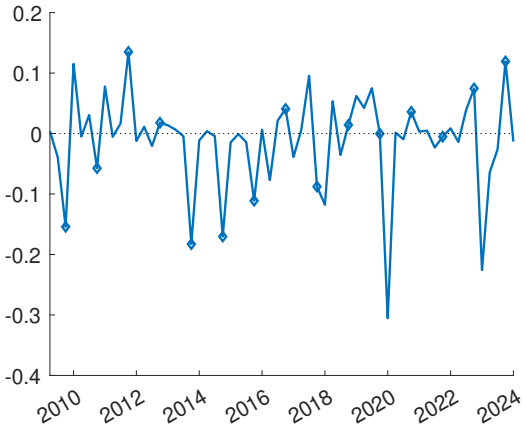


(b) Unemployment rate

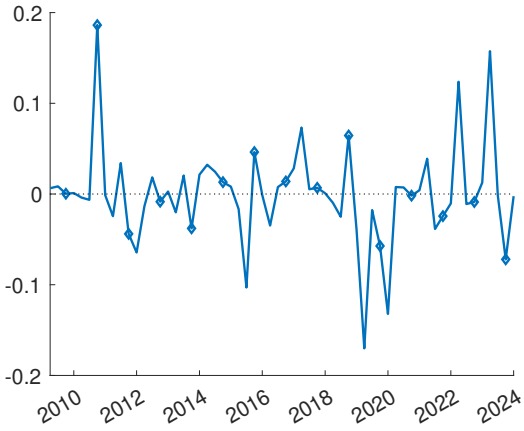


VAR model

(c) Real GDP growth



(d) Unemployment rate



Notes: Posterior medians of noise levels in observed SPF forecasts for the next calendar year ahead. Q4 observations marked by a diamond. Estimates based on data through 2024Q1.

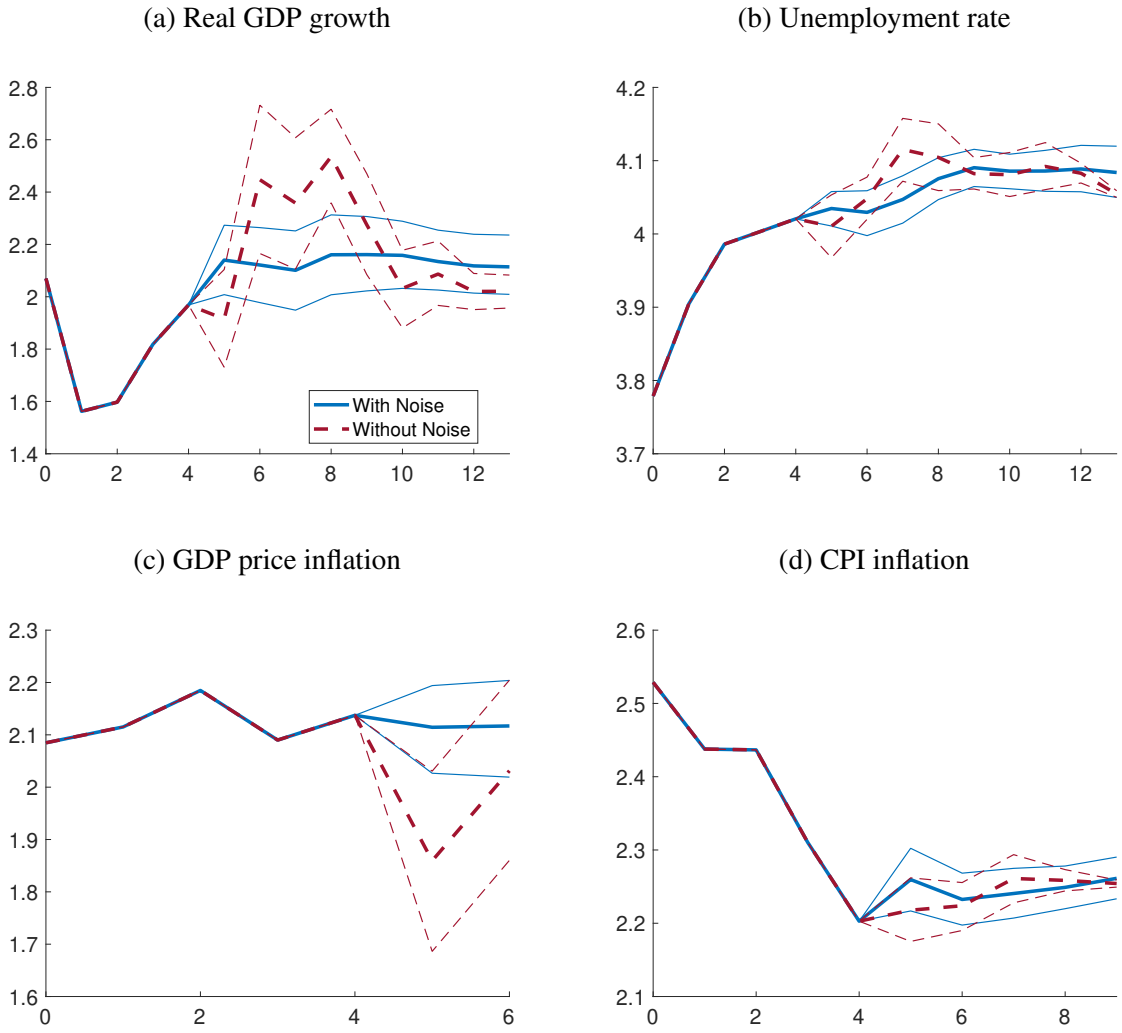
I(c.3) Comparison of results from models with and without noise in annual forecasts

To show how including measurement error around published annual forecasts impacts our forecasts and results, Figures A.7 through A.10 report examples of term structures of (out-of-sample) quarterly forecasts from models with and without measurement noise. Note that these charts end with H as specified for each variable; to facilitate comparisons, we do not report the forecasts out to 16 steps ahead (recall that the point forecasts for $h = H + 1, \ldots, 16$ are equal to the forecasts for $h = H$). In addition, by construction, without noise on short-horizon forecasts, the forecasts for $h = 0, \ldots, 4$ are the same across the noise and no-noise models.

Focusing first on results from MDS specifications, without noise, the quarterly forecasts at longer horizons show more variation from quarter to quarter than do the forecasts from the model with measurement noise in annual forecasts. In particular, the forecasts from the model with noise avoid the tendency of the forecasts from the model without noise to change one way early on and then snap in the opposite direction in the following few quarters. In some instances — e.g., GDP price inflation in the 2024Q1 example and unemployment in the 2019Q4 example — the inclusion of noise in the model can impact the level of the longer-horizon quarterly forecasts as compared to the model without noise. It is also evident that the inclusion of measurement noise on annual forecasts can have some impact on the uncertainty around the estimated quarterly forecast at longer horizons. This is evident in the case of GDP price inflation, with greater uncertainty around the forecasts from the model without noise than the model with noise. However, as noted in the paper, uncertainty around the latent quarterly forecast estimates at longer horizons is small relative to the overall forecast uncertainty reflected in the size of historical forecast errors. While not shown in the interest of brevity, the inclusion of noise in the model does not have much impact on overall forecast uncertainty: The widths of forecast confidence bands are comparable for the with-noise and without-noise model specifications.

Patterns are very broadly similar in forecasts from VAR specifications that allow for bias in SPF forecasts. However, the inclusion of noise around published annual forecasts has a smaller impact with the VAR than the MDS specification. For example, the with-noise and without-noise forecasts of GDP growth are quite similar in the VAR case; the with-noise forecasts are not smoother than the without-noise forecasts as observed in the MDS case. In turn, the forecasts of GDP growth

Figure A.7: Term structures estimated with and without noise (MDS, 2024Q1)

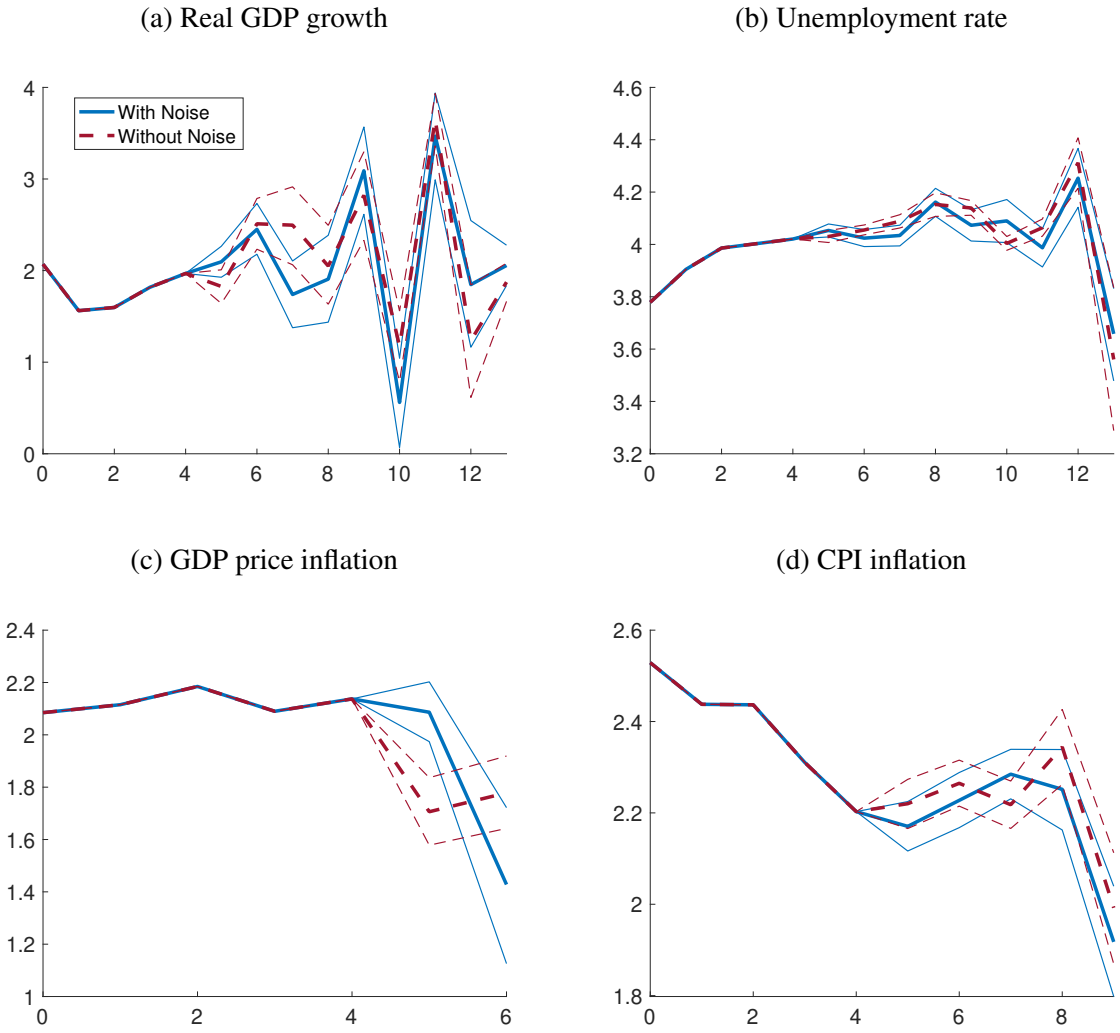


Notes: Posterior medians (and 68% bands) of term structures of SPF-consistent expectations, denoted $\boldsymbol{Y_t}$, obtained from the MDS model with or without noise in the observed SPF calendar-year forecasts. For estimation of the mode without noise, we drop Q4 observations for the next-year forecast (due to perfect overlap with the observed quarterly SPF forecasts.) Estimates based on data through 2024Q1.

from the VAR (in both the with-noise and without-noise estimates) show more variability across quarters than do the forecasts from the MDS specification with noise.

To further assess the role of measurement noise in our models and results, Tables A.3 and A.4 provide the estimated intercepts and slopes of Mincer-Zarnowitz regressions of SPF forecasts published in quarter $t+1$ on SPF forecasts estimated from our model using SPF forecasts up through quarter t. Table A.3 reports results from MDS and VAR specifications without measurement noise; Table A.4 provides corresponding results from specifications with measurement noise (the results also shown in the paper). In these results, the forecasts from models without noise are somewhat

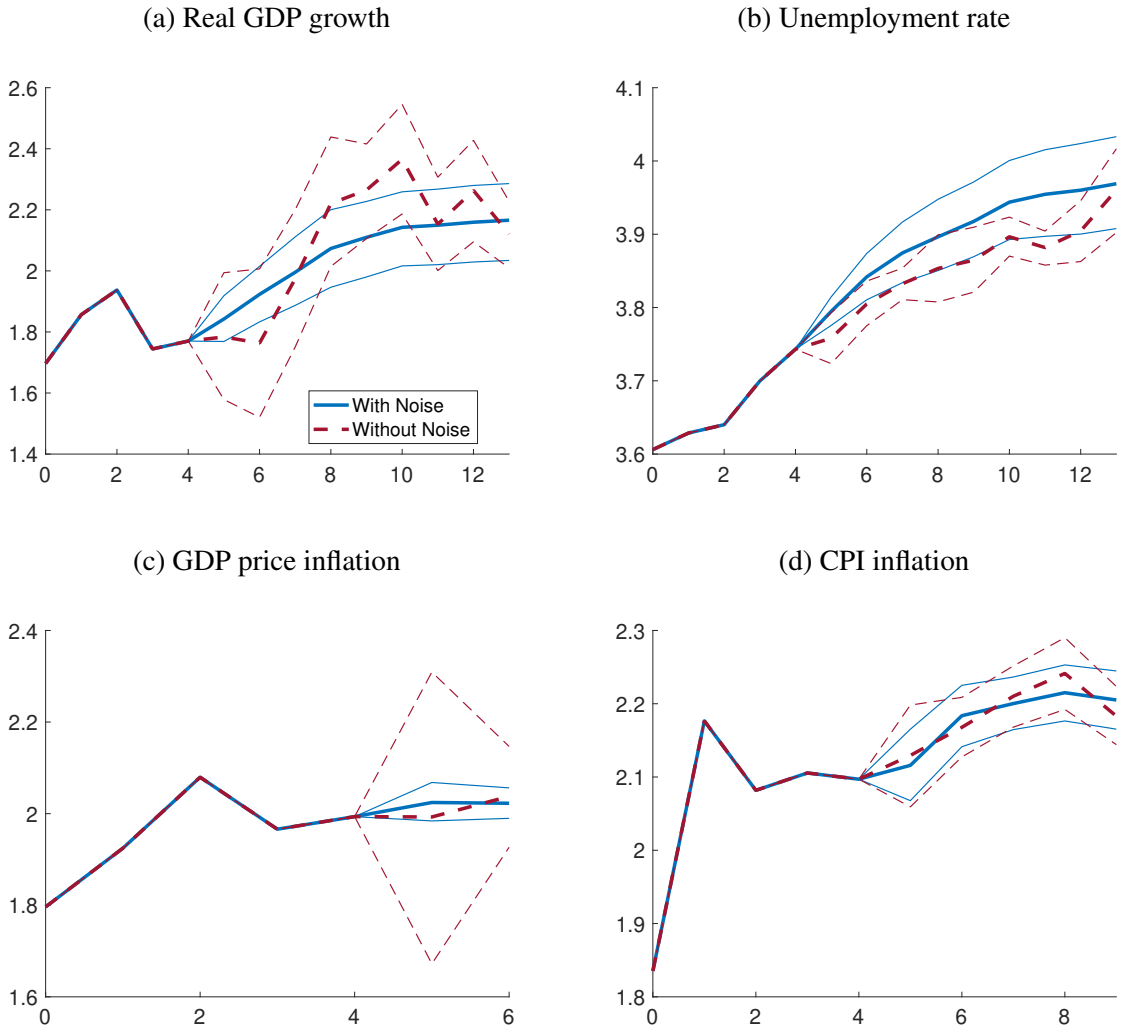
Figure A.8: Term structures estimated with and without noise (VAR, 2024Q1)



Notes: Posterior medians (and 68% bands) of term structures of SPF-consistent expectations, denoted $\boldsymbol{Y}_t$, obtained from the VAR model with or without noise in the observed SPF calendar-year forecasts. For estimation of the mode without noise, we drop Q4 observations for the next-year forecast (due to perfect overlap with the observed quarterly SPF forecasts.) Estimates based on data through 2024Q1.

less efficient predictions of future SPF forecasts, in particular for annual SPF forecasts, as well as the four-quarters-ahead SPF. Overall, there are more rejections of slope coefficients of unity in the no-noise forecasts than the with-noise forecasts. Related, in some cases, the no-noise forecasts yield noticeably lower slope coefficients than the with-noise forecasts, especially with the VAR and less so with the MDS specification. For example, with GDP growth (PGDP inflation) at the four-quarters-ahead horizon ($h = 4$), the slope coefficient estimate is 0.39 (0.47) in the noise-free MDS forecasts and 0.87 (0.91) in the with-noise MDS forecasts. As another example, with GDP growth

Figure A.9: Term structures estimated with and without noise (MDS, 2019Q4)

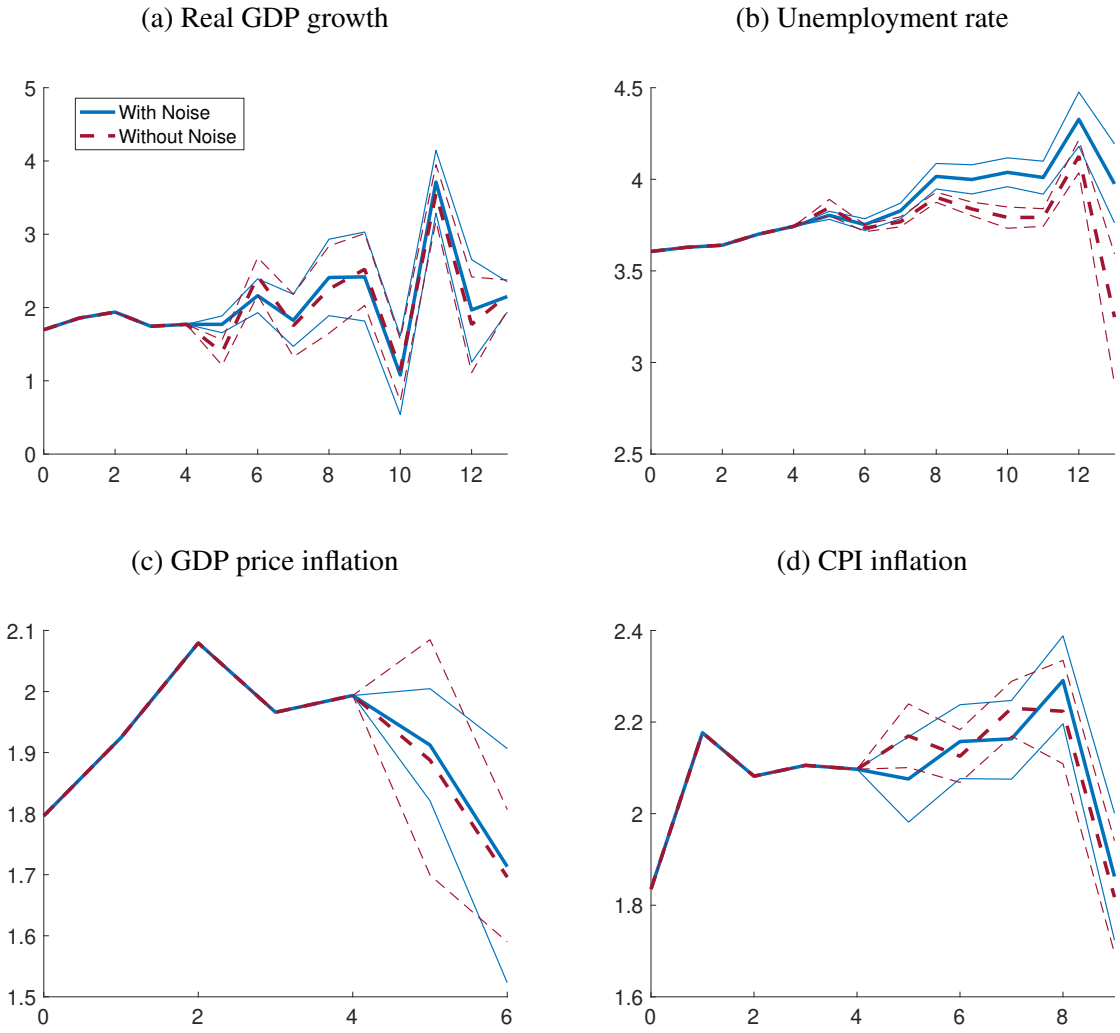


Notes: Posterior medians (and 68% bands) of term structures of SPF-consistent expectations, denoted $\boldsymbol{Y}_t$, obtained from the MDS model with or without noise in the observed SPF calendar-year forecasts. For estimation of the mode without noise, we drop Q4 observations for the next-year forecast (due to perfect overlap with the observed quarterly SPF forecasts.) Estimates based on data through 2019Q4.

(PGDP inflation) at the one-year-ahead horizon ($y = 1$ in the table), the slope coefficient estimate is 0.83 (0.80) in the noise-free VAR forecasts and 0.91 (0.91) in the with-noise VAR forecasts.

While the specification with or without measurement error has some bearing on imputed SPF-consistent expectations, and the model's fit for SPF data, it has less effect on model-based predictive densities for the outcome variable. To illustrate the latter, Figure A.11–A.14 plot the probability integral transforms (PITs) of the forecasts of MDS and VAR models without noise against those from our baseline models with noise. As these figures show, both model variants generate

Figure A.10: Term structures estimated with and without noise (VAR, 2019Q4)



Notes: Posterior medians (and 68% bands) of term structures of SPF-consistent expectations, denoted $\boldsymbol{Y}_t$, obtained from the VAR model with or without noise in the observed SPF calendar-year forecasts. For estimation of the mode without noise, we drop Q4 observations for the next-year forecast (due to perfect overlap with the observed quarterly SPF forecasts.) Estimates based on data through 2019Q4.

fairly similar PITs. Likewise, the (realized) coverage rates for 68% and 90% predictive intervals generated from the model without noise, as reported in Table A.5, are quite similar to those from models with noise as reported in Tables A.11 and A.10 further below (as well as Table 3 in the paper).

While the specification of measurement error notably affects the imputed term structures of expectations (as shown in Figures A.7 through A.10), those effects change the imputed expectations by just about 10-20 basis points. In contrast, the predictive densities for the (quarterly) outcome

variables are much wider; as reported in the paper, the width of their 68% bands regularly amounts to multiple percentage points. These differences in scales also explain the relative similarity in predictive densities obtained from models with and without noise, despite their differing impact on imputed term structures of expectations. All told, the specification of measurement error for annual forecasts matters mainly for improving the model's fit for SPF data than for the outcome variable.

