

Hauber, Philipp

Working Paper

How useful is external information from professional forecasters? Conditional forecasts in large factor models

Suggested Citation: Hauber, Philipp (2021) : How useful is external information from professional forecasters? Conditional forecasts in large factor models, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/251469>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

How useful is external information from professional forecasters? Conditional forecasts in large factor models

Philipp Hauber

Julius-Maximilians-Universität Würzburg, Kiel Institute for the World Economy

Abstract

This paper evaluates forecasts from a factor model estimated with a large real-time dataset of the German economy. The evaluation focuses on a broad cross-section of variables such as activity series including components of the gross domestic product and gross value added, deflators and other price measures as well as several labor market indicators. In addition to unconditional forecasts for these variables, we also investigate to what extent the forecast accuracy improves when we condition on professional forecasters' view on GDP growth and CPI inflation. We find that over the period from 2006 to 2017 the model's unconditional forecasts are broadly in line with autoregressive benchmarks for the majority of the 37 series that we focus on in the evaluation, in some cases performing somewhat better and in others somewhat worse. For a few variables capturing real activity and some price indicators, however, we find large gains in predictive accuracy that persist for forecast horizons of up to two quarters ahead. Conditioning on external information tends to improve the forecast accuracy in some instances but typically only for those series where the unconditional forecasts are already quite accurate. For around a third of the variables under consideration, the differences in forecast accuracy between conditional and unconditional forecasts are statistically significant for density forecasts; for point forecasts on the other hand we find no significant differences. From a methodological point of view, this paper proposes precision-based sampling algorithms to draw from the predictive density - unconditional or conditional on a subset of the system variables - in factor models and other models with unobserved components. Simulations show that these algorithms perform favorably compared to Kalman filter-based alternatives typically used in the literature.

Keywords: factor models, conditional forecasting, precision-based sampling

JEL classification: C11, C53, C55, E37.

Contact details: Julius-Maximilians-Universität Würzburg, Chair of Monetary Economics and International Financial Markets, Sanderring 2, 97070 Würzburg. Email: philipp.hauber@uni-wuerzburg.de. I would like to thank Martin Ademmer for helpful suggestions and Simone Knief for providing access to Kiel University's high-performance computing facilities.

1 Introduction

Factor models feature prominently in macroeconomic now- and forecasting applications. Typically, however, forecast evaluations in this literature focus only on a small subset of the large number of time series that are used to estimate these models. For example, Stock and Watson (2002) evaluate the forecast performance of industrial production, personal income, manufacturing and trade sales and (nonagricultural) employment in the United States. The factor-model based nowcasting literature has largely focused on predicting the real gross domestic product (GDP) (see Giannone et al., 2008; Bok et al., 2018; Schumacher and Breitung, 2008; Kuzin et al., 2013). While arguably the single most important measure of economic activity, analysts or policy-makers are typically also interested in forecasts for many other variables, e.g. concerning the labor market, price measures including deflators, components of GDP or activity measures like industrial production or turnover. Moreover, forecasts of these variables given a path of future real GDP growth and CPI inflation may be used in scenario analysis or as consistency checks to other model-based forecasts. Such conditional forecasts have received little attention in the factor model literature.

This paper therefore extends the literature on factor models by broadening the cross-sectional dimension of the forecast evaluation to include series that are not commonly featured. We consider a total of 37 time series comprising labor market indicators, price measures such as the CPI or PPI as well components of gross domestic product such as consumption, investment and exports as well gross value added of different sectors like construction or manufacturing. In addition to an evaluation of unconditional forecasts for such variables, we also investigate how the predictive accuracy changes when external information in the form of professional forecasters' views on activity and inflation is incorporated. Such an evaluation captures both how well the model describes the joint dynamics of the data as well as the accuracy of the conditioning information and thus gives a complete assessment of the forecast accuracy for the variables under consideration in this study.

Specifically, in the empirical application we rely on external information from the Reuters Poll of professional forecasters which collects the forecasts from private sector institutions and research institutes on a quarterly basis. We supplement these historical forecasts with a large real-time dataset consisting of roughly 80 time series that allows us to exactly replicate the information available at the time the forecast survey was conducted. We then estimate factor models with Bayesian methods to obtain forecasts both unconditional as well as conditional on the GDP growth and CPI inflation forecasts from the Reuters Poll. This allows us to assess how predictive accuracy improves both in terms of point and density forecast accuracy when external information is included.

As such, we also contribute to the literature on conditional forecasting which has mostly focussed on Bayesian vectorautoregressions - another class of popular macroeconometric methods. Banbura et al. (2015) consider conditional forecasts by conditioning on future values they shed light on the stability of dynamic relationships in the Euro area. Similarly, in their empirical application Clark and McCracken (2017) condition on realizations when

testing for bias, efficiency and MSE accuracy of conditional versus unconditional forecasts.

Closer to the questions guiding our research are Krüger et al. (2017), Tallman and Zaman (2020) and Ganics and Odendahl (2021) who combine model-based density forecasts with external information in the form of forecast surveys. Using entropic tilting, these authors tend to find gains in predictive accuracy even for those variables that are not directly tilted to match the survey forecasts. However, common to both papers is the focus on only a few time series within a Bayesian vector autoregression context.¹ Our analysis differs in that we work with factor models rather than vector autoregressions, condition on quarterly point forecasts for GDP growth and CPI inflation and consider the impact on a much broader and more disaggregated range of variables.