Table A.3: Predictability of SPF point forecasts (noise-free model)

	intercept								**slope**							
	RGDP		UNRATE		PGDP		CPI		RGDP		UNRATE		PGDP		CPI	
Forecast	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR
h = 0	**-1.52**	−1.17	0.81	0.91	−0.04	0.06	−0.33	−0.36	1.41	1.35	0.86	0.85	0.99	0.96	1.15	1.21
	(0.76)	(0.73)	(0.52)	(0.58)	(0.11)	(0.13)	(0.40)	(0.39)	(0.24)	(0.26)	(0.08)	(0.09)	(0.05)	(0.06)	(0.16)	(0.16)
h = 1	−0.11	−0.06	0.55	0.73	−0.08	0.04	−0.07	−0.07	1.02	1.02	0.91	0.88	1.02	0.98	1.01	1.03
	(0.28)	(0.25)	(0.43)	(0.48)	(0.09)	(0.09)	(0.15)	(0.13)	(0.09)	(0.08)	(0.06)	(0.08)	(0.04)	(0.04)	(0.07)	(0.06)
h = 2	−0.14	0.05	0.38	0.58	0.11	0.15	0.10	0.16	1.02	0.97	0.94	0.90	**0.94**	0.99	0.94	0.93
	(0.23)	(0.23)	(0.36)	(0.39)	(0.07)	(0.09)	(0.10)	(0.09)	(0.08)	(0.08)	(0.06)	(0.06)	(0.03)	(0.04)	(0.04)	(0.04)
h = 3	0.13	**1.70**	0.26	0.30	0.15	**1.27**	0.15	**0.61**	0.94	**0.37**	0.96	0.96	**0.92**	**0.37**	**0.92**	**0.75**
	(0.26)	(0.18)	(0.32)	(0.38)	(0.08)	(0.08)	(0.09)	(0.17)	(0.09)	(0.06)	(0.05)	(0.06)	(0.04)	(0.03)	(0.04)	(0.07)
h = 4	**1.67**	**1.69**	0.20	0.44	**1.25**	**0.24**	**0.50**	**0.49**	**0.39**	**0.40**	0.97	0.93	**0.47**	**0.89**	**0.80**	**0.82**
	(0.19)	(0.16)	(0.28)	(0.28)	(0.15)	(0.08)	(0.14)	(0.11)	(0.07)	(0.05)	(0.05)	(0.05)	(0.06)	(0.04)	(0.06)	(0.04)
y = 1	0.23	**0.47**	0.23	0.36	0.14	**0.41**	0.12	**0.38**	0.91	**0.83**	0.96	0.94	0.94	**0.80**	0.94	**0.86**
	(0.15)	(0.14)	(0.34)	(0.38)	(0.09)	(0.10)	(0.09)	(0.15)	(0.05)	(0.05)	(0.06)	(0.06)	(0.04)	(0.05)	(0.04)	(0.06)
y = 2	0.30	**0.68**	0.01	0.32	—	—	**0.60**	**0.78**	0.87	**0.73**	1.00	0.94	—	—	**0.74**	**0.66**
	(0.21)	(0.32)	(0.23)	(0.22)			(0.27)	(0.27)	(0.07)	(0.11)	(0.04)	(0.03)			(0.12)	(0.12)
y = 3	**0.33**	**1.08**	0.22	0.29	—	—	—	—	**0.87**	**0.58**	0.95	0.94	—	—	—	—
	(0.15)	(0.35)	(0.22)	(0.21)					(0.06)	(0.14)	(0.04)	(0.04)				

Notes: Estimated slope coefficients of Mincer-Zarnowitz regressions for model-based predictions of next-quarter's published values for SPF forecasts at different forecast horizons. Heteroskedasticity-consistent standard errors in brackets. Bold font distinguishes coefficient estimates significantly different from 0 (intercept) or 1 (slope) with a 10% confidence level. Evaluation window from 1990Q1 to 2023Q4 (and as far as data for SPF forecasts at the different horizons is available).

Table A.4: Predictability of SPF point forecasts (model w/noise)

	intercept								**slope**							
	RGDP		UNRATE		PGDP		CPI		RGDP		UNRATE		PGDP		CPI	
Forecast	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR
h = 0	**-1.52**	−1.09	0.81	1.02	−0.04	0.04	−0.33	−0.31	1.41	1.32	0.86	0.83	0.99	0.98	1.15	1.19
	(0.76)	(0.71)	(0.52)	(0.60)	(0.11)	(0.12)	(0.40)	(0.37)	(0.24)	(0.25)	(0.08)	(0.10)	(0.05)	(0.06)	(0.16)	(0.15)
h = 1	−0.11	−0.04	0.55	0.82	−0.08	0.08	−0.07	0.04	1.02	1.02	0.91	0.86	1.02	0.99	1.01	1.00
	(0.28)	(0.27)	(0.43)	(0.51)	(0.09)	(0.09)	(0.15)	(0.13)	(0.09)	(0.09)	(0.06)	(0.08)	(0.04)	(0.04)	(0.07)	(0.06)
h = 2	−0.14	0.02	0.38	0.59	0.11	**0.27**	0.10	0.14	1.02	0.97	0.94	0.90	**0.94**	**0.85**	0.94	0.95
	(0.23)	(0.23)	(0.36)	(0.42)	(0.07)	(0.08)	(0.10)	(0.11)	(0.08)	(0.08)	(0.06)	(0.07)	(0.03)	(0.03)	(0.04)	(0.05)
h = 3	0.13	0.50	0.26	0.52	0.15	**0.25**	0.15	**0.21**	0.94	**0.80**	0.96	0.91	**0.92**	**0.86**	**0.92**	**0.92**
	(0.26)	(0.26)	(0.32)	(0.36)	(0.08)	(0.08)	(0.09)	(0.09)	(0.09)	(0.09)	(0.05)	(0.06)	(0.04)	(0.03)	(0.04)	(0.04)
h = 4	0.37	**1.25**	0.20	0.48	**0.20**	**0.24**	0.13	**0.25**	**0.87**	**0.57**	0.97	0.93	**0.91**	**0.90**	0.94	**0.89**
	(0.19)	(0.26)	(0.29)	(0.35)	(0.07)	(0.08)	(0.09)	(0.09)	(0.06)	(0.09)	(0.05)	(0.06)	(0.03)	(0.04)	(0.04)	(0.03)
y = 1	0.09	0.20	0.25	0.41	0.12	**0.18**	0.02	0.10	0.94	0.91	0.96	0.93	**0.93**	**0.91**	0.98	0.96
	(0.25)	(0.21)	(0.28)	(0.33)	(0.07)	(0.07)	(0.12)	(0.11)	(0.09)	(0.07)	(0.05)	(0.06)	(0.03)	(0.03)	(0.05)	(0.05)
y = 2	0.16	0.13	0.36	0.26	—	—	0.36	**0.61**	0.94	0.96	0.92	0.95	—	—	0.85	**0.74**
	(0.25)	(0.27)	(0.36)	(0.24)			(0.25)	(0.21)	(0.09)	(0.10)	(0.07)	(0.04)			(0.11)	(0.09)
y = 3	0.11	0.99	**1.03**	0.18	—	—	—	—	0.95	0.59	**0.76**	0.95	—	—	—	—
	(0.16)	(0.59)	(0.37)	(0.19)					(0.06)	(0.24)	(0.08)	(0.04)				

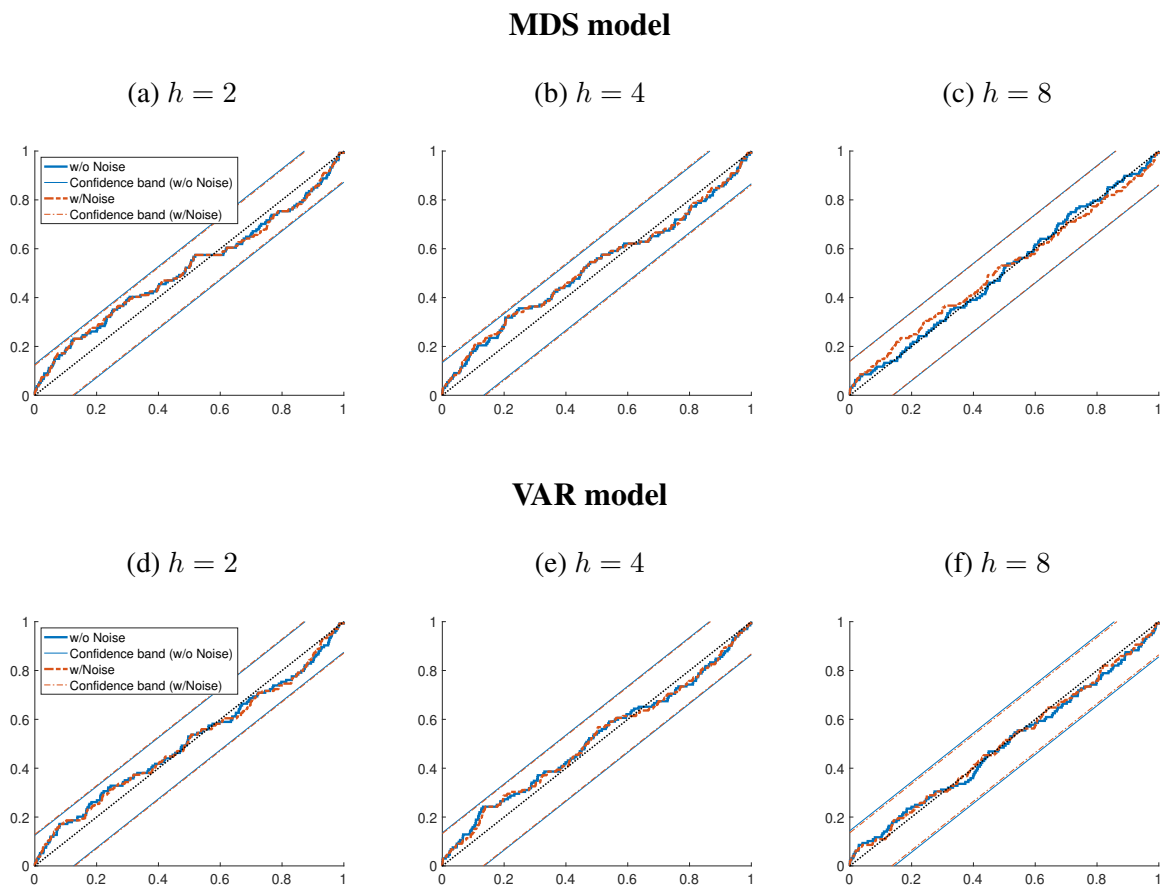
Notes: Estimated slope coefficients of Mincer-Zarnowitz regressions for model-based predictions of next-quarter's published values for SPF forecasts at different forecast horizons. Heteroskedasticity-consistent standard errors in brackets. Bold font distinguishes coefficient estimates significantly different from 0 (intercept) or 1 (slope) with a 10% confidence level. Evaluation window from 1990Q1 to 2023Q4 (and as far as data for SPF forecasts at the different horizons is available).

Table A.5: Coverage rates (model w/o noise, full sample)

	RGDP		UNRATE		PGDP		CPI	
h	68%	90%	68%	90%	68%	90%	68%	90%
PANEL A: MDS Model								
0	48.53***	80.15***	87.50***	96.32***	58.09**	84.56	66.91	91.91
1	54.07***	77.78***	82.96***	96.30***	60.00*	85.19	59.26**	85.19
2	52.99***	79.10***	79.85**	94.78*	60.45	82.09*	58.21**	85.07
3	50.38***	78.20**	75.94	93.98	60.15	85.71	63.16	83.46
4	56.06**	81.82*	71.97	91.67	57.58**	84.85	65.15	84.85
5	54.96**	85.50	69.47	90.08	62.60	87.02	64.89	82.44
6	62.31	85.38	66.92	88.46	69.23	93.08	66.92	84.62
7	66.67	88.37	66.67	89.92	68.22	93.02	65.89	86.05
8	69.53	84.38	66.41	89.06	70.31	93.75	65.62	85.94
9	71.65	84.25	64.57	88.98	70.87	92.91	66.93	87.40
10	70.63	87.30	61.90	88.10	70.63	92.86	67.46	88.10
11	69.60	87.20	64.00	88.00	72.00	92.00	70.40	87.20
12	68.55	87.90	63.71	87.10	75.00	91.13	69.35	87.10
13	69.92	88.62	61.79	87.80	75.61	92.68	69.92	87.80
14	69.67	87.70	60.66	86.89	76.23	94.26	68.03	88.52
15	67.77	88.43	61.16	86.78	76.03	94.21	71.07	87.60
16	71.67	88.33	59.17	87.50	75.83	93.33	71.67	89.17
PANEL B: VAR Model								
0	52.21***	82.35***	80.15***	94.12*	56.62***	84.56	67.65	91.18
1	56.30**	77.78***	82.22***	95.56**	62.96	82.96*	59.26*	85.93
2	56.72**	79.85***	79.10**	94.78*	67.16	85.82	63.43	83.58
3	54.89**	79.70**	78.95*	93.98	64.66	87.22	64.66	86.47
4	55.30**	84.85	72.73	91.67	64.39	86.36	67.42	87.88
5	58.78*	84.73	70.99	90.84	66.41	88.55	68.70	87.79
6	58.46*	84.62	67.69	90.00	66.15	90.00	66.15	86.92
7	58.14*	85.27	63.57	89.92	66.67	90.70	65.12	87.60
8	60.94	82.03	61.72	89.84	66.41	89.84	67.19	87.50
9	62.20	81.10	60.63	88.98	66.14	92.13	70.08	88.19
10	61.90	84.92	57.14	88.10	65.87	91.27	69.84	88.10
11	64.00	82.40	55.20**	88.00	68.00	89.60	68.00	89.60
12	66.13	84.68	54.84*	86.29	67.74	91.13	70.97	89.52
13	68.29	85.37	54.47*	84.55	66.67	91.06	73.17	89.43
14	66.39	84.43	53.28*	85.25	70.49	90.16	71.31	90.16
15	66.94	86.78	52.89*	85.12	73.55	90.91	73.55	89.26
16	66.67	87.50	53.33	85.83	74.17	91.67	72.50	90.00

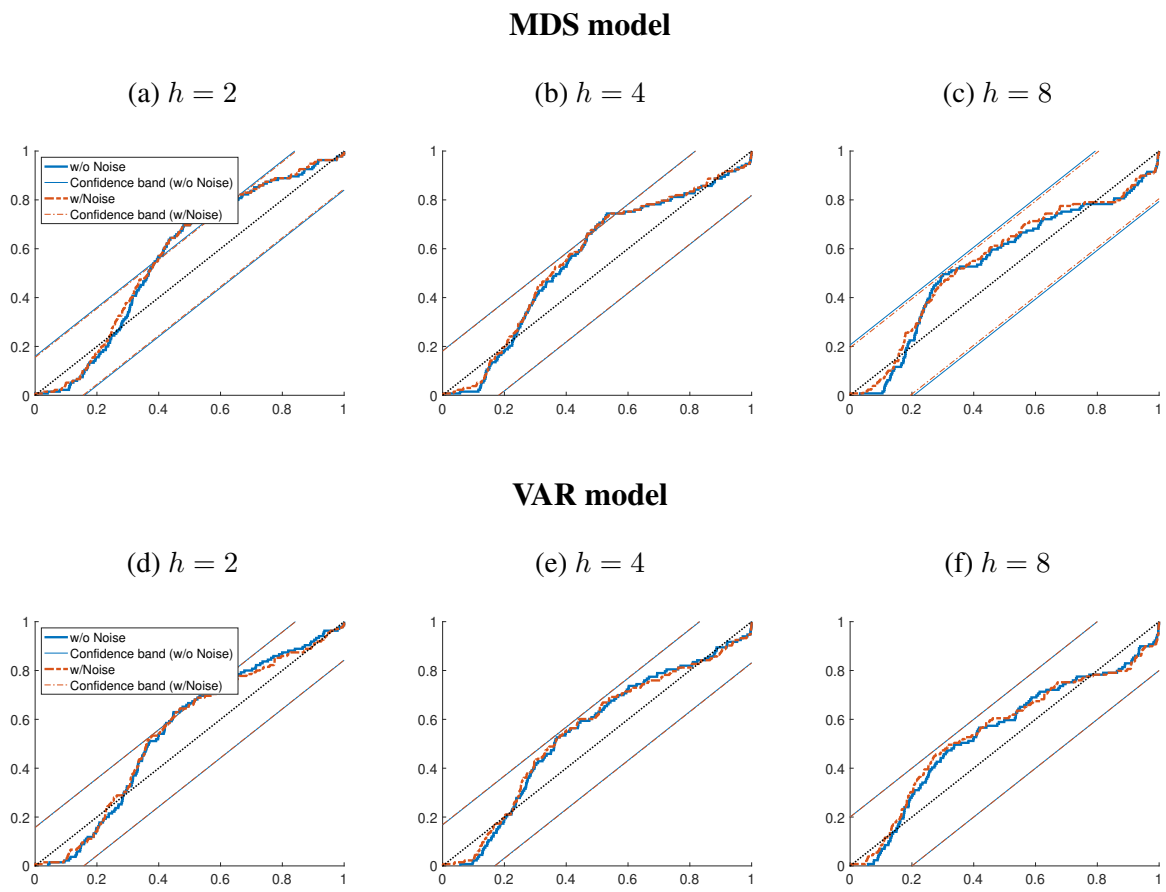
Note: Coverage rates for uncertainty bands with nominal levels of 68% and 90% for out-of-sample forecasts at quarterly forecast horizons, h. Evaluation window from 1990Q1 through 2023Q4 (and as far as realized values are available). Reflecting the availability of annual SPF forecasts, forecasts for inflation in CPI and GDP prices are evaluated only up to $h = 12$, and $h = 8$, respectively. 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 A.11: GDP growth PITs with and without noise



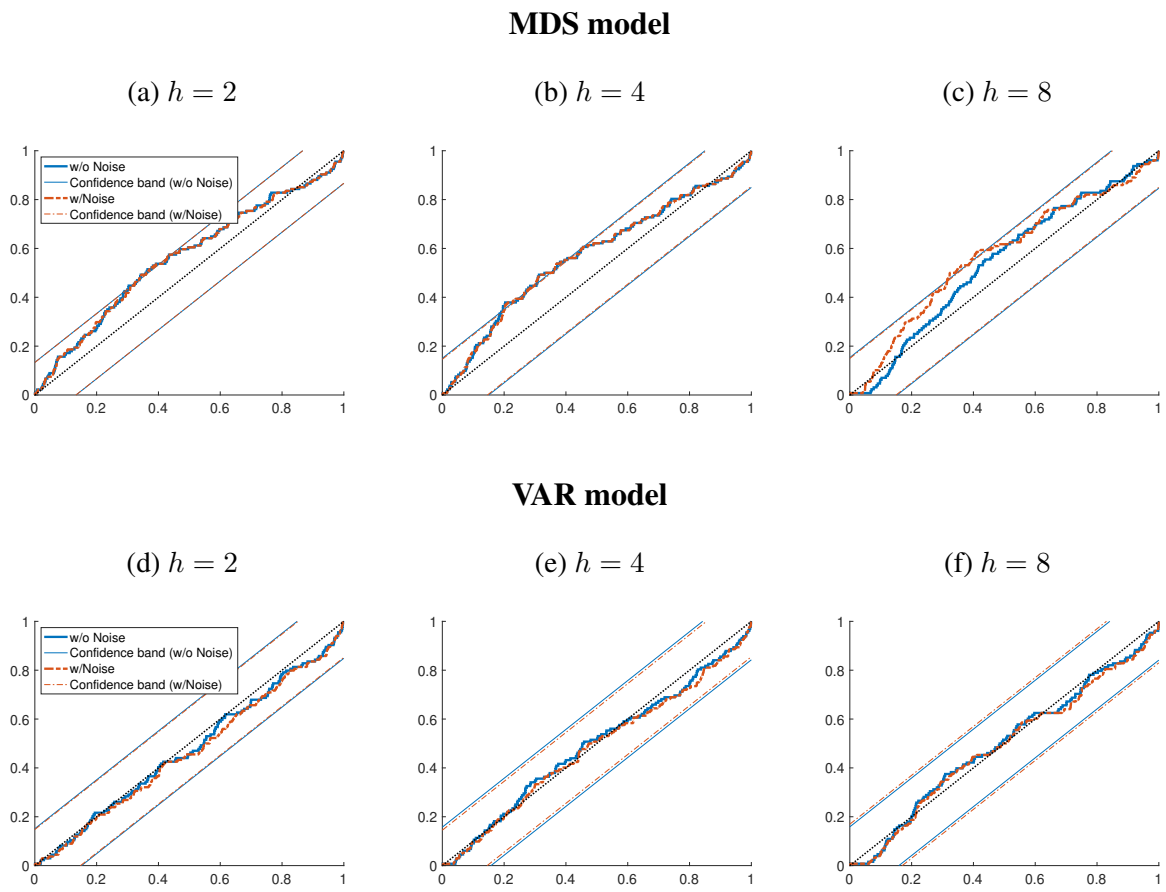
Notes: Empirical cumulative distributions of probability integral transforms (PITs) for GDP growth at selected quarterly forecast horizons. All forecasts are generated out of sample by our MDS and VAR models (with and without noise in measurement equations for annual forecasts), and evaluated over an evaluation window from 1990Q1 through 2023Q4 (and as far as realized values are available). 95% confidence bands for tests of correct calibration from Rossi and Sekhposyan (2019); computed separately for each model, but with nearly identical plot lines.

Figure A.12: Unemployment rate PITs with and without noise



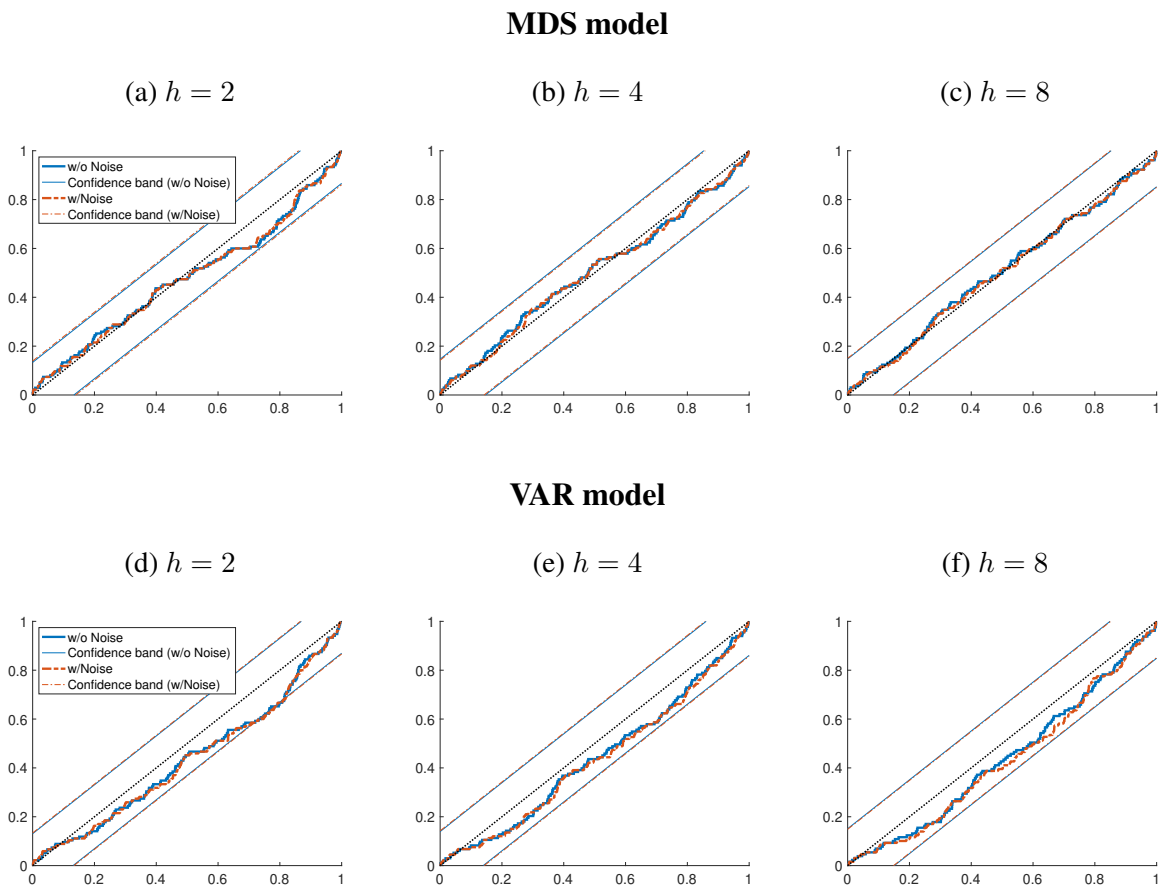
Notes: Empirical cumulative distributions of probability integral transforms (PITs) for Unemployment rate at selected quarterly forecast horizons. All forecasts are generated out of sample by our MDS and VAR models (with and without noise in measurement equations for annual forecasts), and evaluated over an evaluation window from 1990Q1 through 2023Q4 (and as far as realized values are available). 95% confidence bands for tests of correct calibration from Rossi and Sekhposyan (2019); computed separately for each model, but with nearly identical plot lines.