Besides the empirical contribution, this paper also proposes precision-based sampling algorithms for conditional forecasting and scenario analysis. Precision-based sampling algorithms (Chan and Jeliazkov, 2009; McCausland, 2012) that build on the seminal work by Rue (2001) on Gaussian Markov random fields, are increasingly used in macroeconomic applications. Hauber and Schumacher (2021) generalize these precision-samplers to applications with missing data. Since future values of variables can simply be considered as missing, we show how draws from the predictive density - unconditional or conditional on future values of a subset of the system's variables - can be obtained. In a simulation study we compare the performance of these algorithms to the Kalman-filter based implementations proposed by Banbura et al. (2015).

The results of the forecast evaluation show that for the majority of time series under consideration the factor model produces forecasts that are as accurate as an autoregressive benchmarks, with relative root mean squared forecast errors (RMSFE) and continuous ranked probability scores (CRPS) ranging from 0.9 to 1.1. For a few variables we find much larger gains compared to the benchmark and these persist as the forecast horizon increases. When we incorporate external information in the form of GDP growth and CPI inflation forecasts from a survey of professional forecasters, we generally find some improvements in forecast accuracy for the variables of interest from categories such as real activity, prices and the labor market. Only for a smaller number of series, however, do we find that conditioning produces large gains and these are typically the series for which the model already produces accurate unconditional forecasts. While the point and density forecast accuracy is broadly comparable in the sense that relative gains compared to an autoregressive benchmark for a given variable are similar, we find that when evaluating the entire predictive distribution the loss from the conditional forecasts is significantly lower than that of the model's unconditional predictions. For point forecasts, there is virtually no such evidence.

The remainder of this paper is structured as follows: Section 2 illustrates the precision-based sampling algorithms for conditional forecasting, while in Section 3 we discuss the dataset including the forecast survey used to condition the factor model forecasts as well as the evaluation set-up. Section 4 discusses the results and several robustness checks while

¹The latter two also consider medium-sized VAR with 10 and 14 variables, respectively as a robustness check and find similar results to their smaller baseline models.

Section 5 concludes.

2 Precision-based sampling algorithms for conditional forecasting

In this section, we lay out the precision-based sampling algorithms to obtain conditional forecasts. To fix notation and ideas, we first review the general precision-based sampler in Section 2.1 before turning to forecasting applications (Section 2.2). Lastly, we also discuss how repeated draws from the predictive density given parameters can be obtained (Section 2.3).

2.1 Precision-based sampling

We outline the precision-based algorithms through the lens of a dynamic factor model but the analysis can be generalized to other state space models with unobserved components. Factor models feature prominently in the macroeconomic literature. For an overview of their methodological developments and applications, see Stock and Watson (2016). Precision-based sampling applications include Chan and Jeliazkov (2009); McCausland (2015); Kaufmann and Schumacher (2017, 2019).

A *dynamic factor model* is defined as

$$\mathbf{y}_t = \sum_{k=0}^K \lambda_k \mathbf{f}_{t-k} + \mathbf{e}_t \quad (1a)$$

$$\mathbf{f}_t = \sum_{p=1}^P \phi_p \mathbf{f}_{t-p} + \mathbf{v}_t \quad (1b)$$

$$\mathbf{e}_t = \sum_{j=0}^J \psi_j \mathbf{e}_{t-j} + \boldsymbol{\epsilon}_t \quad (1c)$$

where \mathbf{f}_t denotes an $R \times 1$ vector of unobserved factors which summarizes the co-movement of the observables \mathbf{y}_t . The dynamics of the factors are modelled as a vector autoregression of order P . The idiosyncratic components, \mathbf{e}_t are modeled as independent autoregressive processes of order J , i.e. ψ_1, \dots, ψ_J are diagonal matrices. The $N \times R$ loadings matrices λ_k capture the dynamic relationships between observables and factors. Setting $K = 0$, so that the variables only load contemporaneously on the factors, yields a *static factor model*. The innovations $\boldsymbol{\epsilon}_t, \mathbf{v}_t$ are Normal, uncorrelated at all leads and lags and their covariance matrices given by Ω and Σ , respectively.

Stacking the observables, factors and idiosyncratic components over t yields

$$\mathbf{y} = \mathbf{M}\mathbf{f} + \mathbf{e} \quad (2)$$

where

$$\mathbf{M} = \begin{bmatrix} \lambda_0 & & & & \\ \lambda_1 & \lambda_0 & & & \\ \lambda_2 & \lambda_1 & \lambda_0 & & \\ & \ddots & \ddots & \ddots & \\ & & \lambda_K & \dots & \lambda_1 & \lambda_0 \\ & & & \ddots & \ddots & \ddots & \ddots \\ & & & & \lambda_K & \dots & \lambda_1 & \lambda_0 \end{bmatrix}.$$

The vector of idiosyncratic components can be written as $\mathbf{H}_e \mathbf{e} = \varepsilon$ with

$$\mathbf{H}_e = \begin{bmatrix} I_N & & & & \\ -\psi_1 & I_N & & & \\ -\psi_2 & -\psi_1 & I_N & & \\ & & & & \\ -\psi_J & \dots & \dots & -\psi_1 & I_N \\ & & & & \\ & & & -\psi_J & \dots & \dots & -\psi_1 & I_N \end{bmatrix}$$

and $\varepsilon \sim \mathcal{N}(0_{NT}, I_T \otimes \Omega)$. It follows that \mathbf{e} is Normal, with mean 0 and covariance matrix $\mathbf{V}_e = (\mathbf{H}_e^\top (I_T \otimes \Omega^{-1}) \mathbf{H}_e)^{-1}$.

In a similar fashion, the vector of factors is distributed as

$$\mathbf{f} \sim \mathcal{N}(0, (\mathbf{H}_f^\top (I_T \otimes \Sigma^{-1}) \mathbf{H}_f)^{-1})$$

with

$$\mathbf{H}_f = \begin{bmatrix} I_N & & & & \\ -\phi_1 & I_N & & & \\ -\phi_2 & -\phi_1 & I_N & & \\ & & & & \\ -\phi_P & \dots & \dots & -\phi_1 & I_N \\ & & & & \\ & & & -\phi_P & \dots & \dots & -\phi_1 & I_N \end{bmatrix}.$$

Define

$$\mathbf{z} := \begin{bmatrix} \mathbf{f} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} I_{TR} & 0 \\ \mathbf{M} & I_{TN} \end{bmatrix} \begin{bmatrix} \mathbf{f} \\ \mathbf{e} \end{bmatrix}. \quad (3)$$

It follows that \mathbf{z} is Normal with mean 0 and covariance matrix \mathbf{V}_z . The corresponding precision matrix is the inverse of the covariance matrix:

$$\begin{aligned}
\mathbf{Q}_z &= \mathbf{V}_z^{-1} \\
&= \text{Var} \left(\begin{bmatrix} \mathbf{f} \\ \mathbf{y} \end{bmatrix} \right)^{-1} \\
&= \left(\begin{bmatrix} I & 0 \\ \mathbf{M} & I \end{bmatrix} \begin{bmatrix} \mathbf{H}_f^{-1} \mathbf{V}_v \mathbf{H}_f^{-1\top} & 0 \\ 0 & \mathbf{V}_e \end{bmatrix} \begin{bmatrix} I & \mathbf{M}^\top \\ 0 & I \end{bmatrix} \right)^{-1} \\
&= \begin{bmatrix} I & -\mathbf{M}^\top \\ 0 & I \end{bmatrix} \begin{bmatrix} \mathbf{H}_f^\top \mathbf{V}_v \mathbf{H}_f & 0 \\ 0 & \mathbf{V}_e \end{bmatrix} \begin{bmatrix} I & 0 \\ -\mathbf{M} & I \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{H}_f^\top \mathbf{V}_v \mathbf{H}_f + \mathbf{M}^\top \mathbf{V}_e \mathbf{M} & -\mathbf{M}^\top \mathbf{V}_e \\ -\mathbf{V}_e \mathbf{M} & \mathbf{V}_e \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{Q}_f & \mathbf{Q}_{fy} \\ \mathbf{Q}_{fy}^\top & \mathbf{Q}_y \end{bmatrix}
\end{aligned}$$

Bayesian estimation of Equations (1a)–(1c) requires a draw from the distribution of the factors conditional on the observations and parameters Θ , $p(\mathbf{f}_1, \dots, \mathbf{f}_T | \mathbf{y}_1, \dots, \mathbf{y}_T, \Theta)$. "Simulation smoothing" methods based on the Kalman filter/smoothing have been proposed by Carter and Kohn (1994) and Durbin and Koopman (2001) and are widely used in the literature. Precision-based sampling algorithms (Chan and Jeliazkov, 2009; McCausland, 2012) provide an alternative to these methods.

Using standard results for multivariate Normal distributions (e.g. Bishop, 2006, pp. 86–87), we can derive the desired conditional distribution:

$$\begin{aligned}
p(\mathbf{f} | \mathbf{y}, \Theta) &\sim \mathcal{N}(\mathbf{Q}_f^{-1}(\mathbf{Q}_f \boldsymbol{\mu}_f - \mathbf{Q}_{fy}(\mathbf{y} - \boldsymbol{\mu}_y)), \mathbf{Q}_f^{-1}) \\
&:= \mathcal{N}(\boldsymbol{\mu}_{f|y}, \mathbf{Q}_{f|y}^{-1})
\end{aligned}$$

with (see also Chan and Jeliazkov, 2009, eqn. 6-8):

$$\begin{aligned}
\mathbf{Q}_{f|y} &= \mathbf{Q}_f \\
&= \mathbf{H}^\top \mathbf{V}_v^{-1} \mathbf{H} + \mathbf{M}^\top \mathbf{V}_e^{-1} \mathbf{M} \\
\boldsymbol{\mu}_{f|y} &= \mathbf{Q}_f^{-1} \mathbf{Q}_f \tilde{\mathbf{c}} + \mathbf{M}^\top \mathbf{V}_e^{-1} \mathbf{y}
\end{aligned}$$

Draws from this distribution can be obtained efficiently since the $T \cdot S \times R$ matrix \mathbf{Q}_f is banded and its inversion not required. Specifically, calculate the conditional mean following Rue and Held (2005, Algorithm 2.1). This requires the computation of the lower Cholesky factor \mathbf{L} of \mathbf{Q}_f . Then draw $\mathbf{v} \sim \mathcal{N}(0, I_{TR})$, solve $\mathbf{L}^\top \mathbf{x} = \mathbf{v}$ for \mathbf{x} and set $\boldsymbol{\mu}_{f|y} + \mathbf{x}$, yielding a draw of \mathbf{f} conditional on \mathbf{y} (Rue and Held, 2005, Algorithm 2.4).