Figure A.13: GDP price inflation PITs with and without noise



Notes: Empirical cumulative distributions of probability integral transforms (PITs) for GDP price inflation at selected quarterly forecast horizons. All forecasts are generated out of sample by our MDS and VAR models (with and without noise in measurement equations for annual forecasts), and evaluated over an evaluation window from 1990Q1 through 2023Q4 (and as far as realized values are available). 95% confidence bands for tests of correct calibration from Rossi and Sekhposyan (2019); computed separately for each model, but with nearly identical plot lines.

Figure A.14: CPI inflation PITs with and without noise



Notes: Empirical cumulative distributions of probability integral transforms (PITs) for CPI inflation at selected quarterly forecast horizons. All forecasts are generated out of sample by our MDS and VAR models (with and without noise in measurement equations for annual forecasts), and evaluated over an evaluation window from 1990Q1 through 2023Q4 (and as far as realized values are available). 95% confidence bands for tests of correct calibration from Rossi and Sekhposyan (2019); computed separately for each model, but with nearly identical plot lines.

II Details on Bayesian MCMC sampler and priors

II(a) Model summary, priors and MCMC steps

Before turning to a description of priors and MCMC sampling steps, we begin by restating the equations of our general model, with a VAR specification for forecast updates and unconditional bias, as detailed in Section 4 of the paper.

States: The model tracks a term structure of SPF-implied quarterly forecasts, denoted $\boldsymbol{Y_t}$ as detailed in equation (1) of the paper:

$$\boldsymbol{Y_t} \equiv \Big[y_{t-1},\ F_t y_t,\ F_t y_{t+1}, \ldots, F_t y_{t+h}, \ldots, F_t y_{t+H} \Big]' . \tag{1}$$

The dynamics of $\boldsymbol{Y_t}$ are characterized by the following trend-cycle decomposition:

$$\boldsymbol{Y_t} = \tilde{\boldsymbol{Y}}_{\boldsymbol{t}} + \mathbf{1} y_t^* \,, \tag{A.46}$$

$$y_t^* = y_{t-1}^* + w_t^* \,, \qquad w_t^* \sim \mathcal{N}(0, \omega_t^2) \,. \tag{A.47}$$

Section 4 of the paper, in equations (7), (9), and (10), derives the following law of motion for the gap vector:

$$\tilde{\boldsymbol{Y}}_{\boldsymbol{t}} = \left(\boldsymbol{I} - \tilde{\boldsymbol{\Psi}} \right) \bar{\boldsymbol{Y}} + \tilde{\boldsymbol{\Psi}} \tilde{\boldsymbol{Y}}_{\boldsymbol{t-1}} + \tilde{\boldsymbol{\eta}}_{\boldsymbol{t}} \,, \tag{7}$$

$$\text{and} \quad \tilde{\boldsymbol{\eta}}_{\boldsymbol{t}} = \tilde{\boldsymbol{\Pi}} \, \tilde{\boldsymbol{\eta}}_{\boldsymbol{t-1}} + \tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{t}} \,, \ \text{with} \ \ \tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{t}} \sim \mathcal{N}(\mathbf{0}, \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{t}}), \tag{9}$$

$$\Rightarrow \quad \tilde{\boldsymbol{Y}}_{\boldsymbol{t}} = \left(\boldsymbol{I} - \tilde{\boldsymbol{\Psi}} \right) \left(\boldsymbol{I} - \tilde{\boldsymbol{\Pi}} \right) \bar{\boldsymbol{Y}} + \left(\tilde{\boldsymbol{\Psi}} + \tilde{\boldsymbol{\Pi}} \right) \tilde{\boldsymbol{Y}}_{\boldsymbol{t-1}} - \left(\tilde{\boldsymbol{\Psi}} \, \tilde{\boldsymbol{\Pi}} \right) \tilde{\boldsymbol{Y}}_{\boldsymbol{t-2}} + \tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{t}} \,, \tag{10}$$

where $\tilde{\boldsymbol{\Psi}}$ is a matrix of zeros and ones as described in equation (8) of the paper, $\tilde{\boldsymbol{\Pi}}$ a stable matrix to be estimated, and $\bar{\boldsymbol{Y}}$ a vector of average gap values, also to be estimated.

As discussed in the paper, the MDS version of our model is nested in the above, with the restrictions $\bar{\boldsymbol{Y}} = \mathbf{0}$ and $\tilde{\boldsymbol{\Pi}} = \mathbf{0}$.

The measurement equations are:

$$\boldsymbol{Z_t} = \begin{bmatrix} \boldsymbol{Z_{q,t}} \\ \boldsymbol{Z_{a,t}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{C_{q,t}} \\ \boldsymbol{C_{a,t}} \end{bmatrix} \boldsymbol{Y_t} + \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{n_t} \end{bmatrix}, \quad n_{i,t} \sim \mathcal{N}(0, \sigma^2_{i,t}), \tag{A.48}$$

where $\boldsymbol{Z_{q,t}}$ contains the lagged realized value y_{t-1} and observed quarterly fixed-horizon forecasts from the SPF (all assumed to be measured without error) and $\boldsymbol{Z_{a,t}}$ consists of the observed fixed-event annual predictions from the SPF. Further details on $\boldsymbol{Z_{q,t}}$, $\boldsymbol{Z_{a,t}}$ and their measurement loadings $\boldsymbol{C_{q,t}}$ and $\boldsymbol{C_{a,t}}$ are described in Appendix I(b) above.

Shock distributions: Horseshoe models are applied to the time-varying variances of shocks to trend and noise, ω_t^2 and $\sigma_{i,t}^2$, as detailed further below in appendix II(b). As described in Section 4.5 of the paper, the time-varying second moment matrices of the gap shocks are modeled via the following two-block stochastic volatility (SV) process with fat tails. Restating the equations from the paper, we have:

$$\tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{t}} = \begin{bmatrix} \tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{1,t}} \\ \tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{2,t}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} & \tilde{\boldsymbol{K}} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\varepsilon}^*_{\boldsymbol{1,t}} \\ \boldsymbol{\varepsilon}^*_{\boldsymbol{2,t}} \end{bmatrix}, \quad \text{with} \begin{bmatrix} \boldsymbol{\varepsilon}^*_{\boldsymbol{1,t}} \\ \boldsymbol{\varepsilon}^*_{\boldsymbol{2,t}} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}, \begin{bmatrix} \lambda_{1,t} \cdot \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{11}} & \boldsymbol{0} \\ \boldsymbol{0} & \lambda_{2,t} \cdot \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{22}} \end{bmatrix} \right), \tag{15}$$

in which $\tilde{\boldsymbol{K}}$ is a matrix (with dimension $4 \times (H-2)$) of coefficients to be estimated. This SV structure yields the following time-varying variance-covariance matrix of the cyclical shocks:

$$\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{t}} = \begin{bmatrix} \boldsymbol{I} & \tilde{\boldsymbol{K}} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \lambda_{1,t} \cdot \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{11}} & \boldsymbol{0} \\ \boldsymbol{0} & \lambda_{2,t} \cdot \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{22}} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \tilde{\boldsymbol{K}} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix}'. \tag{16}$$

The scalar factors $\lambda_{1,t}$ and $\lambda_{2,t}$ impart time variation and fat tails to the shock vector $\tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{t}}$. Building on, among others, Carriero et al. (2022b), Chan (2020), and Jacquier et al. (2004), we model these

factors as the products of iid inverse-gamma draws and persistent stochastic volatility processes:

$$\lambda_{i,t} = \phi_{i,t} \cdot \tilde{\lambda}_{i,t}\,, \qquad \forall\, i = 1, 2\,, \tag{17}$$

$$\text{with} \quad \phi_{i,t} \sim \mathcal{IG}\left(\frac{\nu_i}{2}, \frac{\nu_i}{2}\right), \qquad \log \tilde{\boldsymbol{\lambda}}_{\boldsymbol{t}} \equiv \begin{bmatrix} \log \tilde{\lambda}_{1,t} \\ \log \tilde{\lambda}_{2,t} \end{bmatrix} = \begin{bmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{bmatrix} \log \tilde{\boldsymbol{\lambda}}_{\boldsymbol{t-1}} + \boldsymbol{\epsilon}^{\boldsymbol{\lambda}}_{\boldsymbol{t}}\,, \tag{18}$$

and $\boldsymbol{\epsilon}_{\boldsymbol{t}}^{\boldsymbol{\lambda}} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Phi})$. The iid inverse-gamma draws add fat tails in the form of a multivariate t distribution with ν_i degrees of freedom to each block. The vector SV process $\log \tilde{\boldsymbol{\lambda}}_{\boldsymbol{t}}$ has correlated shocks and is normalized to a mean of zero, obviating the need for normalizing assumptions on the constant-coefficient matrices $\tilde{\boldsymbol{\Sigma}}_{\mathbf{11}}$ and $\tilde{\boldsymbol{\Sigma}}_{\mathbf{22}}$.

Notation for state vectors, measurement vectors, and parameters: We collect all observed measurements, $\{\boldsymbol{Z_t}\}_{t=1}^{T}$, in the measurement vector $\boldsymbol{Z}$, and all values of $\{y_t^*\}_{t=0}^{T}$, $\{\tilde{\boldsymbol{Y}}_{\boldsymbol{t}}\}_{t=-1}^{T}$, and $\bar{\boldsymbol{Y}}$ in the (linear) state vector $\boldsymbol{Y}$, where $t = T$ denotes the end of the sample, and $t = -1, 0$ points to initial conditions.[6] Conditional on values for the time-varying second moment parameters, $\{\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{t}}\}_{t=1}^{T}$, and $\{\omega_t^2\}_{t=1}^{T}$, the equations above describe a Gaussian state space, with measurements $\boldsymbol{Z}$ and state vector $\boldsymbol{Y}$. In addition, we collect the latent SV and t-mixture-representation states, $\tilde{\lambda}_{i,t}$ and $\phi_{t,i}$, in the vectors $\tilde{\boldsymbol{\lambda}}$, and $\boldsymbol{\phi}$, respectively, and collect the time-varying second moment parameters for trend and noise shocks in the vectors $\boldsymbol{\omega^2}$ and $\boldsymbol{\sigma^2}$. The remaining parameters of the model are the VAR coefficient matrix, $\tilde{\boldsymbol{\Pi}}$, fixed parameters of the block-SV model, $\boldsymbol{K}$, $\tilde{\boldsymbol{\Sigma}}_{\mathbf{11}}$, and $\tilde{\boldsymbol{\Sigma}}_{\mathbf{22}}$, the AR(1) coefficients of the cyclical SV processes, $\boldsymbol{\rho} = [\rho_1,\ \rho_2]'$, and associated variance-covariance matrix of shocks to the SV processes, $\boldsymbol{\Phi}$, as well as the degrees of freedom of the multivariate t distributions for cyclical shocks, $\boldsymbol{\nu} = [\nu_1,\ \nu_2]'$.

In our description of the MCMC sampler below, we denote the sets of constant parameters and time-varying second moment states as follows:

$$\boldsymbol{\Theta} \equiv \left\{\tilde{\boldsymbol{\Pi}}, \boldsymbol{K}, \tilde{\boldsymbol{\Sigma}}_{\mathbf{11}}, \tilde{\boldsymbol{\Sigma}}_{\mathbf{22}}, \boldsymbol{\rho}, \boldsymbol{\Phi}, \boldsymbol{\nu}\right\} , \tag{A.49}$$

$$\boldsymbol{\Omega} = \left\{\tilde{\boldsymbol{\lambda}}, \boldsymbol{\phi}, \boldsymbol{\omega}^2, \boldsymbol{\sigma}^2\right\} . \tag{A.50}$$

When referring to any of these sets while excluding one of its components, we simply refer to $\boldsymbol{\Theta}^{\dagger}$ or $\boldsymbol{\Omega}^{\dagger}$, where the excluded component shall be clear from the context. Moreover, we suppress notation for mixture states involved in the estimation of the SV processes based on the methods of Kim et al. (1998) and Omori et al. (2007), as well as mixture states used in the horsehoe representations of the shocks to trend and noise (with further details provided in Appendix II(b)).

We use the following priors for parameters and initial values of the states:

[6] Our reference to $\boldsymbol{Y}$ as state vector is understood conditional on trajectories for the time-varying second moment parameters, whose stochastic evolution is, of course, also needed to fully describe the state of the model.

- $y_0^* \sim N\left(0, 100^2\right)$, which is an essentially diffuse prior for the initial value of the trend level.
- $\begin{bmatrix} \tilde{\boldsymbol{Y}}_{\boldsymbol{0}} \\ \tilde{\boldsymbol{Y}}_{\boldsymbol{-1}} \end{bmatrix} \sim N\left(\boldsymbol{0}, 25 \cdot \boldsymbol{I}\right)$, which is a fairly wide prior for the initial gap levels.
- $\bar{\boldsymbol{Y}} \sim N\left(\boldsymbol{0}, \boldsymbol{D_Y}\right)$ with $\boldsymbol{D_Y}$ a diagonal matrix with typical element $\mathrm{Var}\left(\bar{Y}_j\right) = 25/j$ so as to imply shrinkage of (unconditional) bias towards zero that is increasing with forecast horizon.
- $\mathrm{vec}\left(\tilde{\boldsymbol{\Pi}}\right) \sim N\left(\boldsymbol{0}, \boldsymbol{D}_{\tilde{\boldsymbol{\Pi}}}\right)$, where $\boldsymbol{D}_{\tilde{\boldsymbol{\Pi}}}$ is a diagonal matrix that implements the structure of a typical Minnesota prior on the VAR coefficients, with overall shrinkage $\mathrm{Var}\left(\tilde{\Pi}_{i,i}\right) = \theta_1$ and cross-variable shrinkage $\mathrm{Var}\left(\tilde{\Pi}_{i,j}\right) = \theta_1 \cdot \theta_2$ for $i \neq j$, and $\theta_1 = .2^2$ and $\theta_2 = .5^2$.
- $\log\left(\tilde{\lambda}_{j,0}\right) \sim \mathcal{N}(0, 100)$ for $j = 1, 2$, which is a fairly uninformative prior for the initial values of the SV factors.
- $\rho_j \sim \mathcal{N}(0.8, 0.2^2)$ for $j = 1, 2$ as in Clark and Ravazzolo (2015) and other studies.
- $\nu_j \sim U(3, 40)$ for $j = 1, 2$, and implemented over a grid of natural numbers as in Jacquier et al. (2004).
- $\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{jj}} \sim IW(N_j, 0.01 \cdot \boldsymbol{I})$ for $i = 1, 2$, where N_j is the number of elements of the j block of the SV model, so that the prior is relatively uninformative, and has no mean.
- $\mathrm{vec}\left(\boldsymbol{K}\right) \sim N\left(\boldsymbol{0}, \boldsymbol{I}\right)$.
- $\boldsymbol{\Phi} \sim IW(2, \boldsymbol{I})$, which is fairly uninformative since with as many degrees of freedom as there are SV shocks, this prior has no mean.

Note that the initial levels of trend and mean bias, y_0^* and $\bar{\boldsymbol{Y}}$, are not separately identifiable and could be normalized, for example by setting $y_0^* = 0$. We choose to estimate both, but report only statistics reflecting their joint effects, as the estimates with the normalization $y_0^* = 0$ displayed poorer convergence properties in our experiments.

Our MCMC sampler iterates over the following steps:

1. $p\left(\boldsymbol{Y} \middle| \boldsymbol{\Omega}, \boldsymbol{\Theta}, \boldsymbol{Z}\right)$. Draws from the latent vector of term-structure expectations can be obtained via standard sampling techniques for a linear Gaussian state space model. For computational

efficiency, we build on the precision-based sampler developed by Mertens (2023), with details described below in Appendix II(c).

2. $p\left(\tilde{\boldsymbol{\Pi}}\middle|\boldsymbol{Y},\boldsymbol{\Omega},\boldsymbol{\Theta}^{\dagger},\boldsymbol{Z}\right)$. We draw from the posterior of the VAR coefficients using a Bayesian updating with normal conjugate prior (and posteriors). Appendix II(d) provides details of an efficient implementation that accounts for the heteroskedasticity in the VAR residuals while exploiting the two-block structure of the SV model. Rejection sampling is employed to ensure that the VAR coefficients remain within the unit circle.

3. $p\left(\boldsymbol{K}\middle|\boldsymbol{Y},\boldsymbol{\Omega},\boldsymbol{\Theta}^{\dagger},\boldsymbol{Z}\right)$ is a standard Bayesian vector regression with normal conjugate prior (and posterior), performed by regressing draws of $\tilde{\boldsymbol{\varepsilon}}_{\mathbf{1},\boldsymbol{t}}$ on $\tilde{\boldsymbol{\varepsilon}}_{\mathbf{2},\boldsymbol{t}}$ after scaling each by $\lambda_{1,t}$.

4. For $j=1,2$, draw $p\left(\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{jj}}\middle|\boldsymbol{Y},\boldsymbol{\Omega},\boldsymbol{\Theta}^{\dagger},\boldsymbol{Z}\right)$, which are standard inverse-Wishart updates based on draws of $\boldsymbol{\varepsilon}^{*}_{\boldsymbol{j},\boldsymbol{t}}/\lambda_{j,t}$.

5. $p\left(\boldsymbol{\nu}\middle|\boldsymbol{Y},\boldsymbol{\Omega},\boldsymbol{\Theta}^{\dagger},\boldsymbol{Z}\right)$ the degrees of freedom for the multivariate student t distributions of the cyclical shocks are sampled over a uniform grid of natural numbers, as described by Jacquier et al. (2004).

6. $p\left(\tilde{\boldsymbol{\phi}}\middle|\boldsymbol{Y},\boldsymbol{\Omega}^{\dagger},\boldsymbol{\Theta},\boldsymbol{Z}\right)$ which are the common inverse-gamma mixture states of the multivariate t distributions for the cyclical shock blocks, and are drawn as described, for example, by Chan (2020), exploiting the conjugacy of their (conditional) inverse Gamma priors.

7. $p\left(\tilde{\boldsymbol{\lambda}}\middle|\boldsymbol{Y},\boldsymbol{\Omega}^{\dagger},\boldsymbol{\Theta},\boldsymbol{Z}\right)$ which are the common SV processes of the cyclical shock blocks, and are drawn as described, for example, by Carriero et al. (2016) and Chan (2020), using the mixture-state SV sampler of Kim et al. (1998), with a 10-point grid as recommended in Omori et al. (2007), while following the advice of Del Negro and Primiceri (2015), regarding the correct ordering of steps in the mixture sampling.

8. $p\left(\boldsymbol{\rho}\middle|\boldsymbol{Y},\boldsymbol{\Omega},\boldsymbol{\Theta}^{\dagger},\boldsymbol{Z}\right)$ is a seemingly-unrelated Bayesian system regression with normal conjugate prior (and posterior), using draws of $\tilde{\lambda}_{j,t}$ for $j=1,2$, and which is conducted using rejection sampling to ensure that the AR(1) coefficients remain within the unit circle.

9. $p\left(\boldsymbol{\Phi}\middle|\boldsymbol{Y},\boldsymbol{\Omega},\boldsymbol{\Theta}^{\dagger},\boldsymbol{Z}\right)$ is a standard inverse-Wishart update based on draws of the shocks to the (log-)SV processes.

10. $p\left(\boldsymbol{\omega}^2 \middle| \boldsymbol{Y}, \boldsymbol{\Omega}^\dagger, \boldsymbol{\Theta}, \boldsymbol{Z}\right)$ and $p\left(\boldsymbol{\sigma}^2 \middle| \boldsymbol{Y}, \boldsymbol{\Omega}^\dagger, \boldsymbol{\Theta}, \boldsymbol{Z}\right)$ are independently sampled using the Gibbs sampling steps described by Makalic and Schmidt (2016) for the horseshoe model. (See also Appendix II(b)).

The MDS version of our model restricts $\bar{\boldsymbol{Y}}$ and $\tilde{\boldsymbol{\Pi}}$ to be zero, so that when estimating the model, we omit $\bar{\boldsymbol{Y}}$ from the state vector $\boldsymbol{Y}$ and drop the sampling step for $\tilde{\boldsymbol{\Pi}}$. In addition, we can omit $\tilde{\boldsymbol{Y}}_{-1}$ from the set of initial conditions, since the gap dynamics are fully determined by a VAR(1) (instead of a VAR(2)) in this case.

The remainder of this appendix describes details of the horseshoe specifications for the shocks to trend and noise, the precision-based sampler for the state space, and the sampling of VAR coefficients when heteroskedasticity in its residuals is described by a two-block SV model.

II(b) Horseshoe shock specifications

II(b.1) Horseshoe model for trend shocks

We model shocks to the trend and measurement errors via a horseshoe model. The horseshoe model has originally been proposed by Carvalho et al. (2010) for modeling sparse regressions, i.e. regressions with a priori potentially many regressors, many of whom are however expected to be irrelevant, with only a few attracting substantial mass a posteriori. As such, while the horseshoe prior places considerable mass on coefficient values of zero, it also has particularly fat tails to generate (few) significantly-sized coefficient estimates.

In our application, we apply the horseshoe model to sequences of shocks (in this case: the trend shocks), instead of regression coefficients where we suspect that most realizations are close to zero while some can also be sizable. In a similar spirit, Prüser (2021) applies a horseshoe model to the shocks of drifting coefficients in a VAR with time-varying parameters. The horseshoe has a conditionally Gaussian representation for the shocks:

$$w_t^* \sim \mathcal{N}\left(0, \omega_t^2\right) \tag{A.51}$$

and achieves its particular form with a hierarchical model for the conditional variance ω_t^2:

$$\Rightarrow \quad \omega_t^2 = \tau_\omega^2 \cdot \vartheta_{w,t}^2 \tag{A.52}$$

$$\text{with} \quad \tau_{\omega}^2 \sim \mathcal{C}^+(0,1)\,, \qquad\qquad \text{and} \quad \vartheta_{w,t}^2 \sim \mathcal{C}^+(0,1)\,, \tag{A.53}$$

where $\mathcal{C}^+(0,1)$ denotes the half-Cauchy distribution. In this horseshoe model, τ_w^2 denotes the global shrinkage (applicable to shocks, w_t^*, at all t) and $\vartheta_{w,t}^2$ denotes local shrinkage (that is specific to the time-t realization of w_t^*). For brevity, we denote the hierarchical model for ω_t^2 as follows:

$$\omega_t^2 \sim \mathcal{HS}\left(\tau_{\omega}^2\right). \tag{A.54}$$

The horseshoe model can be represented via scale mixtures. Moreover, as shown by Makalic and Schmidt (2016), Bayesian estimation via Gibbs sampling becomes straightforward when auxiliary variables are used, and we follow their approach in sampling posterior values for ω_t^2. Estimated trend levels are reported in Appendix V further below.

II(b.2) Measurement error with horseshoe model

With similar motivation to the trend shocks, we apply (separate) horseshoe models to the measurement errors attached to annual SPF forecasts. Moreover, since we suspect that the size of measurement errors varies across quarters of the year, we apply separate horseshoe models to forecasts collected in different quarters of the year. For this purpose, let $q(t) \in \{1,2,3,4\}$ denote a function that maps a time index t into the corresponding quarter of the year.

As described, the measurement error in the annual forecast i at time t by $n_{i,t}$ has a conditionally Gaussian distribution, and for each $q(t)$ we apply separate horseshoe models to the conditional variance of $n_{i,t}$:

$$n_{i,t} \sim \mathcal{N}(0, \sigma_{i,t}^2)\,, \qquad \sigma_{i,t}^2 \sim \mathcal{HS}\left(\tau_{i,q(t)}^2\right) \tag{A.55}$$

$$\Rightarrow \quad \sigma_{i,t}^2 = \tau_{i,q(t)}^2 \cdot \vartheta_{i,t}^2\,, \qquad \tau_{i,q(t)}^2 \sim \mathcal{C}^+(0,1) \qquad \vartheta_{i,t}^2 \sim \mathcal{C}^+(0,1)\,. \tag{A.56}$$

Estimated noise levels are reported in Appendix I(c).

II(c) Precision-based sampling from state space

Step 1 of the MCMC scheme outlined above involves drawing a vector of latent states, $\boldsymbol{Y}$, conditional on observables, $\boldsymbol{Z}$, and draws of model parameters (including time-varying second moment

parameters). This sampling step involves a Gaussian signal extraction, which is efficiently implemented with a precision-based sampler, that extends methods detailed in Mertens (2023), and that is described here.