2.2 Forecasting

In addition to a draw of the factors conditional on the data, macroeconomic applications may require forecasts - both unconditional and conditional on future paths of a subset of the observables - from the model in Equations (1a)–(1c). Examples include projections of inflation which take assumptions regarding the macroeconomic environment into account (Giannone et al., 2014) or external nowcasts as "jump-off" points for longer-term forecasts in reduced form or structural models (Faust and Wright, 2009; Wolters, 2015; Del Negro and Schorfheide, 2013); Knotek and Zaman (2019) show how macroeconomic forecasts can be improved by conditioning on nowcasts for financial variables.

Let $\mathbf{y}^o = [\mathbf{y}_1^\top, \dots, \mathbf{y}_T^\top]^\top$ denote the available observations through T . For the forecast periods $h = 1 : H$, $\mathbf{y}_{c,t} \forall t = T + 1 : T + H$ are the observations which are being conditioned on. The entire conditioning set is given by

$$\mathbf{y}^c = [\mathbf{y}_{c,T+1}^\top, \dots, \mathbf{y}_{c,T+H}^\top]^\top$$

with $N^c = \dim(\mathbf{y}^c)$. Conversely, $\mathbf{y}_{-c,t} = \mathbf{y}_t \notin \mathbf{y}_{c,t}$ denotes the observations for which no conditioning information is available at time t and which are being forecast: $\mathbf{y}^f = [\mathbf{y}_{-c,T+1}^\top, \dots, \mathbf{y}_{-c,T+H}^\top]^\top$.

In what follows, we will not be concerned with the estimation of model parameters and assume that random draws from $p(\Theta|\mathbf{y}^o)$ are readily available. Also, for the sake of exposition, assume that between periods $t = 1$ and $t = T$ all observations are available. Situations where this is not the case, e.g. because some series start at a later date in the sample or because outlying observations have been removed by setting them to missing, can naturally be accommodated.

From a Bayesian perspective, forecasting boils down to sampling from the predictive density given by (see e.g. Geweke and Whiteman, 2006)

$$p(\mathbf{y}^f|\mathbf{y}^o, \mathbf{y}^c) \propto \int_{\Theta} p(\mathbf{y}^f|\mathbf{y}^o, \mathbf{y}^c, \Theta) p(\Theta|\mathbf{y}^o, \mathbf{y}^c) d\Theta. \quad (4)$$

Note that this density takes into account that the conditioning set \mathbf{y}^c contains information about which Θ are more likely a posteriori. Draws from this density can be obtained by applying the insights in Hauber and Schumacher (2021) who generalize the precision sampler outlined above to applications with missing data. Since future values of the variables can simply be considered as missing, draws from the joint distribution of states and forecasts conditional on the data (and possibly the future paths of a subset of the variables) can be sampled efficiently.

To evaluate Equation 4, derive the joint distribution of states and observations from $t = 1 : T + H$:

$$\begin{aligned} \mathbf{z} &= [\mathbf{f}^\top, \mathbf{y}^\top]^\top \\ &= [\mathbf{f}_1, \dots, \mathbf{f}_{T+H}, \mathbf{y}_1, \dots, \mathbf{y}_{T+H}]^\top \end{aligned}$$

It follows that the distribution of \mathbf{z} is Normal with mean $\boldsymbol{\mu}_z = 0$ and precision matrix \mathbf{Q}_z as in Section 2.1. To sample from the conditional distribution $p(\mathbf{y}^f, \mathbf{f} | \mathbf{y}^o, \mathbf{y}^c, \Theta)$ requires a permutation of \mathbf{z} . To this end, define the permutation matrix \mathcal{P} that reorders \mathbf{z} such that the states and elements of \mathbf{y}^f - ordered by time period t - are placed first. That is to say,

$$\mathbf{z}_{\mathcal{P}} := \mathcal{P}\mathbf{z} = [\mathbf{f}_1^\top, \dots, \mathbf{f}_T^\top, \mathbf{f}_{T+1}^\top, \mathbf{y}_{-c, T+1}^\top, \dots, \mathbf{f}_{T+H}^\top, \mathbf{y}_{-c, T+H}^\top, \mathbf{y}_1^\top, \dots, \mathbf{y}_{T+H}^\top, \mathbf{y}_{c, T+1}^\top, \dots, \mathbf{y}_{c, T+H}^\top]^\top.$$

Note that the conditioning arguments $\mathbf{y}^o, \mathbf{y}^c$ are now ordered last. The permuted vector of states and observations is also Normal with (permuted) moments given by

$$\mathbf{z}_{\mathcal{P}} \sim \mathcal{N}(\mathcal{P}\boldsymbol{\mu}_z, (\mathcal{P}\mathbf{Q}_z\mathcal{P}^\top)^{-1})$$

where we have made use of the fact that the inverse of a permutation matrix is equal to its transpose. The precision matrix of $\mathbf{z}_{\mathcal{P}}$ can be partitioned as follows:

$$\begin{aligned} \mathbf{Q}_{z_{\mathcal{P}}} &:= \mathcal{P}\mathbf{Q}_z\mathcal{P}^\top \\ &= \begin{bmatrix} \mathbf{Q}_{\mathbf{f}\mathbf{y}^f} & \mathbf{Q}_{\mathbf{f}\mathbf{y}^f, \mathbf{y}^o} \\ \mathbf{Q}_{\mathbf{f}\mathbf{y}^f, \mathbf{y}^o}^\top & \mathbf{Q}_{\mathbf{y}^o} \end{bmatrix} \end{aligned}$$

where $\mathbf{Q}_{\mathbf{f}\mathbf{y}^f, \mathbf{y}^o}$ is a $(T \cdot R + H \cdot N) \times T \cdot N$ matrix and $\mathbf{Q}_{\mathbf{f}\mathbf{y}^f}$ and $\mathbf{Q}_{\mathbf{y}^o}$ are square with dimensions $T \cdot R + H \cdot N$ and $T \cdot N$, respectively.

Similarly, the mean of $\mathbf{z}_{\mathcal{P}}$ can be partitioned as

$$\boldsymbol{\mu}_{z_{\mathcal{P}}} = \begin{bmatrix} \boldsymbol{\mu}_{\mathbf{f}\mathbf{y}^f} \\ \boldsymbol{\mu}_{\mathbf{y}^o} \end{bmatrix}$$

Denote by \mathbf{z}^* a draw from the conditional distribution (surpressing the dependence on the parameters Θ for the sake of readability)

$$\mathbf{z}_{\mathbf{f}\mathbf{y}^f} | \mathbf{y}^o, \mathbf{y}^c \sim \mathcal{N}(\mathbf{Q}_{\mathbf{f}\mathbf{y}^f}^{-1}(\mathbf{Q}_{\mathbf{f}\mathbf{y}^f}\boldsymbol{\mu}_{\mathbf{f}\mathbf{y}^f} - \mathbf{Q}_{\mathbf{f}\mathbf{y}^f, \mathbf{y}^o}(\mathbf{y}^o - \boldsymbol{\mu}_{\mathbf{y}^o})), \mathbf{Q}_{\mathbf{f}\mathbf{y}^f}).$$

Then by reversing the permutation, i.e

$$\begin{bmatrix} \mathbf{f} \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{T+H} \end{bmatrix} = \mathcal{P}^{-1} \begin{bmatrix} \mathbf{z}^* \\ \mathbf{y}^o \\ \mathbf{y}^c \end{bmatrix}$$

we can back out the draws of \mathbf{f} and $\mathbf{y}^f = [\mathbf{y}_{-c, T+1}, \dots, \mathbf{y}_{-c, T+H}]$.

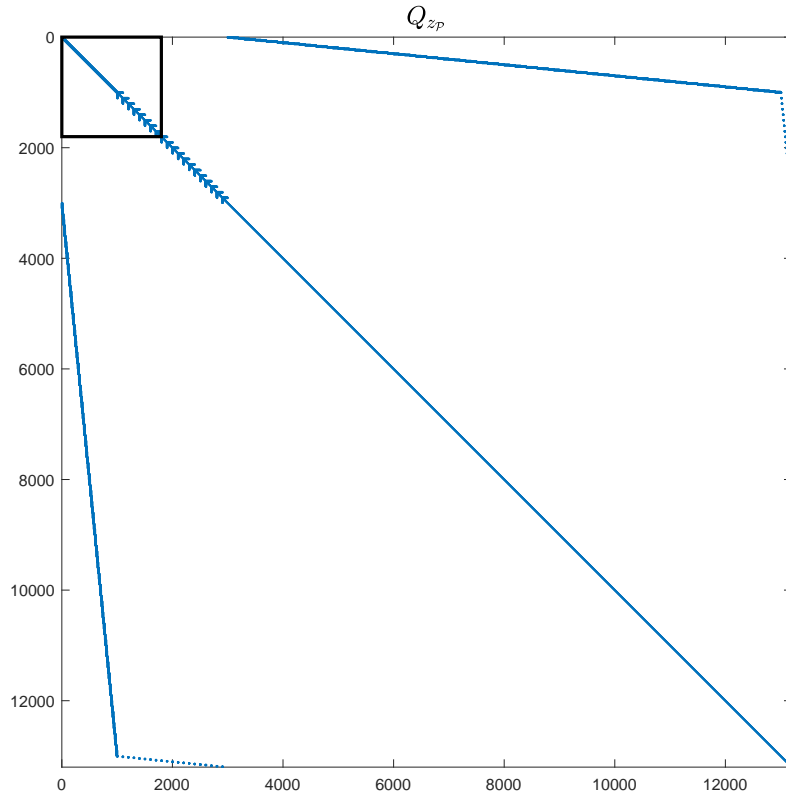
Note that by having placed those components of \mathbf{z} which are conditionally dependent close to each other, the bandwidth of $\mathbf{Q}_{\mathbf{f}\mathbf{y}^f}$ is kept small. Figure 1 shows the precision matrix of $\mathbf{z}_{\mathcal{P}}$ for a factor model with $T = 100$, $N = 100$ and $R = 10$, a forecast horizon of $H = 20$

and conditioning on the first 10 variables², i.e.

$$\mathbf{y}^c = [y_{1,T+1}, \dots, y_{10,T+1}, \dots, y_{1,T+20}, \dots, y_{10,T+20}].$$

The precision matrix $\mathbf{Q}_{\mathbf{y}^f}$ is highlighted by the square.