For ease of notation, we drop dependence of the sampling problem on the various model parameters. As before, $\boldsymbol{Y}$, denotes a vector of all values for $\boldsymbol{Y_t}$ for all t, stacked on top of each other, and likewise for $\boldsymbol{Z}$. In stacked form, the state space can be written as follows:

$$\boldsymbol{\mathcal{A}}\boldsymbol{Y} = \boldsymbol{Y_0} + \boldsymbol{\mathcal{B}}\,\boldsymbol{w_t}\,, \qquad \boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})\,, \tag{A.57}$$

$$\boldsymbol{Z_q} = \boldsymbol{\mathcal{C}_q}\boldsymbol{Y}\,, \tag{A.58}$$

$$\boldsymbol{Z_a} = \boldsymbol{\mathcal{C}_a}\boldsymbol{Y} + \boldsymbol{\mathcal{D}}\boldsymbol{n}\,, \qquad \boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})\,, \tag{A.59}$$

and $\boldsymbol{Z_q}$ and $\boldsymbol{Z_a}$ collecting measurement equations for, respectively, quarterly fixed-horizon data (without measurement error) and annual fixed-event data (with measurement error). Details of the mapping between a dynamic representation, such as the one described in Section 4 of the paper, and this stacked representation are illustrated, for example, in Mertens (2023).

Originally developed in Chib and Jeliazkov (2006) and then Chan and Jeliazkov (2009), precision-based samplers offer a computationally efficient alternative to sample latent states from linear Gaussian models as compared to recursive methods based on Kalman filtering and smoothing, such as the simulation smoother of Durbin and Koopman (2002). Precision-based samplers operate on the inverse variance of the state vector. However, in cases where the measurement vector is assumed to be observed without measurement error, as in (A.58), we face an ill-defined posterior precision:

$$\boldsymbol{\mathcal{C}_q}\,\mathrm{Var}\left(\boldsymbol{Y}|\boldsymbol{Z_q}\right) = \boldsymbol{0} \quad \Longrightarrow \quad \left|\,\mathrm{Var}\left(\boldsymbol{Y}|\boldsymbol{Z_q}\right)\right| = 0\,, \tag{A.60}$$

which cannot be directly handled by conventional precision-based samplers. Mertens (2023) derives a precision-based sampler for the case when all measurement variables are observed without error, thus corresponding to the state space consisting solely of (A.57) and (A.58). The version of our model without error in all measurement equations is isomorphic to merely sampling from $p\left(\boldsymbol{Y}|\boldsymbol{Z_q}\right)$, but using $\boldsymbol{Z}$ in lieu of $\boldsymbol{Z_q}$, and, in this case, the sampling methods of Mertens (2023) can be directly applied.

To sample from the space consisting of (A.57) and (A.58) *and* (A.59), where some measurements, but not all, are observed without error, we build on the methods of Mertens (2023) as follows:

- Consider the posterior moments of $p\left(\boldsymbol{Y}|\boldsymbol{Z_q}\right)$:
  - From a QR decomposition of $\boldsymbol{\mathcal{C}_q}$, obtain the following:

$$\boldsymbol{\mathcal{C}_q} = \boldsymbol{RQ} = \begin{bmatrix} \boldsymbol{R_1} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{Q_1} \\ \boldsymbol{Q_2} \end{bmatrix}, \qquad \boldsymbol{QQ'} = \boldsymbol{I}, \tag{A.61}$$

$$\boldsymbol{\mathcal{Y}} = \begin{bmatrix} \boldsymbol{\mathcal{Y}_1} \\ \boldsymbol{\mathcal{Y}_2} \end{bmatrix} \equiv \begin{bmatrix} \boldsymbol{Q_1} \\ \boldsymbol{Q_2} \end{bmatrix} \boldsymbol{Y}, \qquad \Leftrightarrow \quad \boldsymbol{Y} = \boldsymbol{Q_1'}\,\boldsymbol{\mathcal{Y}_1} + \boldsymbol{Q_2'}\,\boldsymbol{\mathcal{Y}_2}, \tag{A.62}$$

    where $\boldsymbol{R_1}$ is lower triangular, $\boldsymbol{\mathcal{Y}_1}$ describes linear combinations of the state vector $\boldsymbol{Y}$ that are perfectly described by $\boldsymbol{Z_q}$, and $\boldsymbol{\mathcal{Y}_2}$ collects the remaining linear combinations. Since $\boldsymbol{Y}$ is multivariate normal, so is also $\boldsymbol{\mathcal{Y}}$.
  - With $\boldsymbol{\mathcal{Y}_1} = \boldsymbol{R_1^{-1}}\boldsymbol{Z_q}$, the problem of describing $p\left(\boldsymbol{Y}|\boldsymbol{Z_q}\right)$ boils down to

$$p\left(\boldsymbol{\mathcal{Y}_2}|\boldsymbol{Z_q}\right) \sim \mathcal{N}\left(\boldsymbol{\mu_{2|1}}, \boldsymbol{P_{22|1}^{-1}}\right). \tag{A.63}$$

    Derivations for the posterior mean, $\boldsymbol{\mu_{2|1}}$, and precision, $\boldsymbol{P_{22|1}}$, are detailed in Mertens (2023).

- For $p\left(\boldsymbol{\mathcal{Y}_2}|\boldsymbol{Z_q}, \boldsymbol{Z_a}\right)$, set up a conventional precision-based sampling problem with measurement error.
  - The problem is isomorphic to sampling from $p\left(\boldsymbol{\mathcal{Y}_2}|\boldsymbol{Z_a}\right)$, while using $\boldsymbol{\mu_{2|1}}$, and $\boldsymbol{P_{22|1}}$ as prior moments.
  - The transformed measurement equation is:

$$\boldsymbol{Z_a} = \boldsymbol{\mathcal{C}_a}\boldsymbol{Q_1'}\,\boldsymbol{R_1^{-1}}\boldsymbol{Z_q} + \boldsymbol{\mathcal{C}_a}\boldsymbol{Q_2'}\,\boldsymbol{\mathcal{Y}_2} + \boldsymbol{\mathcal{D}n} \tag{A.64}$$

$$\Rightarrow \quad \boldsymbol{\tilde{Z}_a} \equiv \boldsymbol{Z_a} - \boldsymbol{\mathcal{C}_a}\boldsymbol{Q_1'}\,\boldsymbol{R_1^{-1}}\boldsymbol{Z_q} \tag{A.65}$$

$$= \boldsymbol{\tilde{\mathcal{C}}_a}\,\boldsymbol{\mathcal{Y}_2} + \boldsymbol{\mathcal{D}n} \tag{A.66}$$

with $\tilde{\boldsymbol{\mathcal{C}}}_{\boldsymbol{a}} \equiv \boldsymbol{\mathcal{C}}_{\boldsymbol{a}} \boldsymbol{Q}_{\boldsymbol{2}}'$.

– Standard signal extraction formulas lead to the following posterior:

$$p\left(\boldsymbol{\mathcal{Y}_2}|\boldsymbol{Z_q}, \boldsymbol{Z_a}\right) \sim \mathcal{N}\left(\boldsymbol{\mu_2}, \boldsymbol{P_{22}^{-1}}\right) , \tag{A.67}$$

$$\text{with} \quad \boldsymbol{P_{22}} = \boldsymbol{P_{22|1}} + \tilde{\boldsymbol{\mathcal{C}}}_{\boldsymbol{a}}' \left(\boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{D}}'\right)^{-1} \tilde{\boldsymbol{\mathcal{C}}}_{\boldsymbol{a}} , \tag{A.68}$$

$$\text{and} \quad \boldsymbol{P_{22}}\, \boldsymbol{\mu_2} = \boldsymbol{P_{22|1}}\, \boldsymbol{\mu_{2|1}} + \tilde{\boldsymbol{\mathcal{C}}}_{\boldsymbol{a}}' \left(\boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{D}}'\right)^{-1} \tilde{\boldsymbol{Z}}_{\boldsymbol{a}} . \tag{A.69}$$

- Given a draw from $p\left(\boldsymbol{\mathcal{Y}_2}|\boldsymbol{Z_q}, \boldsymbol{Z_a}\right)$ we can construct a draw from $p\left(\boldsymbol{Y}|\boldsymbol{Z_q}, \boldsymbol{Z_a}\right)$ from

$$\boldsymbol{Y} = \boldsymbol{Q_1'}\, \boldsymbol{R_1^{-1}} \boldsymbol{Z_q} + \boldsymbol{Q_2'}\, \boldsymbol{\mathcal{Y}_2} . \tag{A.70}$$

II(d) Sampling of VAR coefficients with two-block SV model

Step 2 of the MCMC scheme outlined above involves drawing the VAR coefficients, $\tilde{\boldsymbol{\Pi}}$, conditional on the latent states, $\boldsymbol{Y}$, and draws of model parameters (including time-varying second moment parameters). As described in the paper, we employ a two-block SV model for the cyclical shocks. In our VAR specification, these cyclical shocks drive the VAR in the detrended SPF forecast updates, given by equation (9) of the paper (and restated above). Estimation of the VAR coefficients for the updates to (detrended) forecasts requires us to account for this source of heteroskedasticity. In general, estimation of VAR models with SV can be computationally intensive, and Carriero et al. (2022a) and Carriero et al. (2019) derive an efficient Gibbs sampling procedure for this case. We build on their sampler, and we develop a variant of it that exploits the specific two-block structure of the SV model. As such Step 2 of the MCMC sampler described above involves multiple steps, which we detail here:

- We condition on draws for $\{\tilde{\boldsymbol{\eta}}_{\boldsymbol{t}}\}_{t=0}^{T}$, as well as $\tilde{\boldsymbol{K}}$, $\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{ii}}$, $\{\lambda_{i,t}\}_{t=1}^{T}$ $\forall\, i = 1, 2$ and seek to sample the slope coefficients $\tilde{\boldsymbol{\Pi}}$ in the following regression, which restates the regressors of (9) in terms of $\boldsymbol{x_t} \equiv \tilde{\boldsymbol{\eta}}_{\boldsymbol{t-1}}$:

$$\tilde{\boldsymbol{\eta}}_{\boldsymbol{t}} = \tilde{\boldsymbol{\Pi}} \boldsymbol{x_t} + \tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{t}}, \tag{A.71}$$

where $\tilde{\varepsilon}_t$ is the vector of cyclical shocks described in (15). With a multivariate normal prior on $\tilde{\mathbf{\Pi}}$ and since the shock vector, $\tilde{\varepsilon}_t$ is multivariate normal, the posterior for $\tilde{\mathbf{\Pi}}$ is multivariate normal as well.[7] Thus, we can equivalently work with a linearly rotated matrix of VAR slopes, which will be convenient to do, as will be shown in the following steps.

- Following the two-block structure of the SV specification, we partition the vector regression into two sets of equations:

$$\begin{bmatrix} \tilde{\boldsymbol{\eta}}_{\mathbf{1},t} \\ \tilde{\boldsymbol{\eta}}_{\mathbf{2},t} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{\Pi}}_{\mathbf{1}} \\ \tilde{\mathbf{\Pi}}_{\mathbf{2}} \end{bmatrix} \boldsymbol{x}_t + \begin{bmatrix} \boldsymbol{I} & \tilde{\boldsymbol{K}} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\varepsilon}^*_{\mathbf{1},t} \\ \boldsymbol{\varepsilon}^*_{\mathbf{2},t} \end{bmatrix}, \tag{A.72}$$

  where $\boldsymbol{\varepsilon}^*_{\boldsymbol{i},\boldsymbol{t}} \sim \mathcal{N}(\mathbf{0}, \lambda_{i,t} \cdot \tilde{\mathbf{\Sigma}}_{\boldsymbol{ii}})$ for $i = 1, 2$ (as in (15)).

- When $\tilde{\boldsymbol{K}} \neq \mathbf{0}$, the VAR shocks are correlated across blocks and it is convenient to rotate the system into a decoupled form:

$$\text{Let} \quad \begin{bmatrix} \boldsymbol{\eta}^*_{\mathbf{1},t} \\ \boldsymbol{\eta}^*_{\mathbf{2},t} \end{bmatrix} \equiv \begin{bmatrix} \boldsymbol{I} & \tilde{\boldsymbol{K}} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\boldsymbol{\eta}}_{\mathbf{1},t} \\ \tilde{\boldsymbol{\eta}}_{\mathbf{2},t} \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} \mathbf{\Pi}^*_{\mathbf{1}} \\ \mathbf{\Pi}^*_{\mathbf{2}} \end{bmatrix} \equiv \begin{bmatrix} \boldsymbol{I} & \tilde{\boldsymbol{K}} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\mathbf{\Pi}}_{\mathbf{1}} \\ \tilde{\mathbf{\Pi}}_{\mathbf{2}} \end{bmatrix} \tag{A.73}$$

$$\text{and we obtain} \quad \begin{bmatrix} \boldsymbol{\eta}^*_{\mathbf{1},t} \\ \boldsymbol{\eta}^*_{\mathbf{2},t} \end{bmatrix} = \begin{bmatrix} \mathbf{\Pi}^*_{\mathbf{1}} \\ \mathbf{\Pi}^*_{\mathbf{2}} \end{bmatrix} \boldsymbol{x}_t + \begin{bmatrix} \boldsymbol{\varepsilon}^*_{\mathbf{1},t} \\ \boldsymbol{\varepsilon}^*_{\mathbf{2},t} \end{bmatrix}. \tag{A.74}$$

$$\text{Since} \quad \begin{bmatrix} \boldsymbol{I} & \tilde{\boldsymbol{K}} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{I} & -\tilde{\boldsymbol{K}} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix}, \tag{A.75}$$

  we have $\boldsymbol{\varepsilon}^*_{\mathbf{2},t} = \tilde{\boldsymbol{\varepsilon}}_{\mathbf{2},t}$, $\boldsymbol{\varepsilon}^*_{\mathbf{1},t} = \tilde{\boldsymbol{\eta}}_{\mathbf{1},t} - \tilde{\boldsymbol{K}}\tilde{\boldsymbol{\eta}}_{\mathbf{2},t}$, $\mathbf{\Pi}^*_{\mathbf{2}} = \tilde{\mathbf{\Pi}}_{\mathbf{2}}$, $\mathbf{\Pi}^*_{\mathbf{1}} = \tilde{\mathbf{\Pi}}_{\mathbf{1}} - \tilde{\boldsymbol{K}}\tilde{\mathbf{\Pi}}_{\mathbf{2}}$.

- If the priors for $\mathbf{\Pi}^*_{\mathbf{1}}$ and $\mathbf{\Pi}^*_{\mathbf{2}}$ are independent, both blocks of the decoupled VAR system in (A.74) could be separately estimated. However, this is generally not the case. In fact, in a typical application (including ours), a researcher might specify independent priors for $\tilde{\mathbf{\Pi}}_{\mathbf{1}}$

[7] The shock vector is multivariate normal conditional on $\boldsymbol{K}$, $\tilde{\mathbf{\Sigma}}_{\boldsymbol{ii}}$, $\{\lambda_{i,t}\}_{t=1}^{T}$ $\forall\, i = 1, 2$.

and $\tilde{\mathbf{\Pi}}_{\mathbf{2}}$. In that case, and with $\tilde{\boldsymbol{K}} \neq \mathbf{0}$, the priors for the rotated slopes, $\mathbf{\Pi}_{\mathbf{1}}^{*}$ and $\mathbf{\Pi}_{\mathbf{2}}^{*}$, will generally be correlated.

- Below, we derive a two-step Gibbs algorithm with the following steps:

  1. $p\left(\mathbf{\Pi}_{\mathbf{1}}^{*}|\mathbf{\Pi}_{\mathbf{2}}^{*}, \boldsymbol{\mathcal{I}}\right)$ with prior $p\left(\mathbf{\Pi}_{\mathbf{1}}^{*}|\mathbf{\Pi}_{\mathbf{2}}^{*}\right)$
  2. $p\left(\mathbf{\Pi}_{\mathbf{2}}^{*}|\mathbf{\Pi}_{\mathbf{1}}^{*}, \boldsymbol{\mathcal{I}}\right)$ with prior $p\left(\mathbf{\Pi}_{\mathbf{2}}^{*}|\mathbf{\Pi}_{\mathbf{1}}^{*}\right)$,

  where $\boldsymbol{\mathcal{I}} = \{\{\tilde{\boldsymbol{\eta}}_{\boldsymbol{t}}\}_{t=0}\}_{t=1}^{T}, \tilde{\boldsymbol{K}}, \tilde{\mathbf{\Sigma}}_{\boldsymbol{ii}}, \{\lambda_{i,t}\}_{t=1}^{T}\ \forall\, i = 1, 2\}$. These Gibbs steps are similar to the triangular algorithm of Carriero et al. (2022a), but specialized to the two-block case, and they proceed by operating on the rotated slopes (and their priors), rather than by adjusting the VAR equations with additional regressors (as in Carriero et al. (2022a)). Of course, for full estimation of our state space model, these two Gibbs steps are to be wrapped into a larger Gibbs sampler with additional steps for inference on the latent states, SV processes, etc. and as described elsewhere in our paper (or its appendix).

- In the general case, with $\tilde{\boldsymbol{K}} \neq \mathbf{0}$, we generally have $p\left(\mathbf{\Pi}_{\mathbf{1}}^{*}|\mathbf{\Pi}_{\mathbf{2}}^{*}\right) \neq p\left(\mathbf{\Pi}_{\mathbf{1}}^{*}\right)$ and $p\left(\mathbf{\Pi}_{\mathbf{2}}^{*}|\mathbf{\Pi}_{\mathbf{1}}^{*}\right) \neq p\left(\mathbf{\Pi}_{\mathbf{2}}^{*}\right)$ and these conditional priors are reevaluated at each step of the Gibbs sampler as described further below.[8]

- Apart from deriving $p\left(\mathbf{\Pi}_{\boldsymbol{i}}^{*}, \mathbf{\Pi}_{\boldsymbol{j\neq i}}^{*}|\boldsymbol{\mathcal{I}}\right)\ \forall\, i, j = 1, 2$, each of the two Gibbs steps amounts to a standard Gaussian vector regression subject to a scalar SV factor, $\lambda_{i,t}$, and can be efficiently sampled with standard methods.

- A given set of draws for $\mathbf{\Pi}_{\mathbf{1}}^{*}$ and $\mathbf{\Pi}_{\mathbf{2}}^{*}$ can then swiftly be transformed into draws for $\tilde{\mathbf{\Pi}}_{\mathbf{1}}$ and $\tilde{\mathbf{\Pi}}_{\mathbf{2}}$ by using (A.73). The joint draw for $\tilde{\mathbf{\Pi}}_{\mathbf{1}}$ and $\tilde{\mathbf{\Pi}}_{\mathbf{2}}$ is accepted only if the resulting transition matrix for the VAR, $\tilde{\mathbf{\Pi}}$, is stable.

Derivation of priors for the rotated slopes: To derive the conditional priors, $p\left(\mathbf{\Pi}_{\boldsymbol{i}}^{*}, \mathbf{\Pi}_{\boldsymbol{j\neq i}}^{*}|\boldsymbol{\mathcal{I}}\right)$, we begin with the (given) prior for the original VAR slopes, $\tilde{\mathbf{\Pi}}$. Since $\tilde{\mathbf{\Pi}}$ is a matrix (its sub-blocks $\tilde{\mathbf{\Pi}}_{\mathbf{1}}$ and $\tilde{\mathbf{\Pi}}_{\mathbf{2}}$ are also matrices), we consider priors for their vectorized forms. To keep better track

[8] Note: For the case of $\tilde{\boldsymbol{K}} = \mathbf{0}$, we have $p\left(\mathbf{\Pi}_{\mathbf{1}}^{*}|\mathbf{\Pi}_{\mathbf{2}}^{*}\right) = p\left(\mathbf{\Pi}_{\mathbf{1}}^{*}\right)$ and $p\left(\mathbf{\Pi}_{\mathbf{2}}^{*}|\mathbf{\Pi}_{\mathbf{1}}^{*}\right) = p\left(\mathbf{\Pi}_{\mathbf{2}}^{*}\right)$, and the system is perfectly decoupled. A single iteration over both steps generates a direct draw from the joint distribution of $p\left(\mathbf{\Pi}_{\mathbf{1}}^{*}, \mathbf{\Pi}_{\mathbf{2}}^{*}|\boldsymbol{\mathcal{I}}\right)$.

of the system's block structure, it is convenient to express the prior in terms of vectorizing the *transposed* slopes matrix, $\tilde{\boldsymbol{\Pi}}'$:[9]

$$\operatorname{vec}\left(\tilde{\boldsymbol{\Pi}}'\right) = \begin{bmatrix} \operatorname{vec}\left(\tilde{\boldsymbol{\Pi}}_{\mathbf{1}}'\right) \\ \operatorname{vec}\left(\tilde{\boldsymbol{\Pi}}_{\mathbf{2}}'\right) \end{bmatrix}. \tag{A.76}$$

Throughout, we assume a Gaussian prior for $\operatorname{vec}\left(\tilde{\boldsymbol{\Pi}}'\right)$ that is mean zero and has a (block-) diagonal variance-covariance matrix. Expressed in terms of precision matrices (i.e. inverse variance-covariance matrices), the prior has the following form:[10]

$$\begin{bmatrix} \operatorname{vec}\left(\tilde{\boldsymbol{\Pi}}_{\mathbf{1}}'\right) \\ \operatorname{vec}\left(\tilde{\boldsymbol{\Pi}}_{\mathbf{2}}'\right) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{P}_{\mathbf{11}}^{-\mathbf{1}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{P}_{\mathbf{22}}^{-\mathbf{1}} \end{bmatrix} \right). \tag{A.77}$$

The vector of rotated slopes is related to the vector of the original slopes as follows:[11]

$$\operatorname{vec}\left(\boldsymbol{\Pi}^{*\prime}\right) = \left(\begin{bmatrix} \boldsymbol{I} & -\tilde{\boldsymbol{K}} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix} \otimes \boldsymbol{I} \right) \operatorname{vec}\left(\tilde{\boldsymbol{\Pi}}'\right) \tag{A.78}$$

$$= \begin{bmatrix} \boldsymbol{I} & -\boldsymbol{\mathcal{K}} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix} \operatorname{vec}\left(\tilde{\boldsymbol{\Pi}}'\right), \qquad \text{with} \quad \boldsymbol{\mathcal{K}} \equiv \tilde{\boldsymbol{K}} \otimes \boldsymbol{I} \tag{A.79}$$

$$\Rightarrow \quad \operatorname{vec}\left(\boldsymbol{\Pi}_{\mathbf{1}}^{*\prime}\right) = \operatorname{vec}\left(\tilde{\boldsymbol{\Pi}}_{\mathbf{1}}'\right) - \boldsymbol{\mathcal{K}} \operatorname{vec}\left(\tilde{\boldsymbol{\Pi}}_{\mathbf{2}}'\right), \qquad \operatorname{vec}\left(\boldsymbol{\Pi}_{\mathbf{2}}^{*\prime}\right) = \operatorname{vec}\left(\tilde{\boldsymbol{\Pi}}_{\mathbf{2}}'\right). \tag{A.80}$$

The joint prior for $\boldsymbol{\Pi}_{\mathbf{1}}^{*}$ and $\boldsymbol{\Pi}_{\mathbf{2}}^{*}$ can then be expressed (and again using precision matrices) as

$$\begin{bmatrix} \operatorname{vec}\left(\boldsymbol{\Pi}_{\mathbf{1}}^{*\prime}\right) \\ \operatorname{vec}\left(\boldsymbol{\Pi}_{\mathbf{2}}^{*\prime}\right) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{P}_{\mathbf{11}}^{*} & \boldsymbol{P}_{\mathbf{12}}^{*} \\ \boldsymbol{P}_{\mathbf{21}}^{*} & \boldsymbol{P}_{\mathbf{22}}^{*} \end{bmatrix}^{-1} \right), \tag{A.81}$$

$$\text{where} \quad \begin{bmatrix} \boldsymbol{P}_{\mathbf{11}}^{*} & \boldsymbol{P}_{\mathbf{12}}^{*} \\ \boldsymbol{P}_{\mathbf{21}}^{*} & \boldsymbol{P}_{\mathbf{22}}^{*} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{\mathcal{K}} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix}' \begin{bmatrix} \boldsymbol{P}_{\mathbf{11}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{P}_{\mathbf{22}} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{\mathcal{K}} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix} \tag{A.82}$$

[9] Expressing the slopes' prior in terms of $\tilde{\boldsymbol{\Pi}}'$ also corresponds to much of the general literature on VAR systems, including Kadiyala and Karlsson (1997), and Carriero et al. (2019).

[10] Typical Minnesota-style priors, like the one used in our application, have diagonal variance-covariance matrices, thus also diagonal precision matrices $\boldsymbol{P}_{\mathbf{11}}$ and $\boldsymbol{P}_{\mathbf{22}}$. The extension to non-diagonal priors and variance-covariance matrices is straightforward.

[11] Throughout, each use of $\boldsymbol{I}$ denotes an identity matrix of conformable size, so that repeated uses of $\boldsymbol{I}$ need not refer to identically-sized identity matrices.

$$= \begin{bmatrix} \boldsymbol{P_{11}} & \boldsymbol{P_{11}\mathcal{K}} \\ \boldsymbol{\mathcal{K}' P_{11}} & \boldsymbol{\mathcal{K}' P_{11}\mathcal{K} + P_{22}} \end{bmatrix}, \tag{A.83}$$

and, using standard signal-extraction formulas, the conditional priors then follow as:

$$\operatorname{vec}\left(\boldsymbol{\Pi_1^{*\prime}}\right) \Big| \operatorname{vec}\left(\boldsymbol{\Pi_2^{*\prime}}\right) \sim \mathcal{N}\left(-\boldsymbol{\mathcal{K}} \operatorname{vec}\left(\boldsymbol{\Pi_2^{*\prime}}\right), \boldsymbol{P_{11}^{-1}}\right), \tag{A.84}$$

$$\operatorname{vec}\left(\boldsymbol{\Pi_2^{*\prime}}\right) \Big| \operatorname{vec}\left(\boldsymbol{\Pi_1^{*\prime}}\right) \sim \mathcal{N}\left(-\boldsymbol{P_{22}^{-1}\mathcal{K}' P_{11}} \operatorname{vec}\left(\boldsymbol{\Pi_1^{*\prime}}\right), \left(\boldsymbol{\mathcal{K}' P_{11}\mathcal{K} + P_{22}}\right)^{-1}\right). \tag{A.85}$$

For MCMC estimation, an additional step is added to the sampler described above for the MDS version of the SV model. Between steps 1 and 2 of the sampler, we draw from

$$p\left(\tilde{\boldsymbol{\Pi}} \middle| \boldsymbol{Z}, \boldsymbol{Y}, \boldsymbol{\lambda}, \sigma_*^2, \tilde{\boldsymbol{\Sigma}}, \sigma_\nu^2\right) = p\left(\tilde{\boldsymbol{\Pi}} \middle| \boldsymbol{Y}, \boldsymbol{\lambda}, \tilde{\boldsymbol{\Sigma}}\right),$$

which is a standard conjugate-normal Bayesian regression update. (Furthermore, $\tilde{\boldsymbol{\Pi}}$ is added to the conditioning sets of the other steps).

III Coibion-Gorodnichenko slopes implied by VAR model

This appendix describes our calculations of model-implied slopes for regressions testing the efficiency of SPF forecasts known from the work of Coibion and Gorodnichenko (2015). We derive these slopes within our VAR model for forecast updates. To restate relevant parts of the model, we denote the vector of forecast updates (including the lagged nowcast error, and change in long-run forecast) by $\boldsymbol{\eta}_t$, and the VAR model with SV specifies the following:

$$\boldsymbol{\eta_t} \equiv F_t \begin{bmatrix} y_{t-1} \\ y_t \\ y_{t+1} \\ \vdots \\ y_{t+H-1} \\ y_{t+H} \end{bmatrix} - F_{t-1} \begin{bmatrix} y_{t-1} \\ y_t \\ y_{t+1} \\ \vdots \\ y_{t+H-1} \\ y_{t+H-1} \end{bmatrix} \tag{A.86}$$

$$= \tilde{\boldsymbol{\eta}}_{\boldsymbol{t}} + \mathbf{1} w_t^* + \bar{\boldsymbol{\eta}}, \tag{A.87}$$

$$\text{with} \quad \tilde{\boldsymbol{\eta}}_{\boldsymbol{t}} = \tilde{\boldsymbol{\Pi}}\, \tilde{\boldsymbol{\eta}}_{\boldsymbol{t-1}} + \tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{t}}\,, \tag{A.88}$$

$$\text{and} \quad \tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{t}} \sim N(\mathbf{0}, \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{t}})\,, \tag{A.89}$$

$$w_t^* \sim N(0, \omega_t^2)\,. \tag{A.90}$$

Since the SPF is assumed to know the lagged realized value, we have $F_t y_{t-1} = y_{t-1}$, and the top element of $\boldsymbol{\eta_t}$ is identical to $e_{t-1} = y_{t-1} - F_{t-1}y_{t-1}$. The VAR's transition matrix, $\tilde{\boldsymbol{\Pi}}$, is required to be stable (i.e. all eigenvalues inside the unit circle).

For each MCMC draw of the model parameters, we derive regression slopes from the population moments of the model. For the time-varying variance parameters, $\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{t}}$ and ω_t^2, we use fixed values (per MCMC draw) as follows: For the trend shocks, that are generated from a horseshoe model, $\omega_t^2 \sim \mathcal{HS}(\tau_w^2)$, we use the global scale parameter τ_w^2. This choice is motivated by the notion that the local scale parameter represents occasional shifts in trend levels that are ignored for sake of computing CG slopes. In lieu of $\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{t}}$, at a given MCMC draw, we use the median values over the sampled (time-series) paths for $\lambda_{1,t}$ and $\lambda_{2,t}$ to construct the corresponding (constant) value for $\tilde{\boldsymbol{\Sigma}}$.

III(a) CG regressions in population

Given values for $\tilde{\boldsymbol{\Sigma}}$, ω^2, and $\tilde{\boldsymbol{\Pi}}$, we can solve for the variance of the unconditional (or steady-state) variance of $\boldsymbol{\eta_t}$ as follows:

$$\text{Var}\,(\boldsymbol{\eta_t}) \equiv \boldsymbol{\Gamma} = \tilde{\boldsymbol{\Gamma}} + \mathbf{1}\mathbf{1}'\omega\,, \qquad \text{with} \quad \tilde{\boldsymbol{\Gamma}} = \tilde{\boldsymbol{\Pi}}\, \tilde{\boldsymbol{\Gamma}}\, \tilde{\boldsymbol{\Pi}}' + \tilde{\boldsymbol{\Sigma}}\,, \tag{A.91}$$

where the second equation is a standard Lyapunov equation. Given a (positive definite) solution for $\tilde{\boldsymbol{\Gamma}}$, autocovariances of $\boldsymbol{\eta_t}$ follow as $\text{Cov}\,(\boldsymbol{\eta}_{t+h}, \boldsymbol{\eta_t}) = \tilde{\boldsymbol{\Pi}}^h\, \tilde{\boldsymbol{\Gamma}}\ \forall h > 0$.[12]

Based on the population variances $\text{Var}\,(\boldsymbol{\eta_t})$ and autocovariances $\text{Cov}\,(\boldsymbol{\eta}_{t+h}, \boldsymbol{\eta_t})$, we want to compute regression slopes, b_h, for the following "CG" regression known from Coibion and Gorodnichenko (2015):

$$(1 - F_t)y_{t+h} = a_h + b_h \cdot (F_t - F_{t-1})y_{t+h} + e_{t+h} \tag{A.92}$$

for horizons $h = 0, 1, 2, \ldots$ etc., with

[12] Equation (A.91) is a standard Lyapunov equation and can be solved analytically. The solution exists since $\tilde{\boldsymbol{\Pi}}$ is stable, and it is positive definite since $\tilde{\boldsymbol{\Sigma}}$ is.

$$b_h = \frac{\text{Cov}\left((1-F_t)y_{t+h}, (F_t - F_{t-1})y_{t+h}\right)}{\text{Var}\left((F_t - F_{t-1})y_{t+h}\right)}. \tag{A.93}$$

The regressor of the CG regression represents a sum of forecast updates at different points in time:[13]

$$(1-F_t)y_{t+h} = \sum_{k=0}^{h}(F_{t+k+1} - F_{t+k})y_{t+h}\,. \tag{A.94}$$

Let $\boldsymbol{e}(k)$ denote a selection vector, defined for $k = -1, 0, 1 \ldots, H$, that selects the $(k+2)$th element of a vector of length $H+2$ so that $(F_t - F_{t-1})y_{t+k} = \boldsymbol{e}(k)'\,\boldsymbol{\eta_t}$ ($\forall\, k < H$), and we get:

$$\text{Cov}\left((1-F_t)y_{t+h}, (F_t - F_{t-1})y_{t+h}\right) = \sum_{k=0}^{h}\text{Cov}\left(F_{t+k+1} - F_{t+k})y_{t+h}, (F_t - F_{t-1})y_{t+h}\right) \tag{A.95}$$

$$= \left(\sum_{k=0}^{h}\boldsymbol{e}(h-k-1)'\tilde{\boldsymbol{\Pi}}^{k+1}\right)\tilde{\boldsymbol{\Gamma}}\,\boldsymbol{e}(h)\,, \tag{A.96}$$

$$\text{and} \quad \text{Var}\left((F_t - F_{t-1})y_{t+h}\right) = \boldsymbol{e}(h)'\,\boldsymbol{\Gamma}\,\boldsymbol{e}(h) \tag{A.97}$$

$$= \boldsymbol{e}(h)'\,\tilde{\boldsymbol{\Gamma}}\,\boldsymbol{e}(h) + \omega^2 \tag{A.98}$$

and the CG slope can be computed as follows:

$$b_h = \frac{\left(\sum_{k=0}^{h}\boldsymbol{e}(h-k-1)'\tilde{\boldsymbol{\Pi}}^{k+1}\right)\tilde{\boldsymbol{\Gamma}}\,\boldsymbol{e}(h)}{\boldsymbol{e}(h)'\,\tilde{\boldsymbol{\Gamma}}\,\boldsymbol{e}(h) + \omega^2}\,. \tag{A.99}$$

When pooling the CG slopes across $j = 0, 2, \ldots, h$, the pooled slope is:[14]

$$\bar{b}_{0:h} = \frac{\sum_{j=1}^{h}\left[\left(\sum_{k=0}^{j}\boldsymbol{e}(h-k-1)'\tilde{\boldsymbol{\Pi}}^{k+1}\right)\tilde{\boldsymbol{\Gamma}}\,\boldsymbol{e}(j)\right]}{\sum_{j=1}^{h}\left[\boldsymbol{e}(j)'\,\tilde{\boldsymbol{\Gamma}}\,\boldsymbol{e}(j)\right] + h\cdot\omega^2}\,. \tag{A.100}$$

Since the common trend is a martingale (and thus an efficient forecast), the CG coefficients will be smaller the larger the trend-shock variance, ω^2.

[13] Recall that $F_{t+h+1}y_{t+h} = y_{t+h}$.

[14] See, for example, Chapter 4 of Hayashi (2000).

III(b) Estimated CG slopes

Table A.6 reports the slopes of Coibion-Gorodnichenko regressions for each variable implied by our VAR specifications fit to SPF forecasts. We compute the population slopes, as described above, for each draw of model parameters obtained from Bayesian MCMC estimation of the VAR model and report the posterior moments of the pooled slopes, $\bar{b}_{0:h}$, for $h = 3$. The horizons considered correspond to what is available in terms of observed SPF fixed-horizon forecasts, and thus commonly used in empirical work.[15]

Table A.6: Slopes of Coibion-Gorodnichenko regressions

	2019Q4			2023Q4		
Variable	5%	50%	95%	5%	50%	95%
RGDP	0.01	0.11	0.22	0.02	0.12	0.23
UNRATE	0.08	0.18	0.30	0.06	0.15	0.27
PGDP	0.00	0.13	0.29	0.04	0.18	0.35
CPI	0.11	0.22	0.33	0.14	0.25	0.36

Notes: Posterior moments of model-implied slopes, b_h, in predictability regressions of Coibion and Gorodnichenko (2015), $(1 - F_t)y_{t+h} = b_h \cdot (F_t - F_{t-1})y_{t+h} + e_{t+h}$, pooled for $h = 0, 1, 2, 3$. Estimated from our VAR model, using full-sample data through 2019Q4 and 2023Q4, respectively.

These estimates are generally in line with the literature that finds positive coefficients of small-to-modest magnitudes, indicating some departures from full rationality in professional forecasts. For the sample of forecasts through 2023Q4, across variables the posterior medians range from 0.11 (RGDP) to 0.24 (CPI inflation), with 90 percent credible sets that do not include 0, except in the case of the unemployment rate. Quantitatively, our estimates are broadly comparable to those in surveys such as Angeletos et al. (2021).[16] Based on this evidence, it appears that our VAR model of SPF forecasts can capture reasonably well the empirical extent of non-rationality emphasized by Coibion and Gorodnichenko and subsequent studies.

IV Model-implied IMA representation for outcome process

Our state space model embeds a process for the outcome variable, y_t, that is driven by multiple shocks, most of which represent expectational updates at various horizons. As described in Clark

[15] Note that the right-hand side of (A.92) involves a forecast for $h + 1$ steps ahead, so that computation of $\bar{b}_{0:3}$ involves all observable fixed-horizon forecasts.

[16] For example, for a sample of (SPF) unemployment and inflation forecasts for 1984-2017 at a horizon of h = 3, these authors report a coefficient of 0.292 (their Table 1).

et al. (2020), such a representation can emerge from a wide class of models with multiple drivers, including DSGE models, that have a (conditionally linear) state-space representation. For ease of reference, we can also describe the model-implied process for y_t in terms of a univariate representation driven by innovations defined relative to the history of y_t alone:

$$\varepsilon_t \equiv y_t - E(y_t|y^{t-1})\,, \qquad \text{with} \quad y^{t-1} \equiv \{y_{t-1}, y_{t-2}, \ldots\}\,. \tag{A.101}$$

Here we show that our MDS model implies a univariate IMA$(1, H)$ process for y_t, which can be derived from the innovations representation of our state space model. Further below, we also show that, due to the added persistence in its forecast updates, the VAR model implies a univariate IMA$(1, \infty)$ representation.

In this context, the conditional expectations operator, $E(\cdot|y^{t-1})$, is understood to condition also on given values for the model's parameters (like $\tilde{\boldsymbol{\Sigma}}_{\mathbf{11}}$, $\tilde{\boldsymbol{\Sigma}}_{\mathbf{22}}$, $\tilde{\boldsymbol{K}}$, and τ_w^2, or $\tilde{\boldsymbol{\Pi}}$ in the case of the VAR model). For simplicity, we derive a time-invariant process of y_t that abstracts from variations in stochastic volatility. To do so, and similar to our derivation of CG slopes in Appendix III(b), we work with a time-invariant representation of our state space model, obtained by holding fixed the time-varying second moments of the model, $\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{t}}$ and ω_t^2. All computations are performed draw-by-draw from the model's MCMC output. As in Appendix III(b), at every MCMC draw, we employ a fixed value for $\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{t}}$ that is constructed using the median values over the sampled (time-series) paths for $\lambda_{1,t}$ and $\lambda_{2,t}$, and we replace $\omega_t^2 \sim \mathcal{HS}(\tau_w^2)$ by the corresponding draw of the global scale parameter τ_w^2.

IV(a) Univariate process for y_t implied by MDS model

The MDS model is characterized by the following state equations (and assuming constant variances, as discussed above):

$$\tilde{\boldsymbol{Y}}_{\boldsymbol{t}} = \tilde{\boldsymbol{\Psi}}\tilde{\boldsymbol{Y}}_{\boldsymbol{t-1}} + \tilde{\boldsymbol{\eta}}_{\boldsymbol{t}}\,, \qquad \tilde{\boldsymbol{\eta}}_{\boldsymbol{t}} \sim \mathcal{N}(\mathbf{0}, \tilde{\boldsymbol{\Sigma}})\,, \tag{A.102}$$

$$y_t^* = y_{t-1}^* + w_t^*\,, \qquad w_t^* \sim \mathcal{N}(0, \tau_w^2)\,. \tag{A.103}$$

We derive an innovations representation based on the following univariate measurement equation:

$$y_t = \boldsymbol{e_1}\tilde{\boldsymbol{Y}}_{\boldsymbol{t+1}} + y^*_{t+1}\,, \tag{A.104}$$

where $\boldsymbol{e_1}$ is a row vector that selects the first element of $\tilde{\boldsymbol{Y}}_{\boldsymbol{t+1}}$.[17] To derive the innovations process for y^t, we need to define the projections of state variables onto y^t:

$$\tilde{\boldsymbol{Y}}_{\boldsymbol{t+1|t}} \equiv E\left(\tilde{\boldsymbol{Y}}_{\boldsymbol{t+1}} \mid y^t\right) = \tilde{\boldsymbol{\Psi}}\tilde{\boldsymbol{Y}}_{\boldsymbol{t|t-1}} + \tilde{\boldsymbol{\kappa}}\,\varepsilon_t\,, \tag{A.105}$$

$$= \left(\boldsymbol{I} - \tilde{\boldsymbol{\Psi}}L\right)^{-1}\tilde{\boldsymbol{\kappa}}\,\varepsilon_t\,, \tag{A.106}$$

and
$$y^*_{t+1|t} \equiv E\left(y^*_{t+1} \mid y^t\right) = y^*_{t|t-1} + \kappa^*\,\varepsilon_t\,, \tag{A.107}$$

$$\Leftrightarrow \qquad (1-L)\,y^*_{t+1|t} = \kappa^*\,\varepsilon_t\,, \tag{A.108}$$

where $\tilde{\boldsymbol{\kappa}}$ and κ^* are steady-state Kalman gains for each component of the state vector. The lag operator L is understood to shift both the timing and conditioning sets of the projections, as in $L \cdot \tilde{\boldsymbol{Y}}_{\boldsymbol{t|t}} = \tilde{\boldsymbol{Y}}_{\boldsymbol{t-1|t-1}}$ and analogously for $y^*_{t|t}$ and ε_t. To derive the steady-state Kalman gains, consider the following state space:

$$\boldsymbol{\mathcal{S}_{t+1}} = \begin{bmatrix} \tilde{\boldsymbol{Y}}_{\boldsymbol{t+1}} \\ y^*_{t+1} \end{bmatrix} = \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{S}_t} + \boldsymbol{\mathcal{B}}\boldsymbol{w_{t+1}}\,, \qquad \boldsymbol{w_t} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})\,, \tag{A.109}$$

$$y_t = \boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{S}_{t+1}}\,, \tag{A.110}$$

with

$$\boldsymbol{\mathcal{A}} = \begin{bmatrix} \tilde{\boldsymbol{\Psi}} & \boldsymbol{0} \\ \boldsymbol{0} & 1 \end{bmatrix}, \qquad \boldsymbol{\mathcal{B}} = \begin{bmatrix} \tilde{\boldsymbol{\Sigma}}^{1/2} & 0 \\ 0 & \tau_w \end{bmatrix}, \qquad \boldsymbol{\mathcal{C}} = \begin{bmatrix} \boldsymbol{e_1} & 1 \end{bmatrix}. \tag{A.111}$$

And the Kalman gains are obtained as follows:

$$\boldsymbol{\kappa} = \begin{bmatrix} \tilde{\boldsymbol{\kappa}} \\ \kappa^* \end{bmatrix} = \boldsymbol{\mathcal{V}}\boldsymbol{\mathcal{C}}'\left(\boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{V}}\boldsymbol{\mathcal{C}}'\right)^{-1}, \tag{A.112}$$

where $\boldsymbol{\mathcal{V}}$ solves the following Riccati equation:

[17] Reflecting the timing assumptions in our state space model, y_t is captured only by $t+1$ state variables. These assumption reflect the data flow of economic data releases relative to the SPF at time $t+1$.

$$\boldsymbol{\mathcal{V}} = \boldsymbol{\mathcal{A}}\left(\boldsymbol{\mathcal{V}} - \boldsymbol{\mathcal{V}}\boldsymbol{\mathcal{C}}'\left(\boldsymbol{\mathcal{C}}\,\boldsymbol{\mathcal{V}}\,\boldsymbol{\mathcal{C}}'\right)^{-1}\boldsymbol{\mathcal{C}}\,\boldsymbol{\mathcal{V}}\right)\boldsymbol{\mathcal{A}}' + \boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}'\,. \tag{A.113}$$

To derive the univariate innovations process for y_t, note that the state equations imply $\tilde{\boldsymbol{Y}}_{\boldsymbol{t+1|t}} = \tilde{\boldsymbol{\Psi}}\tilde{\boldsymbol{Y}}_{\boldsymbol{t|t-1}}$ and $y^*_{t+1|t} = y^*_{t|t-1}$. With $y_t = \varepsilon_t + \boldsymbol{e_1}\,\tilde{\boldsymbol{\Psi}}\,\tilde{\boldsymbol{Y}}_{\boldsymbol{t|t-1}} + y^*_{t|t-1}$, we obtain:

$$(1-L)y_t = \left(1 - (1-\kappa^*)\,L + (1-L)\,\boldsymbol{e_1}\tilde{\boldsymbol{\Psi}}\,\left(\boldsymbol{I} - \tilde{\boldsymbol{\Psi}}L\right)^{-1}\tilde{\boldsymbol{\kappa}}\,L\right)\varepsilon_t\,. \tag{A.114}$$

To determine the number of MA lags on the right-hand side of (A.114), it is useful to note that for $\tilde{\boldsymbol{\Psi}}$ as defined in equation (8) of the paper, we have $\tilde{\boldsymbol{\Psi}}^{H+1+j} = \boldsymbol{0}$ for all $j > 0$:[18]

$$\left(\boldsymbol{I} - \tilde{\boldsymbol{\Psi}}L\right)^{-1} = \sum_{j=0}^{H+1}(\tilde{\boldsymbol{\Psi}}\,L)^i \quad \Rightarrow \quad \tilde{\boldsymbol{\Psi}}\left(\boldsymbol{I} - \tilde{\boldsymbol{\Psi}}L\right)^{-1}L = \sum_{j=1}^{H+1}\tilde{\boldsymbol{\Psi}}^j\,L^j\,. \tag{A.115}$$

In total, it follows that there are $H+2$ MA lags on the right-hand side of (A.114), and the process of y_t is an integrated moving-average (IMA) process with integration order of one and H MA lags, or, in short, IMA$(1, H+2)$. As discussed above, there is some choice in selecting the length, H, of the cyclical state vector, $\tilde{\boldsymbol{Y}}_{\boldsymbol{t}}$, used to track the term structure of SPF consistent forecasts. As shown here, this choice of H is reflected in the number of MA lags in the model-implied univariate process for y_t.

Moreover, it is straightforward to verify that for all $j < H+2$ we have $\boldsymbol{e_1}\tilde{\boldsymbol{\Psi}}^j = \boldsymbol{e_{j+1}}$, where $\boldsymbol{e_j}$ is an $H+2$ dimensional row vector that selects the jth row.[19] Denoting the jth element of the Kalman gain vector $\tilde{\boldsymbol{\kappa}}$ by $\tilde{\kappa}_j \equiv \boldsymbol{e_j}\tilde{\boldsymbol{\kappa}}$, the IMA representation simplifies as follows:

$$(1-L)y_t = \left(1 - (1-\kappa^*)\cdot L + \sum_{j=1}^{H+1}\boldsymbol{e_1}\,\tilde{\boldsymbol{\Psi}}^j\,\tilde{\boldsymbol{\kappa}}\cdot(L^j - L^{j+1})\right)\cdot\varepsilon_t \tag{A.116}$$

$$= \left(1 - (1-\kappa^*)\cdot L + \sum_{j=1}^{H+1}\tilde{\kappa}_{j+1}\cdot(L^j - L^{j+1})\right)\cdot\varepsilon_t\,. \tag{A.117}$$

In the MDS model, the moving average coefficients in the IMA representation of y_t thus correspond to the Kalman gains involved in forming projections of trend and cyclical states that track SPF-consistent expectations at different forecast horizons. When the trend gain, κ^*, is close to

[18] Recall that $\boldsymbol{Y_t}$ contains not only forecasts for horizons $h = 1, 2, \ldots, H$ but also the nowcast and the lagged realized value, y_{t-1}, so that $\boldsymbol{Y_t}$ has $H+2$ elements, and the transition matrix $\tilde{\boldsymbol{\Psi}}$ is of dimension $(H+2)\times(H+2)$.