Figure 1: Precision matrix in the case of a dynamic factor model



Note: The figure shows the precision matrix of the permuted vector of observations and states $\mathbf{z}_{\mathcal{P}}$ for a factor model with $T = 100$, $N = 100$ and $R = 10$. Highlighted is the $(T + H) \cdot R + N^c \times (T + H) \cdot R + N^c$ partition corresponding to $\mathbf{z}_{\mathbf{y}^f}$.

Similar to Waggoner and Zha (1999, Algorithm 1) and Banbura et al. (2015, p.745), draws from the predictive density $p(\mathbf{y}^f | \mathbf{y}^o, \mathbf{y}^c)$ can then be obtained by the following algorithm:

²For the sake of illustration we set $P = 1$, $J = 0$, $S = 0$.

Algorithm 1: Draws from the predictive density $p(\mathbf{y}^f | \mathbf{y}^o, \mathbf{y}^c)$

Initialize Θ^0 from $p(\Theta | \mathbf{y}^o)$. For $g = 1 : (G_0 + G)$

1. Construct the system matrices $\mathbf{H}, \mathbf{M}, \mathbf{V}_\varepsilon, \mathbf{V}_v, \tilde{\mathbf{c}}, \mathbf{d}$ given the parameters $\Theta^{(g-1)}$
2. Compute the mean $\boldsymbol{\mu}_z$ and precision matrix \mathbf{Q}_z of the joint distribution of $\mathbf{z} = [\boldsymbol{\alpha}, \mathbf{y}]$
3. Permute \mathbf{z} , yielding $\mathbf{z}_{\mathcal{P}'} \sim \mathcal{N}(\boldsymbol{\mu}_{z_{\mathcal{P}'}} , \mathbf{Q}_{z_{\mathcal{P}'}}^{-1})$ with partitions $[\mathbf{z}_{\alpha y^f}^\top, \mathbf{y}^o, \mathbf{y}^c]^\top$
4. Sample from the conditional distribution $\mathbf{z}_{\alpha y^f} | \mathbf{y}^o, \mathbf{y}^c, \Theta^{(g-1)}$ using (Rue and Held, 2005, Algorithm 2.1 and 2.4) and reverse the permutation to back out the draws of \mathbf{y}^f and $\boldsymbol{\alpha}$
5. Draw $\Theta^{(g)}$ from $p(\Theta | \mathbf{y}^o, \mathbf{y}^c, \mathbf{y}^{f(g)}, \boldsymbol{\alpha}^{(g)})$

Discard the first G_0 draws as burn-in.

If the sample information and prior dominate the information in \mathbf{y}^c , the impact of the conditioning arguments on the posterior distribution of the parameters is likely to be negligible and $p(\Theta | \mathbf{y}^o, \mathbf{y}^c) \approx p(\Theta | \mathbf{y}^o)$ (Del Negro and Schorfheide, 2013; Banbura et al., 2015). In such applications, draws from Equation 4 can then be obtained by skipping step 5 at each iteration in **Algorithm 1**. Similarly, when \mathbf{y}^c is empty, Step 5 can obviously be skipped as well and the algorithm returns unconditional forecasts, i.e. draws from the density $p(\mathbf{y}^f | \mathbf{y}^o)$.

While the discussion above has focussed on dynamic factor models, it can easily be extended to general state space models of the form:

$$\mathbf{y}_t = \mathbf{Z} \boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_t \sim \mathcal{N}(0, \mathbf{H}) \quad (5a)$$

$$\boldsymbol{\alpha}_{t+1} = \mathbf{T} \boldsymbol{\alpha}_t + \boldsymbol{\eta}_t, \boldsymbol{\eta}_t \sim \mathcal{N}(0, \mathbf{Q}) \quad (5b)$$

by setting $P = 1, J = S = 0$.

Lastly, note that Hauber and Schumacher (2021) also propose a sequential sampler that first draws the missing observations conditional on the states and then the states given a complete data set. This approach performs somewhat better in terms of runtime and could in principle be used in Algorithm 1, which in any case requires a Gibbs sampling step if the parameters are updated given the conditioning set. However, the gains in performance from faster sampling of states and forecasts may not outweigh the longer burn-in required due to the additional block in the Gibbs Sampler.

2.3 Soft conditions and repeated samples from $p(\mathbf{y}^f | \mathbf{y}^o, \Theta^{(g)})$

So far we have considered conditional forecasts that fix the future values of some of the endogeneous variables at single points. Waggoner and Zha (1999) labels these "hard" conditions to distinguish them from soft conditions where each element of \mathbf{y}^c is merely restricted to lie within a pre-specified range, i.e. $\mathbf{y}_l^f \leq \mathbf{y}^f \leq \mathbf{y}_u^f$. Draws from the unconditional predictive density $p(\mathbf{y}^f | \mathbf{y}^o, \Theta^{(g)})$ can be obtained until the conditions are satisfied.³

The precision-sampling algorithms outlined above are particularly suited to this task since the bottleneck in the calculations is the Cholesky factorization of the band matrix \mathbf{Q}_{f,y^f} which requires $((T + H)R + NH) \cdot b_w^2$ floating point operations where b_w is the lower bandwidth of \mathbf{Q}_{f,y^f} (Golub and Van Loan, 2013, 4.3.5). However, it only needs to be computed once, irrespective of the number of desired draws (Rue, 2001). Repeated samples thus come at a significantly reduced computational burden, consisting only of $(T + H)S + NH$ independent draws from the standard Normal distribution, band backward substitution - which requires $2((T + H)S + NH) \cdot b_w$ floating point operations (Golub and Van Loan, 2013, 4.3.2)- and a vector addition.