[19] Reflecting differences in context, the definitions of selection vectors $\boldsymbol{e_j}$ here and $\boldsymbol{e}(h)$ in Section III(b) are similar, but differ in that $\boldsymbol{e}(h)$ is a column vector selecting the $(k+2)th$ element.

zero, the unit-root factor $(1-L)$ nearly cancels on both sides of (A.117). Of course, the case of $\kappa^* \approx 0$ corresponds to situations where the trend shock variance is negligible relative to the variability of cyclical shocks in the model, and y_t is essentially a stationary variable, and its process is close to an MA$(H-1)$. In terms of the level of y_t we can also write:

$$y_t = \left(1 + \sum_{j=1}^{H+1} \tilde{\kappa}_{j+1} \cdot L^j\right) \cdot \varepsilon_t + y^*_{t-1|t-1}\,, \tag{A.118}$$

$$\text{with} \quad y^*_{t|t} = \kappa^* \cdot \sum_{j=0}^{\infty} \varepsilon_{t-j}. \tag{A.119}$$

IV(b) Univariate process for y_t implied by VAR model

The derivation of a univariate innovations representation for y_t in the VAR model is similar to the steps described above for the MDS case. The main difference is that (ignoring constants) the gap process for $\tilde{\boldsymbol{Y}}_{\boldsymbol{t}}$ has a richer VMA representation:

$$\tilde{\boldsymbol{Y}}_{\boldsymbol{t}} = \tilde{\boldsymbol{\Phi}}(L)^{-1}\,\tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{t}}\,, \qquad \text{with} \quad \tilde{\boldsymbol{\Phi}}(L) = \left(\boldsymbol{I} - \tilde{\boldsymbol{\Pi}}L\right)\left(\boldsymbol{I} - \tilde{\boldsymbol{\Psi}}L\right). \tag{A.120}$$

The roots of $\tilde{\boldsymbol{\Phi}}(L)$ are the union of the roots of $(\boldsymbol{I} - \tilde{\boldsymbol{\Pi}}L)$ and $(\boldsymbol{I} - \tilde{\boldsymbol{\Psi}}L)$, and the former are generally non-zero so that the VMA lags of $\tilde{\boldsymbol{\Phi}}(L)^{-1}$ vanish only asymptotically. Nevertheless, the derivation of the process for y_t is isomorphic for the MDS case and results in the following IMA$(1,\infty)$ representation:[20]

$$(1-L)y_t = \left(1 - (1-\kappa^*)\,L + (1-L)\,\boldsymbol{e_1}\left(\tilde{\boldsymbol{\Pi}} + \tilde{\boldsymbol{\Psi}}\right)\,\tilde{\boldsymbol{\Phi}}(L)^{-1}\tilde{\boldsymbol{\kappa}}\,L\right)\varepsilon_t\,. \tag{A.122}$$

$$y_t = \left(1 + \boldsymbol{e_1}\left(\tilde{\boldsymbol{\Pi}} + \tilde{\boldsymbol{\Psi}}\right)\,\tilde{\boldsymbol{\Phi}}(L)^{-1}\tilde{\boldsymbol{\kappa}}\,L\right)\cdot\varepsilon_t + y^*_{t|t-1}\,. \tag{A.123}$$

IV(c) Estimates of the IMA process for y_t

To illustrate estimates of the IMA process for y_t, we compute impulse responses of y_t in response to an innovation of the IMA process (denoted ε_t in the derivations above). Figures A.15 through A.18

[20] To derive the Kalman gains for the VAR model, we use the following augmented state space matrices:

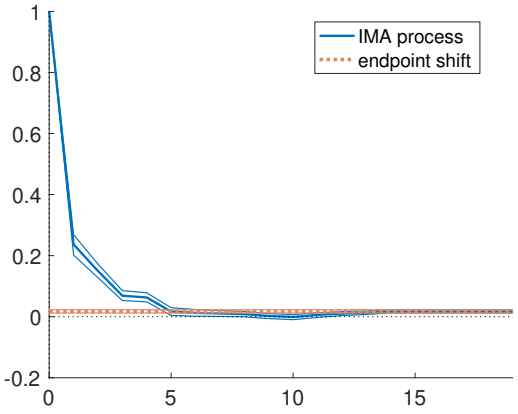
$$\boldsymbol{\mathcal{A}} = \begin{bmatrix} \tilde{\boldsymbol{\Psi}} + \tilde{\boldsymbol{\Pi}} & -\tilde{\boldsymbol{\Psi}}\,\tilde{\boldsymbol{\Pi}} & \boldsymbol{0} \\ \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & 1 \end{bmatrix}, \qquad \boldsymbol{\mathcal{B}} = \begin{bmatrix} \tilde{\boldsymbol{\Sigma}}^{1/2} & \boldsymbol{0} \\ \boldsymbol{0} & 0 \\ \boldsymbol{0} & \tau_w \end{bmatrix}, \qquad \boldsymbol{\mathcal{C}} = \begin{bmatrix} \boldsymbol{e_1} & 0 & 1 \end{bmatrix}. \tag{A.121}$$

report these impulse responses for each variable, based on estimates from MDS and VAR models and using data through 2019Q4 or 2024Q1. The estimates differ mainly across variables, and are fairly unchanged when including or excluding data since the onset of the COVID-19 pandemic. The general contours of the estimated impulse responses for each variable are also quite similar using the MDS or VAR models.

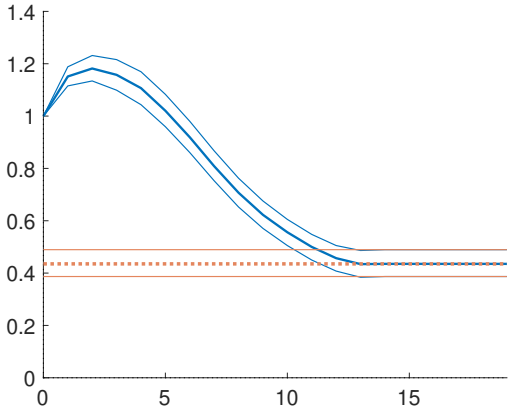
By construction, all responses are equal to 1 on impact. For GDP growth and CPI and GDP price inflation, the estimated responses then return within a couple of quarters to the (new) steady state, which is barely changed for GDP growth and GDP price inflation. Even for CPI inflation, a unit surprise in the realized data raises the endpoint of the term structure of expectations by no more than 10 basis points. Of course, all three of these variables measure rates of change, and the relatively low persistence embodied in their estimated impulse responses is consistent with that. In contrast, the unemployment is a more persistent variable, which is also borne out by the estimated hump shape in its impulse responses. In both the MDS and VAR models, the unemployment rate response peaks about three quarters after impact and at responses that exceed the impact value by about one to two fifths. For the MDS model, the peak of its hump-shaped response is followed by a relatively gradual decline, that declines below 0.6 after only about 8 quarters. In contrast, the VAR model generates a sharper peak in the unemployment rate response followed by a swifter decline (reaching the new steady state after about 8 quarters).

Figure A.15: Univariate process for y_t (MDS, 2019Q4)

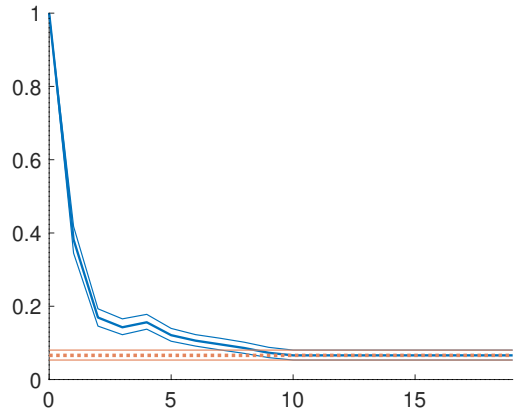
(a) Real GDP growth



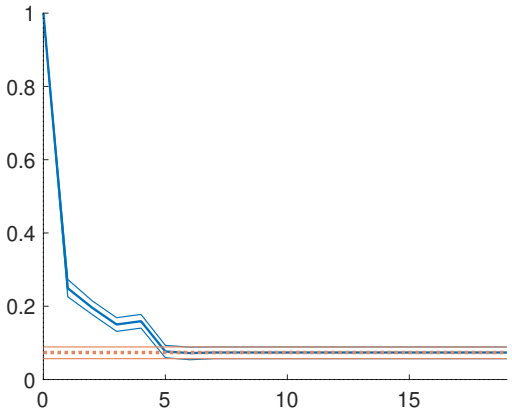
(b) Unemployment rate



(c) CPI inflation



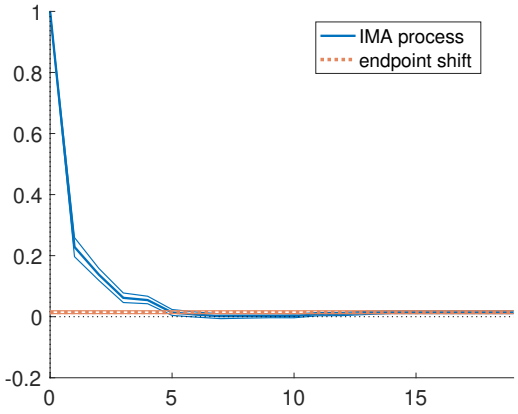
(d) GDP price inflation



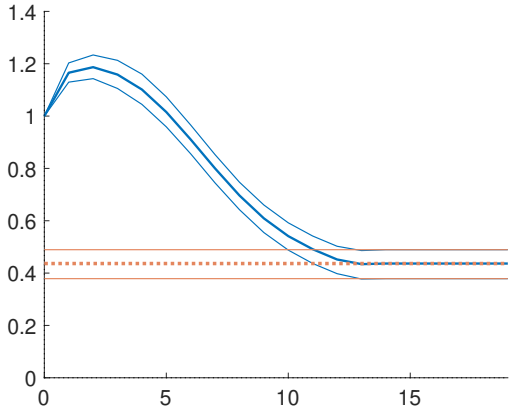
Notes: Impulse responses of univariate innovations representation for y_t implied by our state space for the MDS model. Posterior median and 68% uncertainty bands obtained from data through 2019Q4. The dashed (orange) line depicts the shift in endpoint induced by an innovation to y_t.

Figure A.16: Univariate process for y_t (MDS, 2024Q1)

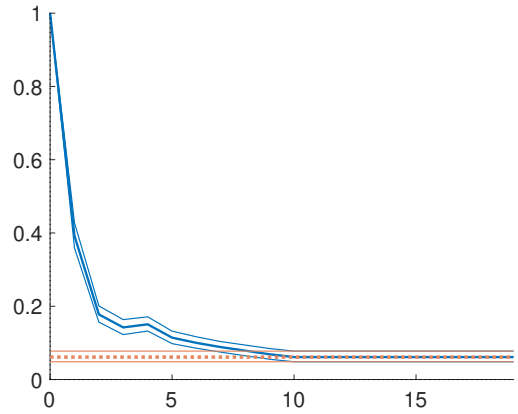
(a) Real GDP growth



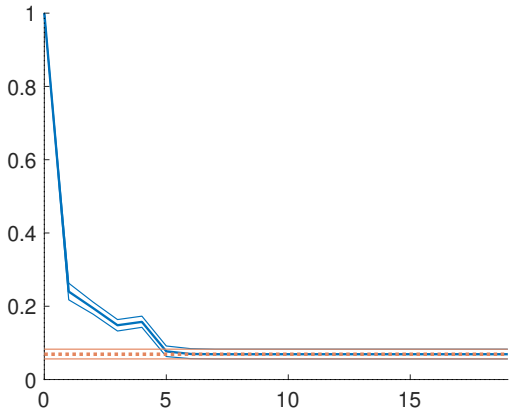
(b) Unemployment rate



(c) CPI inflation



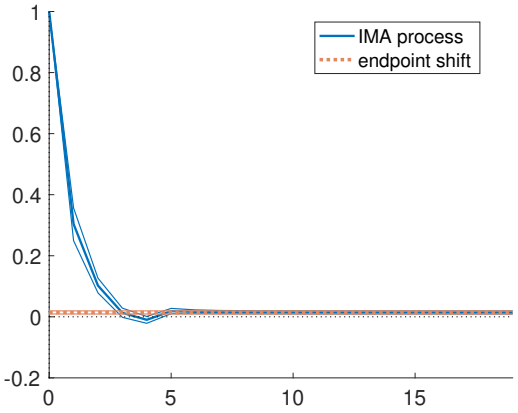
(d) GDP price inflation



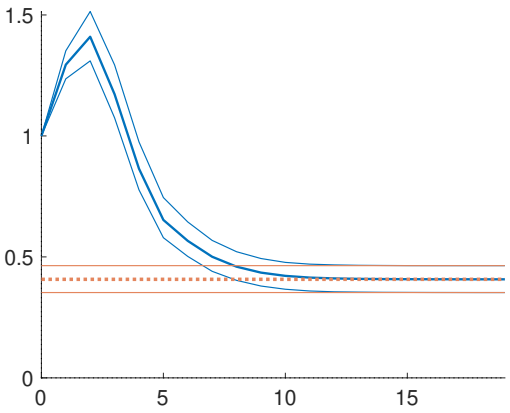
Notes: Impulse responses of univariate innovations representation for y_t implied by our state space for the MDS model. Posterior median and 68% uncertainty bands obtained from data through 2024Q1. The dashed (orange) line depicts the shift in endpoint induced by an innovation to y_t.

Figure A.17: Univariate process for y_t (VAR, 2019Q4)

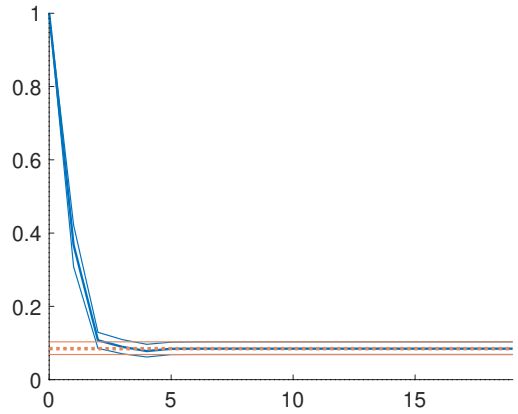
(a) Real GDP growth



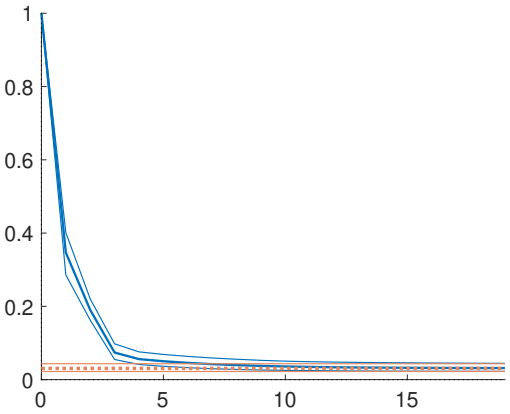
(b) Unemployment rate



(c) CPI inflation



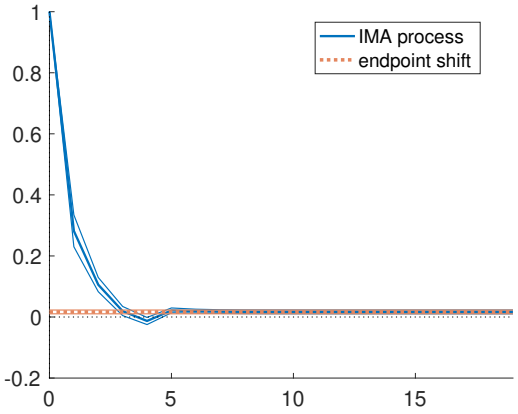
(d) GDP price inflation



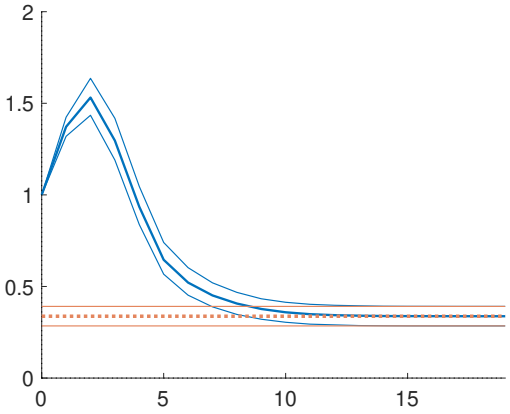
Notes: Impulse responses of univariate innovations representation for y_t implied by our state space for the MDS model. Posterior median and 68% uncertainty bands obtained from data through 2019Q4. The dashed (orange) line depicts the shift in endpoint induced by an innovation to y_t.

Figure A.18: Univariate process for y_t (VAR, 2024Q1)

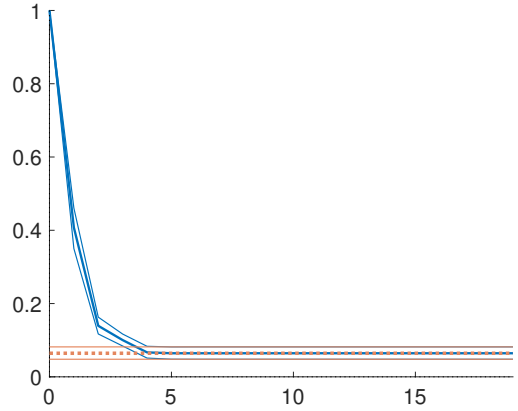
(a) Real GDP growth



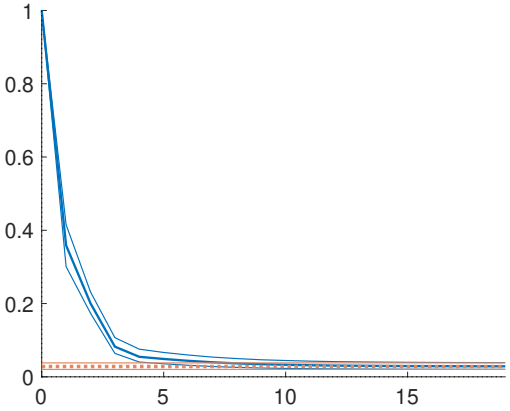
(b) Unemployment rate



(c) CPI inflation



(d) GDP price inflation



Notes: Impulse responses of univariate innovations representation for y_t implied by our state space for the MDS model. Posterior median and 68% uncertainty bands obtained from data through 2024Q1. The dashed (orange) line depicts the shift in endpoint induced by an innovation to y_t.

V Additional results

This section provides additional results, covering additional variables and further details that were not shown in the paper for the sake of brevity. These include: tables of results for a sample ending in 2019Q4, rather than the paper's sample end of 2023Q4, for Mincer-Zarnowitz predictability regressions (Tables A.7 and A.8, for models with and without noise), the relative accuracy of forecasts from the MDS and VAR models (Table A.9), and coverage rates (Table A.10). Those results are fairly similar to what is reported in the paper and above for data including the pandemic sample. In addition, Table A.11 complements the results for the VAR model reported in the paper (for the sample ending in 2023Q4), with similar results from the MDS model.

The results also include additional figures testing the uniformity of the empirical CDFs of PITs for forecasts of CPI and PGDP inflation (Figures A.19 and A.20) and comparing cumulative log scores — i.e., marginal likelihoods — of the MDS and VAR specifications (Figure A.21). In addition, Figure A.22 reports end-of-sample estimates, obtained from recursive out-of-sample forecast simulations, of the MDS model's shifting endpoints, y_t^*, for each variable, that track closely observed long-run forecasts from the SPF (these are 10-years-ahead average forecasts available for GDP growth and CPI inflation). Figure A.23 reports similar estimates generated from the VAR model.

Table A.7: Predictability of SPF point forecasts (pre COVID)

	intercept								**slope**							
	RGDP		UNRATE		PGDP		CPI		RGDP		UNRATE		PGDP		CPI	
Forecast	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR
h = 0	−0.66	−0.49	−0.03	−0.08	0.04	0.13	−0.11	−0.11	1.17	1.16	1.00	1.02	0.94	0.92	1.03	1.08
	(0.43)	(0.39)	(0.07)	(0.08)	(0.11)	(0.11)	(0.44)	(0.43)	(0.16)	(0.15)	(0.01)	(0.01)	(0.05)	(0.05)	(0.18)	(0.18)
h = 1	−0.49	−0.15	−0.07	−0.01	−0.02	0.13	−0.01	0.09	1.14	1.03	1.01	1.00	0.98	0.95	0.97	0.96
	(0.35)	(0.33)	(0.09)	(0.09)	(0.08)	(0.08)	(0.16)	(0.14)	(0.12)	(0.12)	(0.02)	(0.02)	(0.03)	(0.04)	(0.07)	(0.06)
h = 2	−0.09	0.21	−0.10	−0.05	**0.13**	**0.26**	0.12	0.16	1.00	0.90	1.02	1.01	**0.92**	**0.84**	0.92	0.94
	(0.23)	(0.21)	(0.10)	(0.10)	(0.06)	(0.07)	(0.10)	(0.12)	(0.08)	(0.07)	(0.02)	(0.02)	(0.03)	(0.03)	(0.04)	(0.05)
h = 3	0.14	**0.31**	−0.13	−0.08	0.11	**0.22**	0.15	**0.21**	0.94	**0.86**	1.03	1.02	**0.93**	**0.86**	**0.91**	**0.91**
	(0.14)	(0.12)	(0.11)	(0.12)	(0.07)	(0.08)	(0.09)	(0.09)	(0.05)	(0.04)	(0.02)	(0.02)	(0.03)	(0.03)	(0.04)	(0.04)
h = 4	**0.29**	**0.94**	−0.18	−0.08	**0.16**	**0.22**	0.12	**0.23**	**0.89**	**0.68**	1.04	1.02	**0.92**	**0.91**	0.94	**0.89**
	(0.12)	(0.19)	(0.11)	(0.17)	(0.07)	(0.08)	(0.09)	(0.09)	(0.04)	(0.07)	(0.02)	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)
y = 1	−0.24	−0.15	−0.17	−0.12	**0.16**	**0.22**	0.01	0.10	1.06	1.04	1.03	1.03	**0.90**	**0.88**	0.98	0.95
	(0.30)	(0.28)	(0.09)	(0.11)	(0.07)	(0.07)	(0.12)	(0.12)	(0.10)	(0.10)	(0.02)	(0.02)	(0.03)	(0.03)	(0.05)	(0.05)
y = 2	−0.19	−0.19	**-0.27**	−0.08	—	—	0.64	**0.94**	1.07	1.07	1.03	1.00	—	—	0.73	**0.59**
	(0.13)	(0.15)	(0.11)	(0.12)			(0.36)	(0.23)	(0.05)	(0.06)	(0.02)	(0.02)			(0.16)	(0.10)
y = 3	0.02	−0.10	**0.52**	−0.06	—	—	—	—	0.99	1.04	**0.86**	0.99	—	—	—	—
	(0.19)	(0.20)	(0.16)	(0.20)					(0.07)	(0.08)	(0.03)	(0.04)				

Notes: Estimated slope coefficients of Mincer-Zarnowitz regressions for model-based predictions of next-quarter's published values for SPF forecasts at different forecast horizons. Heteroskedasticity-consistent standard errors in brackets. Bold font distinguishes coefficient estimates significantly different from 0 (intercept) or 1 (slope) with a 10% confidence level. Evaluation window from 1990Q1 to 2019Q4 (and as far as data for SPF forecasts at the different horizons is available).

Table A.8: Predictability of SPF point forecasts (pre COVID, noise-free model)

	intercept								**slope**							
	RGDP		UNRATE		PGDP		CPI		RGDP		UNRATE		PGDP		CPI	
Forecast	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR	MDS	VAR
h = 0	−0.66	−0.46	−0.03	−0.13	0.04	0.15	−0.11	−0.12	1.17	1.15	1.00	1.02	0.94	0.90	1.03	1.08
	(0.43)	(0.39)	(0.07)	(0.08)	(0.11)	(0.11)	(0.44)	(0.45)	(0.16)	(0.15)	(0.01)	(0.01)	(0.05)	(0.05)	(0.18)	(0.19)
h = 1	−0.49	−0.12	−0.07	−0.05	−0.02	0.09	−0.01	−0.01	1.14	1.02	1.01	1.01	0.98	0.94	0.97	0.99
	(0.35)	(0.32)	(0.09)	(0.09)	(0.08)	(0.09)	(0.16)	(0.14)	(0.12)	(0.11)	(0.02)	(0.02)	(0.03)	(0.04)	(0.07)	(0.06)
h = 2	−0.09	0.15	−0.10	0.00	**0.13**	0.12	0.12	0.17	1.00	0.93	1.02	1.00	**0.92**	1.01	0.92	0.92
	(0.23)	(0.22)	(0.10)	(0.13)	(0.06)	(0.10)	(0.10)	(0.10)	(0.08)	(0.08)	(0.02)	(0.02)	(0.03)	(0.05)	(0.04)	(0.04)
h = 3	0.14	**1.74**	−0.13	−0.16	0.11	**1.10**	0.15	**0.61**	0.94	**0.35**	1.03	1.04	**0.93**	**0.41**	**0.91**	**0.75**
	(0.14)	(0.15)	(0.11)	(0.24)	(0.07)	(0.07)	(0.09)	(0.18)	(0.05)	(0.06)	(0.02)	(0.05)	(0.03)	(0.02)	(0.04)	(0.07)
h = 4	**1.65**	**1.76**	−0.13	0.07	**1.13**	**0.18**	**0.48**	**0.49**	**0.39**	**0.38**	1.03	0.99	**0.52**	**0.90**	**0.80**	**0.82**
	(0.19)	(0.14)	(0.12)	(0.17)	(0.16)	(0.08)	(0.15)	(0.11)	(0.07)	(0.05)	(0.02)	(0.03)	(0.07)	(0.04)	(0.06)	(0.05)
y = 1	**0.42**	**0.51**	−0.16	−0.18	**0.17**	**0.42**	0.12	**0.38**	**0.83**	**0.81**	1.03	1.03	**0.92**	**0.78**	0.94	**0.85**
	(0.17)	(0.13)	(0.13)	(0.16)	(0.08)	(0.09)	(0.09)	(0.15)	(0.06)	(0.05)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.06)
y = 2	0.13	0.29	**-0.33**	−0.03	—	—	**0.98**	**1.20**	0.94	0.89	**1.05**	0.99	—	—	**0.57**	**0.48**
	(0.10)	(0.24)	(0.08)	(0.09)			(0.32)	(0.30)	(0.04)	(0.09)	(0.02)	(0.02)			(0.14)	(0.13)
y = 3	0.13	0.35	0.06	0.11	—	—	—	—	0.94	0.87	0.97	0.97	—	—	—	—
	(0.16)	(0.31)	(0.22)	(0.20)					(0.06)	(0.12)	(0.04)	(0.04)				

Notes: Estimated slope coefficients of Mincer-Zarnowitz regressions for model-based predictions of next-quarter's published values for SPF forecasts at different forecast horizons. Heteroskedasticity-consistent standard errors in brackets. Bold font distinguishes coefficient estimates significantly different from 0 (intercept) or 1 (slope) with a 10% confidence level. Evaluation window from 1990Q1 to 2019Q4 (and as far as data for SPF forecasts at the different horizons is available).

Table A.9: Relative Forecast Accuracy of MDS vs VAR models (pre COVID)

	RMSE				CRPS			
h	RGDP	UNRATE	PGDP	CPI	RGDP	UNRATE	PGDP	CPI
0	1.00	1.16	1.02	0.92*	1.00	0.96	1.01	0.94***
1	1.00	1.06	1.00	1.01	1.00	0.96	0.99	1.01
2	0.99	1.01	0.98	1.01	0.99	0.96	0.97	1.01
3	1.00	0.99	0.98	1.00	1.01	0.96	0.97	1.01
4	1.00	0.99	0.99	0.99	1.00	0.97	0.98	1.01
5	1.01	1.00	1.00	0.99	1.02	0.99	0.99	1.00
6	1.01	1.01	1.01	0.99	1.02	1.00	1.00	1.00
7	1.01	1.01	1.00	0.98	1.01	1.01	0.99	1.00
8	1.01	1.02	1.00	0.98	1.01	1.02	0.99	0.99
9	1.00	1.02	1.01	0.98	1.01	1.02	1.00	0.99
10	1.00	1.02	1.01	0.99	1.00	1.02	1.00	1.00
11	1.01	1.01	1.00	0.98	1.00	1.01	0.99	1.00
12	1.01	1.01	1.01	0.98	1.01	1.01	1.00	0.99
13	1.01	1.00	1.01	0.98	1.01	1.00	1.00	0.99
14	1.01	1.00	1.01	0.98	1.01	0.99	1.00	0.99
15	1.01	0.99	1.01	0.97	1.02	0.99	1.00	0.99
16	1.02	0.99	1.01	0.97	1.03*	0.98	0.99	0.99

Note: Relative RMSE and CRPS of VAR model (with MDS in denominator). Quarterly forecast horizons, h. Evaluation window from 1990Q1 through 2019Q4 (and as far as realized values are available). 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 A.10: Coverage rates (pre COVID)

	RGDP		UNRATE		PGDP		CPI	
h	68%	90%	68%	90%	68%	90%	68%	90%
PANEL A: MDS Model								
0	47.50***	77.50***	88.33***	98.33***	57.50**	87.50	66.67	90.83
1	52.94***	78.15***	84.03***	98.32***	62.18	88.24	60.50*	87.39
2	52.54***	80.51**	82.20***	96.61***	62.71	88.14	62.71	88.14
3	49.57***	81.20*	77.78	94.87	64.10	90.60	64.10	88.03
4	56.90**	84.48	71.55	93.10	61.21	89.66	70.69	89.66
5	58.26*	87.83	70.43	93.04	60.87	91.30	69.57	89.57
6	57.02**	86.84	71.05	93.86	61.40	90.35	70.18	89.47
7	59.29*	88.50	68.14	92.92	61.06	94.69**	69.91	91.15
8	62.50	86.61	62.50	91.96	64.29	95.54***	71.43	91.96
9	61.26	87.39	57.66	91.89	66.67	94.59**	73.87	91.89
10	61.82	87.27	56.36	90.91	65.45	98.18***	74.55	91.82
11	63.30	87.16	55.96	90.83	68.81	98.17***	75.23	92.66
12	68.52	90.74	58.33	89.81	71.30	97.22***	74.07	92.59
13	67.29	88.79	57.94	88.79	74.77	98.13***	76.64	93.46
14	66.04	90.57	58.49	88.68	78.30**	98.11***	77.36	93.40
15	67.62	89.52	56.19	86.67	79.05**	98.10***	77.14	94.29
16	69.23	91.35	58.65	87.50	77.88*	99.04***	78.85	94.23
PANEL B: VAR Model								
0	50.00***	81.67***	76.67*	96.67***	58.33**	87.50	66.67	92.50
1	56.30**	80.67***	79.83***	94.12	63.87	85.71	62.18	87.39
2	58.47*	83.90*	78.81**	95.76*	67.80	90.68	66.10	88.14
3	52.99***	82.91*	78.63*	94.02	67.52	91.45	67.52	90.60
4	57.76*	87.93	69.83	93.97	66.38	91.38	67.24	87.93
5	60.87	88.70	66.96	93.04	69.57	93.91	73.04	92.17
6	60.53	88.60	64.91	92.98	69.30	93.86	70.18	92.98
7	66.37	89.38	63.72	92.92	69.03	95.58**	70.80	92.92
8	64.29	88.39	60.71	91.07	71.43	94.64	74.11	91.96
9	64.86	87.39	56.76	90.99	72.97	92.79	72.07	92.79
10	65.45	89.09	56.36	89.09	71.82	96.36***	72.73	91.82
11	69.72	87.16	51.38**	87.16	76.15	97.25***	77.98	92.66
12	66.67	88.89	49.07**	86.11	75.93*	98.15***	75.93	93.52
13	72.90	88.79	47.66**	86.92	73.83	98.13***	76.64	92.52
14	70.75	91.51	48.11**	85.85	76.42	98.11***	76.42	93.40
15	68.57	91.43	46.67**	84.76	76.19	98.10***	76.19	94.29
16	71.15	91.35	46.15*	83.65	78.85*	99.04***	78.85	94.23

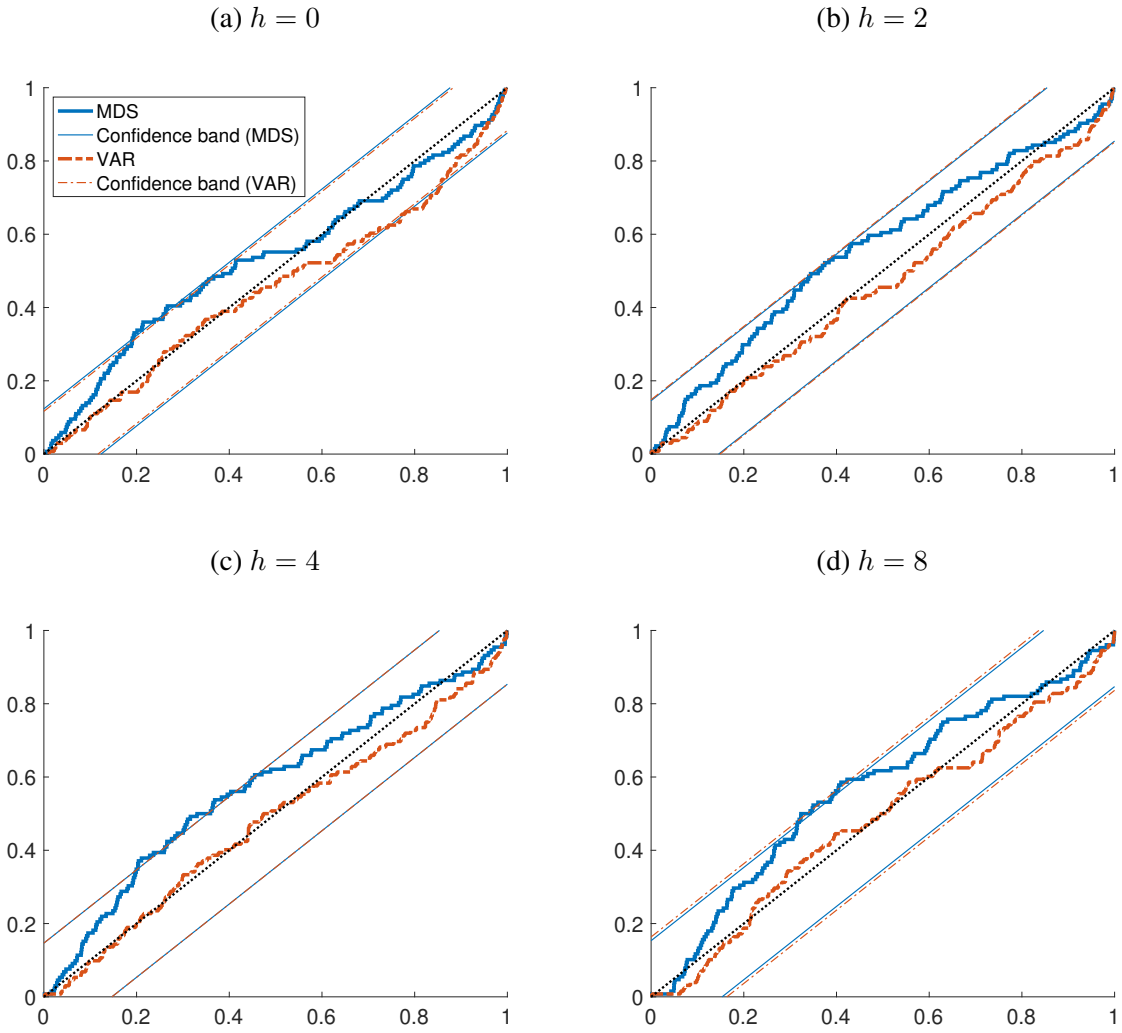
Note: Coverage rates for uncertainty bands with nominal levels of 68% and 90% for out-of-sample forecasts at quarterly forecast horizons, h. Evaluation window from 1990Q1 through 2019Q4 (and as far as realized values are available). Reflecting the availability of annual SPF forecasts, forecasts for inflation in CPI and GDP prices are evaluated only up to $h = 12$, and $h = 8$, respectively. 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 A.11: Coverage rates (full sample)

	RGDP		UNRATE		PGDP		CPI	
h	68%	90%	68%	90%	68%	90%	68%	90%
PANEL A: MDS Model								
0	48.53***	78.68***	86.76***	96.32***	55.15***	83.09*	66.18	88.97
1	52.59***	77.04***	82.22***	97.04***	60.00*	83.70	59.26**	84.44
2	52.99***	80.60***	79.85**	94.78*	60.45	83.58	60.45	84.33
3	50.38***	81.20*	75.94	92.48	61.65	85.71	61.65	84.21
4	56.82**	82.58	71.21	90.15	59.09**	84.85	67.42	85.61
5	58.02**	84.73	69.47	89.31	57.25**	86.26	65.65	84.73
6	56.15**	84.62	69.23	89.23	56.92*	85.38	65.38	85.38
7	57.36**	83.72	66.67	88.37	56.59**	89.15	65.89	86.82
8	60.16	82.03	61.72	88.28	60.16	90.62	66.41	86.72
9	58.27**	82.68	57.48	88.19	62.20	88.98	68.50	87.40
10	59.52*	82.54	56.35	87.30	60.32	92.06	69.05	87.30
11	60.80*	82.40	56.00	87.20	63.20	90.40	69.60	88.00
12	65.32	86.29	57.26	86.29	65.32	89.52	68.55	87.90
13	63.41	84.55	57.72	86.18	68.29	90.24	70.73	88.62
14	62.30	85.25	58.20	86.07	71.31	90.16	71.31	88.52
15	62.81	85.12	57.85	84.30	71.90	90.91	71.07	89.26
16	65.00	86.67	60.00	86.67	70.83	90.83	72.50	89.17
PANEL B: VAR Model								
0	50.00***	81.62***	76.47*	94.85**	55.88***	84.56*	66.18	91.18
1	55.56***	79.26***	77.78**	93.33	62.22	82.22**	60.74	84.44
2	58.21**	82.84**	77.61*	94.03	65.67	85.82	63.43	84.33
3	53.38***	82.71*	78.20*	91.73	64.66	86.47	64.66	86.47
4	56.06**	85.61	68.94	90.91	62.88	86.36	64.39	84.09
5	59.54	86.26	66.41	89.31	64.89	87.79	68.70	87.02
6	58.46**	86.15	63.85	88.46	64.62	87.69	66.15	88.46
7	64.34	86.05	62.02	88.37	64.34	89.92	65.89	87.60
8	62.50	84.38	60.16	87.50	65.62	88.28	68.75	86.72
9	62.20	83.46	56.69*	86.61	66.93	86.61	66.14	88.19
10	62.70	84.92	56.35*	85.71	65.87	89.68	67.46	87.30
11	66.40	83.20	52.00**	83.20	69.60	88.80	72.00	88.00
12	63.71	84.68	50.00**	82.26	69.35	89.52	70.16	87.90
13	69.11	84.55	48.78**	82.93	67.48	90.24	70.73	87.80
14	66.39	86.89	49.18**	82.79	69.67	89.34	69.67	88.52
15	64.46	86.78	48.76**	82.64	69.42	88.43	69.42	89.26
16	66.67	87.50	48.33*	81.67	71.67	90.00	71.67	89.17

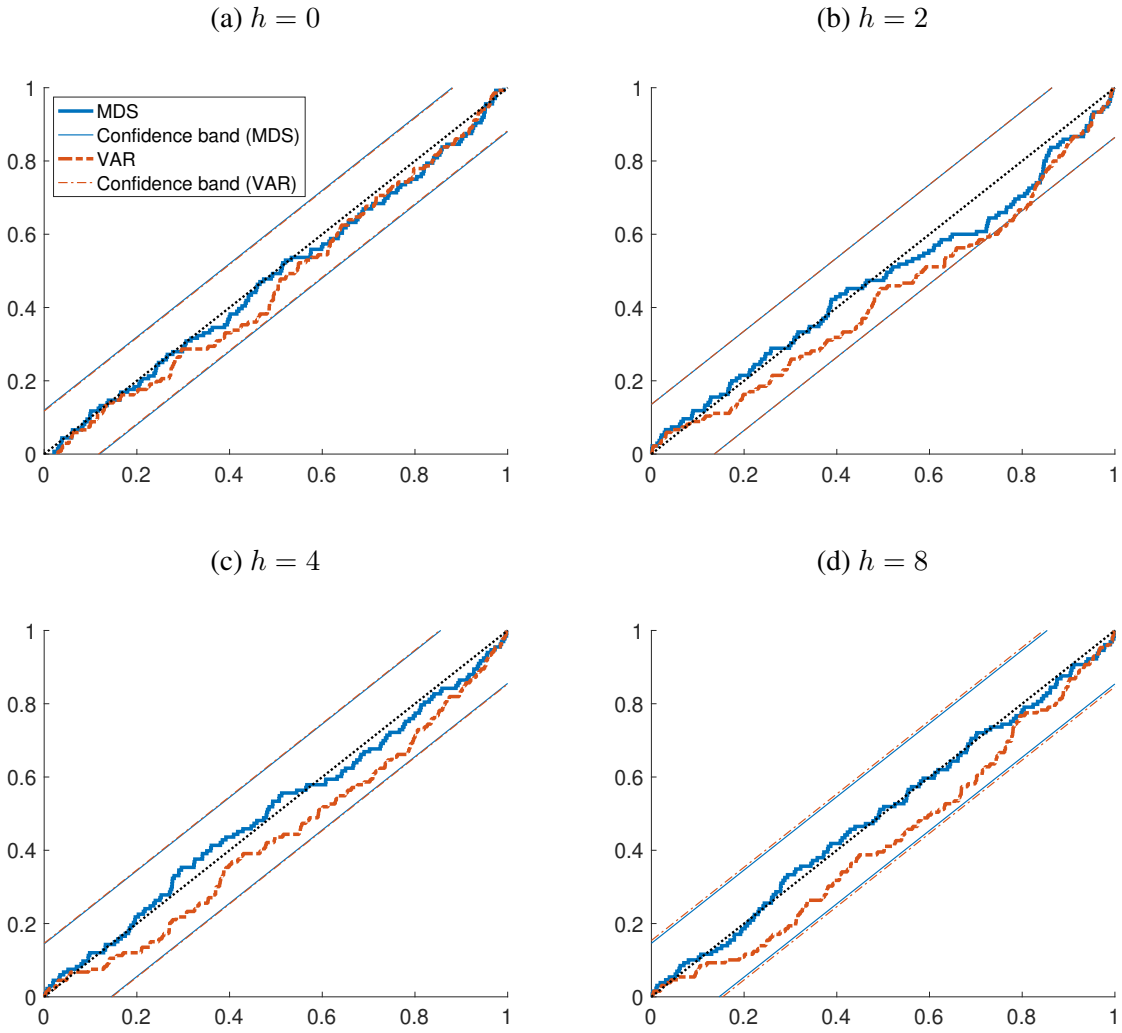
Note: Coverage rates for uncertainty bands with nominal levels of 68% and 90% for out-of-sample forecasts at quarterly forecast horizons, h. Evaluation window from 1990Q1 through 2023Q4 (and as far as realized values are available). Reflecting the availability of annual SPF forecasts, forecasts for inflation in CPI and GDP prices are evaluated only up to $h = 12$, and $h = 8$, respectively. 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 A.19: GDP price inflation PITs



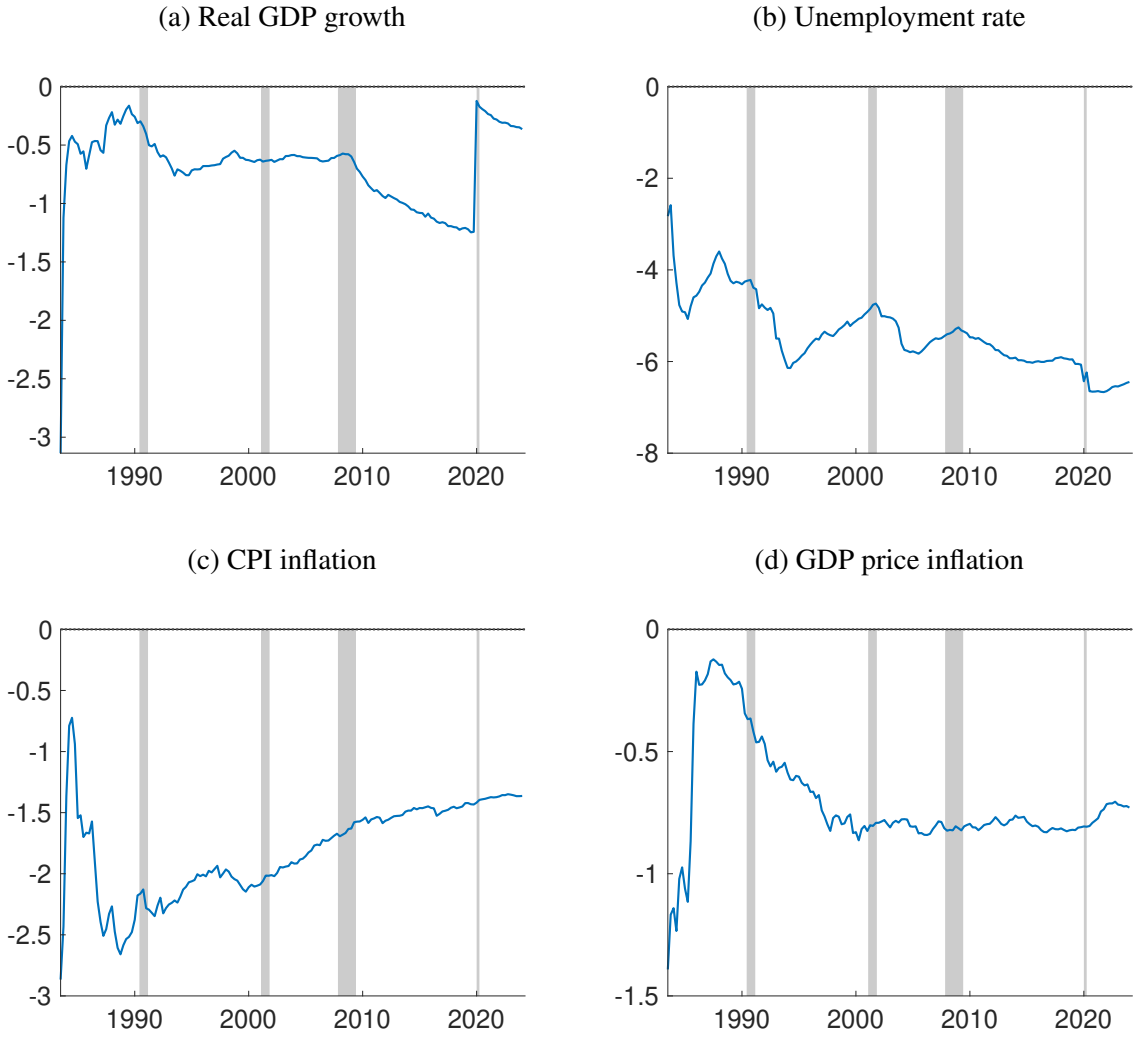
Notes: Empirical cumulative distributions of probability integral transforms (PITs) for inflation in the GDP price index at selected quarterly forecast horizons. All forecasts are generated out of sample by our MDS and VAR models, and evaluated over an evaluation window from 1990Q1 through 2023Q4 (and as far as realized values are available). 95% confidence bands for tests of correct calibration from Rossi and Sekhposyan (2019); computed separately for each model, but with nearly identical plot lines.