2.4 Comparison with Kalman filter-based simulation smoothers

In simulations to determine the computational efficiency of the algorithms outlined above in comparison to Kalman filter-based simulation smoothers, we find that the former is computationally more efficient in terms of runtime for both large and small factor models and forecast horizons. The gains in computational efficiency are particularly pronounced for soft conditioning, where repeated draws from the predictive density are required. See Appendix 8 for details of the simulation study.

3 Real-time evaluation of unconditional and conditional forecasts

In this section, we present the data used in the forecast evaluation and provide details on the Bayesian estimation. Furthermore, we discuss the set-up of the real-time evaluation.

3.1 Data

3.1.1 Real-time dataset for the German economy

To estimate the factor model, we construct a large quarterly real-time dataset covering different aspects of the macroeconomy. Beyond real gross domestic product, the main building

³Similarly, hard and soft conditions can be combined, e.g. if inflation forecasts are produced conditional on an oil price path and subject to the restriction that the Federal Funds rate is positive. In such a scenario, draws from $p(\mathbf{y}^f | \mathbf{y}^o, \mathbf{y}^c, \Theta^{(g)})$ are obtained until the soft restrictions are satisfied.

block of the dataset are other series from the national accounts such as the entire expenditure side components, e.g. private and public consumption, gross fixed capital formation, export and imports. From the production side of GDP we can only include gross value added in the industrial and construction sector as well as trade, transport and hospitality as real-time vintages for other sectors are only available from 2011 onwards.⁴

In addition to the chained volume indices we also include the corresponding deflators of the expenditure and production side components. Furthermore, our dataset contains series on real activity such as industrial production, turnover or orders as well as construction, headline CPI and PPI indices as well as the corresponding core indices and a measure of prices in the construction sector as well as labor market indicators such as economy-wide employment and wages and hours worked in the industrial and construction sector.

In addition, data are supplemented with financial indicators such as interest rates, exchange rates and stock market indices. Lastly, we also include a few survey-based indicators provided by the European Commission (ESI) and covering the industrial, construction and services sectors as well as a measure of firms' employment expectations (EEI). Note that for the latter two sources, there is no real-time problem, as financial indicators are not revised subsequent to their original publication. Similarly, the European Commission's survey indicators are only revised due the seasonal adjustment procedure. That is to say, the survey responses for May 2006 as reported in 2018 might differ (slightly) from those originally published as the seasonal factors have since been updated. While this violates the real-time assumption, such revisions are likely to be small. We therefore abstract from these subsequent changes in what follows and rely on the seasonally adjusted indices.

Altogether, our dataset contains 57 series. All of them are transformed to stationarity prior to estimation. In most cases, this involves taking the first difference of the logarithm of the original series, though in some cases, e.g. interest rates or survey indicators, we take simple differences.⁵ Details of the time series, their sources as well as the transformation employed can be found in Appendix 6.

3.1.2 Professional forecasts from the Reuters Poll

As conditioning information, we rely on the Reuters Poll of forecasts for German GDP and CPI inflation. The survey compiles the views of around twenty different professional forecasters from the private sector and research institutes. It is conducted once a quarter during the first month, i.e. January, April, July and October. Until mid-2014, the respondents submit their views during the first few days of the month; from then on the survey date is around the middle of the month.⁶

For GDP, the forecasts are given in terms of quarter-on-quarter (q/q) growth rates and we can thus use the median forecast directly to obtain conditional forecasts from the model out-

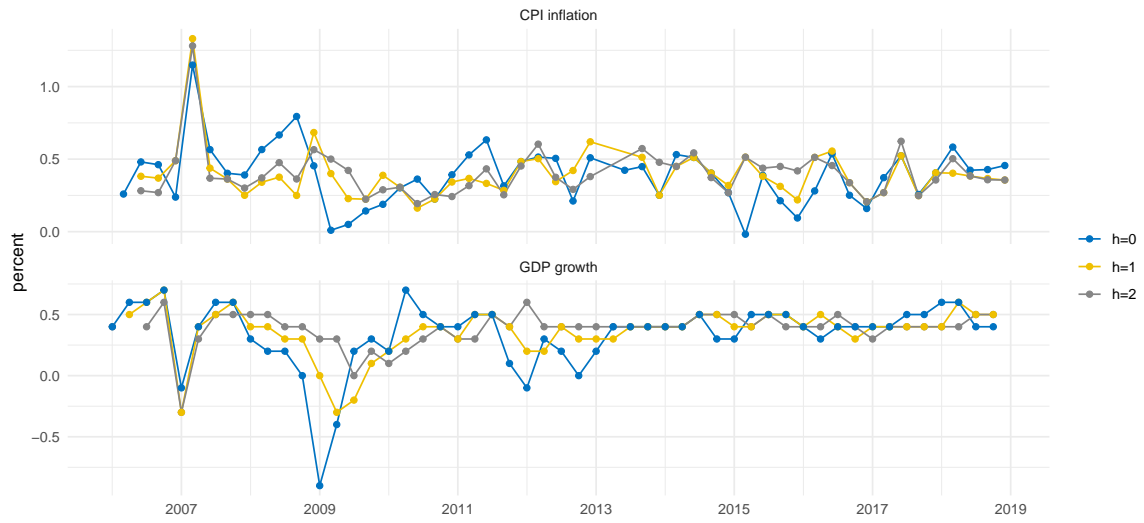
⁴The three mentioned sectors comprise roughly 50 percent of total gross value added.