Figure A.20: CPI inflation PITs



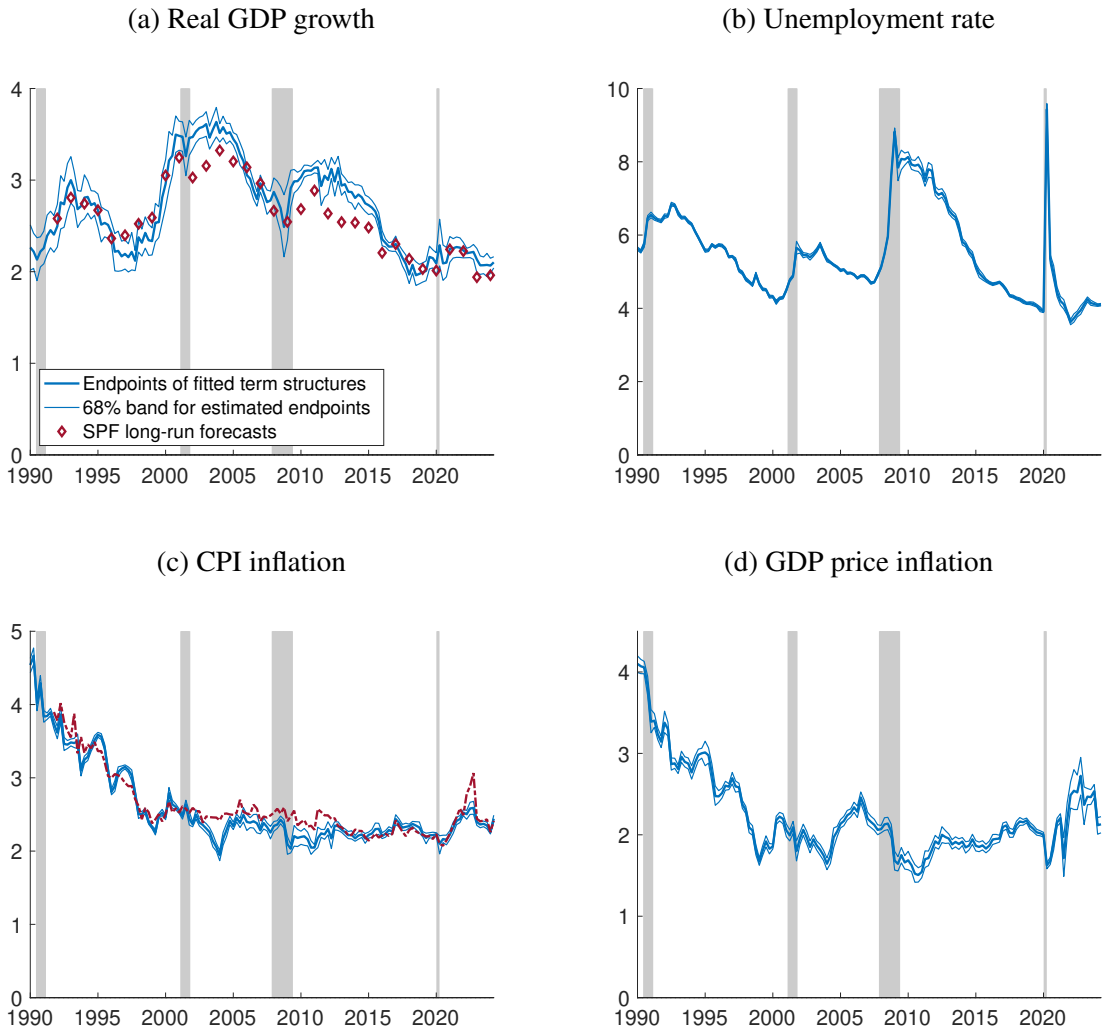
Notes: Empirical cumulative distributions of probability integral transforms (PITs) for CPI inflation at selected quarterly forecast horizons. All forecasts are generated out of sample by our MDS and VAR models, and evaluated over an evaluation window from 1990Q1 through 2023Q4 (and as far as realized values are available). 95% confidence bands for tests of correct calibration from Rossi and Sekhposyan (2019); computed separately for each model, but with nearly identical plot lines.

Figure A.21: Log scores



Notes: Figures show recursive means (across time) of the differences in 1-step-ahead log predictive scores for the MDS less the VAR model (negative entries mean the MDS model has the better score). These score differences are closely related to differences in log marginal likelihoods; the likelihoods equal sums of 1-step-ahead log predictive scores. Shaded areas depict NBER recessions.

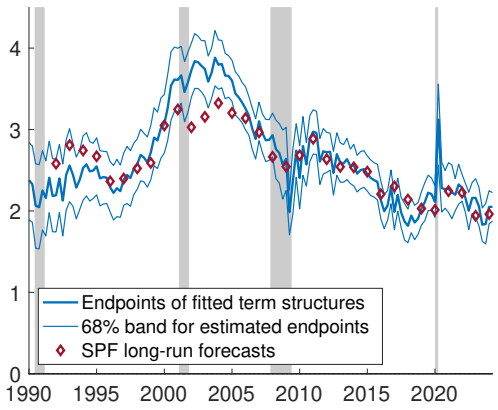
Figure A.22: Endpoint estimates (MDS)



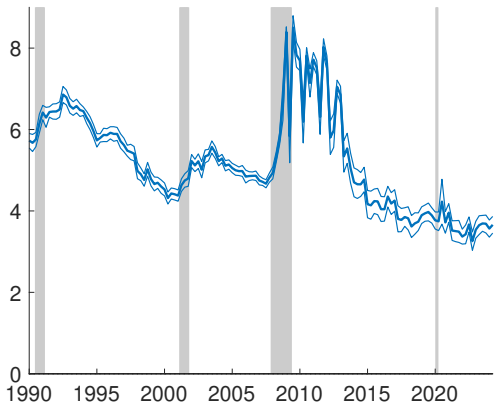
Notes: End-of-sample estimates of endpoints, y_t^*, for the MDS model. Estimates reflect posterior means and 68% bands, obtained from MCMC model estimates over growing samples (all using data since 1968Q4), as used in our out-of-sample forecast simulations. For GDP growth and CPI, corresponding long-run forecasts from the SPF (for 10-year ahead average growth) are shown as well.

Figure A.23: Endpoint estimates (VAR)

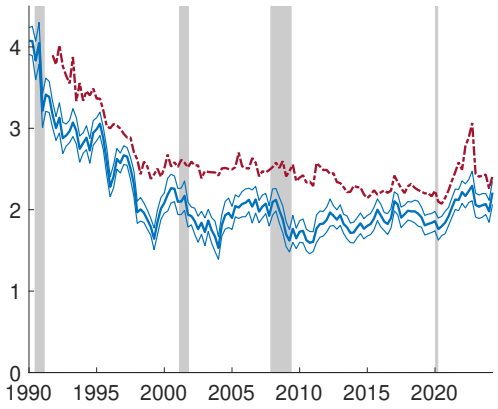
(a) Real GDP growth



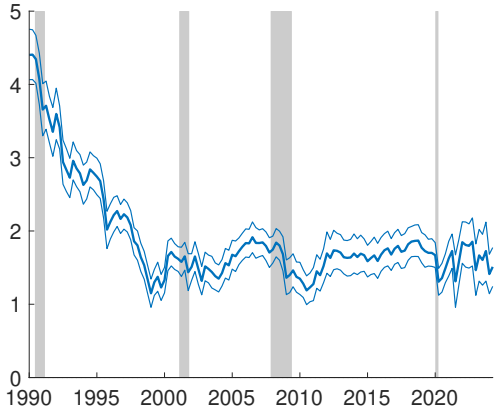
(b) Unemployment rate



(c) CPI inflation



(d) GDP price inflation



Notes: End-of-sample estimates of endpoints, y_t^*, for the VAR model. Estimates reflect posterior means and 68% bands, obtained from MCMC model estimates over growing samples (all using data since 1968Q4), as used in our out-of-sample forecast simulations. For GDP growth and CPI, corresponding long-run forecasts from the SPF (for 10-year ahead average growth) are shown as well.

References

Angeletos, George-Marios, Zhen Huo and Karthik A. Sastry. (2021). "Imperfect macroeconomic expectations: Evidence and theory". *NBER Macroeconomics Annual*, 35, pp. 1-86. https://doi.org/10.1086/712313

Aruoba, S. Boragan. (2020). "Term structures of inflation expectations and real interest rates". *Journal of Business & Economic Statistics*, 38(3), pp. 542-553. https://doi.org/10.1080/07350015.2018.1529599

Carriero, Andrea, Joshua C. C. Chan, Todd E. Clark and Massimiliano Marcellino. (2022a). "Corrigendum to: Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors". *Journal of Econometrics*, 227(2), pp. 506-512. https://doi.org/10.1016/j.jeconom.2021.11.010

Carriero, Andrea, Todd E. Clark and Massimiliano Marcellino. (2016). "Common drifting volatility in large Bayesian VARs". *Journal of Business & Economic Statistics*, 34(3), pp. 375-390. https://doi.org/10.1080/07350015.2015.1040116

Carriero, Andrea, Todd E. Clark and Massimiliano Marcellino. (2019). "Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors". *Journal of Econometrics*, 212(1), pp. 137-154. https://doi.org/10.1016/j.jeconom.2019.04.024

Carriero, Andrea, Todd E. Clark, Massimiliano Marcellino and Elmar Mertens. (2022b). "Addressing COVID-19 outliers in BVARs with stochastic volatility". *Review of Economics and Statistics*, forthcoming. https://doi.org/10.1162/rest_a_01213

Carvalho, Carlos M., Nicholas G. Polson and James G. Scott. (2010). "The horseshoe estimator for sparse signals". *Biometrika*, 97(2), pp. 465-480. https://doi.org/10.1093/biomet/asq017

Chan, Joshua C. C. (2020). "Large Bayesian VARs: A flexible Kronecker error covariance structure". *Journal of Business & Economic Statistics*, 38(1), pp. 68-79. https://doi.org/10.1080/07350015.2018.1451336

Chan, Joshua C. C., and Ivan Jeliazkov. (2009). "Efficient simulation and integrated likelihood estimation in state space models". *International Journal of Mathematical Modelling and Numerical Optimization*, 1(1/2), pp. 101-120. https://doi.org/10.1504/ijmmno.2009.030090

Chib, Siddhartha, and Ivan Jeliazkov. (2006). "Inference in semiparametric dynamic models for binary longitudinal data". *Journal of the American Statistical Association*, 101(474), pp. 685-700. https://doi.org/10.1198/016214505000000871

Clark, Todd E., Michael W. McCracken and Elmar Mertens. (2020). "Modeling time-varying uncertainty of multiple-horizon forecast errors". *The Review of Economics and Statistics*, 102(1), pp. 17-33. https://doi.org/10.1162/rest_a_00809

Clark, Todd E., and Francesco Ravazzolo. (2015). "Macroeconomic forecasting performance under alternative specifications of time-varying volatility". *Journal of Applied Econometrics*, 30(4), pp. 551-575. https://doi.org/10.1002/jae.2379

Coibion, Olivier, and Yuriy Gorodnichenko. (2015). "Information rigidity and the expectations formation process: A simple framework and new facts". *American Economic Review*, 105(8), pp. 2644-2678. https://doi.org/10.1257/aer.20110306

Crump, Richard K., Stefano Eusepi, Emanuel Moench and Bruce Preston. (2023). "The term structure of expectations". In Bachmann, Rüdiger, Giorgio Topa and Wilbert van der Klaauw (eds.), *Handbook of Economic Expectations*, chap. 17. Academic Press, pp. 507-540. https://doi.org/10.1016/B978-0-12-822927-9.00025-2

Del Negro, Marco, and Giorgio E. Primiceri. (2015). "Time varying structural vector autoregressions and monetary policy: A corrigendum". *Review of Economic Studies*, 82(4), pp. 1342-1345. https://doi.org/10.1093/restud/rdv024

Durbin, J., and S.J. Koopman. (2002). "A simple and efficient simulation smoother for state space time series analysis". *Biometrika*, 89(3), pp. 603-615. https://doi.org/10.1093/biomet/89.3.603

Hayashi, Fumio. (2000). Econometrics. Princeton University Press. 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(1), pp. 185-212. https://doi.org/10.1016/j.jeconom.2003.09.001

Kadiyala, K. Rao, and Sune Karlsson. (1997). "Numerical methods for estimation and inference in Bayesian VAR-models". *Journal of Applied Econometrics*, 12(2), pp. 99-132. https://doi.org/10.1002/(sici)1099-1255(199703)12:2<99::aid-jae429>3.3.co;2-1

Kim, Sangjoon, Neil Shephard and Siddhartha Chib. (1998). "Stochastic volatility: Likelihood inference and comparison with ARCH models". *The Review of Economic Studies*, 65(3), pp. 361-393. https://doi.org/10.1111/1467-937X.00050

Makalic, Enes, and Daniel F. Schmidt. (2016). "A simple sampler for the horseshoe estimator". *IEEE Signal Processing Letters*, 23(1), pp. 179-182. https://doi.org/10.1109/LSP.2015.2503725

Mariano, Roberto S., and Yasutomo Murasawa. (2003). "A new coincident index of business cycles based on monthly and quarterly series". *Journal of Applied Econometrics*, 18(4), pp. 427-443. https://doi.org/10.1002/jae.695

Mertens, Elmar. (2023). "Precision-based sampling for state space models that have no measurement error". *Journal of Economic Dynamics and Control*, 154(104720). https://doi.org/10.1016/j.jedc.2023.104720

Omori, Yasuhiro, Siddhartha Chib, Neil Shephard and Jouchi Nakajima. (2007). "Stochastic volatility with leverage: Fast and efficient likelihood inference". *Journal of Econometrics*, 140(2), pp. 425-449. https://doi.org/10.1016/j.jeconom.2006.07.008

Patton, Andrew J., and Allan Timmermann. (2011). "Predictability of output growth and inflation: A multi-horizon survey approach". *Journal of Business & Economic Statistics*, 29(3), pp. 397-410. https://doi.org/10.1198/jbes.2010.08347

Prüser, Jan. (2021). "The horseshoe prior for time-varying parameter VARs and monetary policy". *Journal of Economic Dynamics and Control*, 129(C). https://doi.org/10.1016/j.jedc.2021.104188

Rossi, Barbara, and Tatevik Sekhposyan. (2019). "Alternative tests for correct specification of conditional predictive densities". *Journal of Econometrics*, 208(2), pp. 638-657. https://doi.org/10.1016/j.jeconom.2018.07.008

BANCO DE ESPAÑA PUBLICATIONS

WORKING PAPERS

2310 ANDRÉS ALONSO-ROBISCO, JOSÉ MANUEL CARBÓ and JOSÉ MANUEL MARQUÉS: Machine Learning methods in climate finance: a systematic review.

2311 ALESSANDRO PERI, OMAR RACHEDI and IACOPO VAROTTO: The public investment multiplier in a production network.

2312 JUAN S. MORA-SANGUINETTI, JAVIER QUINTANA, ISABEL SOLER and ROK SPRUK: Sector-level economic effects of regulatory complexity: evidence from Spain.

2313 CORINNA GHIRELLI, ENKELEJDA HAVARI, ELENA MERONI and STEFANO VERZILLO: The long-term causal effects of winning an ERC grant.

2314 ALFREDO GARCÍA-HIERNAUX, MARÍA T. GONZÁLEZ-PÉREZ and DAVID E. GUERRERO: How to measure inflation volatility. A note.

2315 NICOLÁS ABBATE, INÉS BERNIELL, JOAQUÍN COLEFF, LUIS LAGUINGE, MARGARITA MACHELETT, MARIANA MARCHIONNI, JULIÁN PEDRAZZI and MARÍA FLORENCIA PINTO: Discrimination against gay and transgender people in Latin America: a correspondence study in the rental housing market.

2316 SALOMÓN GARCÍA: The amplification effects of adverse selection in mortgage credit suply.

2317 METTE EJRNÆS, ESTEBAN GARCÍA-MIRALLES, METTE GØRTZ and PETTER LUNDBORG: When death was postponed: the effect of HIV medication on work, savings and marriage.

2318 GABRIEL JIMÉNEZ, LUC LAEVEN, DAVID MARTÍNEZ-MIERA and JOSÉ-LUIS PEYDRÓ: Public guarantees and private banks' incentives: evidence from the COVID-19 crisis.

2319 HERVÉ LE BIHAN, DANILO LEIVA-LEÓN and MATÍAS PACCE: Underlying inflation and asymmetric risks.

2320 JUAN S. MORA-SANGUINETTI, LAURA HOSPIDO and ANDRÉS ATIENZA-MAESO: The numbers of equality regulation. Quantifying regulatory activity on non-discrimination and its relationship with gender gaps in the labour market.

2321 ANDRES ALONSO-ROBISCO and JOSÉ MANUEL CARBÓ: Analysis of CBDC Narrative of Central Banks using Large Language Models.

2322 STEFANIA ALBANESI, ANTÓNIO DIAS DA SILVA, JUAN F. JIMENO, ANA LAMO and ALENA WABITSCH: New technologies and jobs in Europe.

2323 JOSÉ E. GUTIÉRREZ: Optimal regulation of credit lines.

2324 MERCEDES DE LUIS, EMILIO RODRÍGUEZ and DIEGO TORRES: Machine learning applied to active fixed-income portfolio management: a Lasso logit approach.

2325 SELVA BAHAR BAZIKI, MARÍA J. NIETO and RIMA TURK-ARISS: Sovereign portfolio composition and bank risk: the case of European banks.

2326 ANGEL-IVAN MORENO and TERESA CAMINERO: Assessing the data challenges of climate-related disclosures in european banks. A text mining study.

2327 JULIO GÁLVEZ: Household portfolio choices under (non-)linear income risk: an empirical framework.

2328 NATASCHA HINTERLANG: Effects of Carbon Pricing in Germany and Spain: An Assessment with EMuSe.

2329 RODOLFO CAMPOS, SAMUEL PIENKNAGURA and JACOPO TIMINI: How far has globalization gone? A tale of two regions.

2330 NICOLÁS FORTEZA and SANDRA GARCÍA-URIBE: A Score Function to Prioritize Editing in Household Survey Data: A Machine Learning Approach.

2331 PATRICK MACNAMARA, MYROSLAV PIDKUYKO and RAFFAELE ROSSI: Taxing consumption in unequal economies.

2332 ESTHER CÁCERES and MATÍAS LAMAS: Dividend Restrictions and Search for Income.

2333 MARGARITA MACHELETT: Gender price gaps and competition: Evidence from a correspondence study.

2334 ANTON NAKOV and CARLOS THOMAS: Climate-conscious monetary policy.

2335 RICARDO BARAHONA, STEFANO CASSELLA and KRISTY A. E. JANSEN: Do teams alleviate or exacerbate the extrapolation bias in the stock market?

2336 JUAN S. MORA-SANGUINETTI and ANDRÉS ATIENZA-MAESO: "Green regulation": A quantification of regulations related to renewable energy, sustainable transport, pollution and energy efficiency between 2000 and 2022.

2401 LAURA HOSPIDO, NAGORE IRIBERRI and MARGARITA MACHELETT: Gender gaps in financial literacy: a multi-arm RCT to break the response bias in surveys.

2402 RUBÉN DOMÍNGUEZ-DÍAZ, SAMUEL HURTADO and CAROLINA MENÉNDEZ: The medium-term effects of investment stimulus.

2403 CLODOMIRO FERREIRA, JOSÉ MIGUEL LEIVA, GALO NUÑO, ÁLVARO ORTIZ, TOMASA RODRIGO and SIRENIA VAZQUEZ: The heterogeneous impact of inflation on households' balance sheets.

2404 JORGE ABAD, GALO NUÑO and CARLOS THOMAS: CBDC and the operational framework of monetary policy.
2405 STÉPHANE BONHOMME and ANGELA DENIS: Estimating individual responses when tomorrow matters.
2406 LAURA ÁLVAREZ-ROMÁN, SERGIO MAYORDOMO, CARLES VERGARA-ALERT and XAVIER VIVES: Climate risk, soft information and credit supply.
2407 JESÚS FERNÁNDEZ-VILLAVERDE, JOËL MARBET, GALO NUÑO and OMAR RACHEDI: Inequality and the zero lower bound.
2408 PABLO BURRIEL, MAR DELGADO-TÉLLEZ, CAMILA FIGUEROA, IVÁN KATARYNIUK and JAVIER J. PÉREZ: Estimating the contribution of macroeconomic factors to sovereign bond spreads in the euro area.
2409 LUIS E. ROJAS and DOMINIK THALER: The bright side of the doom loop: banks' sovereign exposure and default incentives.
2410 SALOMÓN GARCÍA-VILLEGAS and ENRIC MARTORELL: Climate transition risk and the role of bank capital requirements.
2411 MIKEL BEDAYO and JORGE E. GALÁN: The impact of the Countercyclical Capital Buffer on credit: Evidence from its accumulation and release before and during COVID-19.
2412 EFFROSYNI ADAMOPOULOU, LUIS DÍEZ-CATALÁN and ERNESTO VILLANUEVA: Staggered contracts and unemployment during recessions.
2413 LUIS FÉRNANDEZ LAFUERZA and JORGE E. GALÁN: Should macroprudential policy target corporate lending? Evidence from credit standards and defaults.
2414 STÉPHANE BONHOMME and ANGELA DENIS: Estimating heterogeneous effects: applications to labor economics.
2415 LUIS GUIROLA, LAURA HOSPIDO and ANDREA WEBER: Family and career: An analysis across Europe and North America.
2416 GERALD P. DWYER, BILJANA GILEVSKA, MARÍA J. NIETO and MARGARITA SAMARTÍN: The effects of the ECB's unconventional monetary policies from 2011 to 2018 on banking assets.
2417 NICOLÁS FORTEZA, ELVIRA PRADES and MARC ROCA: Analysing the VAT cut pass-through in Spain using web-scraped supermarket data and machine learning.
2418 JOSÉ-ELÍAS GALLEGOS: HANK beyond FIRE: Amplification, forward guidance, and belief shocks.
2419 DANIEL ALONSO: Stabilisation properties of a SURE-like European unemployment insurance.
2420 FRANCISCO GONZÁLEZ, JOSÉ E. GUTIÉRREZ and JOSÉ MARÍA SERENA: Shadow seniority? Lending relationships and borrowers' selective default.
2421 ROBERTO BLANCO, MIGUEL GARCÍA-POSADA, SERGIO MAYORDOMO and MARÍA RODRÍGUEZ-MORENO: Access to credit and firm survival during a crisis: the case of zero-bank-debt firms.
2422 FERNANDO CEREZO, PABLO GIRÓN, MARÍA T. GONZÁLEZ-PÉREZ and ROBERTO PASCUAL: The impact of sovereign debt purchase programmes. A case study: the Spanish-to-Portuguese bond yield spread.
2423 EDGAR SILGADO-GÓMEZ: Sovereign uncertainty.
2424 CLODOMIRO FERREIRA, JULIO GÁLVEZ and MYROSLAV PIDKUYKO: Housing tenure, consumption and household debt: life-cycle dynamics during a housing bust in Spain.
2425 RUBÉN DOMÍNGUEZ-DÍAZ and SAMUEL HURTADO: Green energy transition and vulnerability to external shocks.
2426 JOSEP GISBERT and JOSÉ E. GUTIÉRREZ: Bridging the gap? Fintech and financial inclusion.
2427 RODOLFO G. CAMPOS, MARIO LARCH, JACOPO TIMINI, ELENA VIDAL and YOTO V. YOTOV: Does the WTO Promote Trade? A Meta-analysis.
2428 SONER BASKAYA, JOSÉ E. GUTIÉRREZ, JOSÉ MARÍA SERENA and SERAFEIM TSOUKAS: Bank supervision and non-performing loan cleansing.
2429 TODD E. CLARK, GERGELY GANICS, and ELMAR MERTENS: Constructing fan charts from the ragged edge of SPF forecasts.