⁵The Deutsche Bundesbank's real-time database includes vintages of chained volume indices and nominal values for the national accounts. The implicit deflators are calculated by dividing the nominal series with the chained volume index and taking the first difference of the logarithm of the resulting series.

⁶A complete list of the dates when the forecast polls were conducted is available upon request.

lined above. The inflation forecasts, however, are reported as year-over-year (y/y) growth rates of the consumer price index. To bring the forecasts in line with the treatment of CPI inflation in the model, we have to transform the individual forecasts. Specifically, we interpret the y/y forecasts of inflation, $x_t^{y/y}$, as describing the four-quarter change in the quarterly consumer price index X , i.e. $x_t^{y/y} = \log(X_t/X_{t-4})$. Thus, given a sequence of y/y forecasts we can back out the implicit forecasts of the (log of the) consumer price index and derive the corresponding q/q forecasts, i.e. $x_t^{q/q} = \log(X_t/X_{t-1})$ for each participant in the survey. We then use the median of these transformed values in the conditional forecasting exercise.

Figure 2: GDP growth and CPI inflation forecasts from the Reuters Poll of professional forecasters



Note: The figure shows the median forecast from the Reuters Poll of professional forecasters for quarter-on-quarter change in the consumer price index (top panel) and quarter-on-quarter GDP growth. The forecast horizon h is in quarters and relative to the reference period. For details, see the main text. *Source:* Thomson Reuters, author's calculations.

Figure 2 plots the median forecast of q/q CPI inflation (top row) and GDP growth (bottom row). The lines correspond to different forecast horizons relative to the reference period and are given in quarters. That is to say, the $h = 1$ forecast for the first quarter in 2010 was made at the start of October 2009.

3.2 Estimation

The parameters of the factor model outlined above are estimated with Bayesian methods. Draws from the joint posterior distribution of factors and parameters are obtained via a Gibbs Sampler which alternately draws from the conditional posterior of the factors given the parameters and *vice versa*. For the estimation, we set $J = 1$ and $S = 0$ and $R = 2$ and run the Gibbs Sampler for a total of 20000 iterations. We discard the first 10000 as burn-in and of the remaining, store every second draw, yielding 5000 from the posterior distribution of the parameters, $p(\Theta|y)$. As the priors used in the estimation as well as the resulting

conditional posterior distributions are standard in the literature, we relegate a discussion to Appendix 7.

3.3 Evaluation set-up

Given a vintage reflecting the data available to forecasters at the time the Reuters Poll was conducted and corresponding draws from the posterior distribution $p(\Theta|y)$, we generate draws from the predictive density as described in Section 2.1 for all national accounts, price, labor market and activity indicators. These forecasts are unconditional as well as conditional on the values from the Reuters Poll for GDP growth and CPI inflation for the current and following two quarters, i.e. $h = \{0, 1, 2\}$. The evaluation sample starts in 2006 from which point on real-time vintages are available on a broad basis. It ends in the fourth quarter of 2017 as afterwards the number of missing entries in the Reuters Poll increases considerably.

Within the sample, there are also a few data irregularities. For example, in the third quarter of 2013 no CPI inflation forecasts were recorded in the Reuters Poll. We therefore exclude 2012Q3, 2012Q4 and 2013Q1 from the evaluation. Note also that for two labor market series there are some vintages which are released outside of the regular publication calendar. For example, the employment vintage scheduled for release at the end of December was not published until the first week of January, most likely due to the original publication date falling on a bank holiday. As such, when the professional forecasters were asked by Reuters to submit their forecasts by January 1st, the latest available information for this series was published at the end of November. For hours worked in the manufacturing sector, all vintages between March 3rd 2009 and March 4th 2010 are missing. As these irregularities alter the information set and effective forecast horizon compared to the other quarters, we do not consider the affected vintages for these two series. For the remaining series, however, we have a total of 41 quarters to be evaluated for all horizons.

To benchmark the performance of the different priors, we consider a simple univariate Bayesian autoregressive model of order 2 for all the series. i.e.

$$y_{i,t} = \beta_0 + \beta_1 y_{i,t-1} + \beta_2 y_{i,t-2} + \epsilon_t; \epsilon_t \sim \mathcal{N}(0, \sigma^2) \quad (6)$$

The parameters $\beta = [\beta_0, \beta_1, \beta_2]$ and σ^2 are estimated via a Gibbs Sampler with diffuse priors given by $\beta \sim \mathcal{N}(0, 10 \cdot I_3)$ and $\sigma^2 \sim \mathcal{G}^{-1}(3, 0.01)$. Draws from the h-step ahead predictive density are obtained by plugging in draws from the posterior distribution of the coefficients and iterating Equation 6 forward.

Both the point and density forecast performance is evaluated. For the former, we compute the root mean squared forecast error (RMSFE) defined as

$$\text{RMSFE} = \left(\frac{1}{S} \sum_{s=1}^S (\hat{y}_{i,T+s} | \Omega_v - y'_{i,T+s})^2 \right)^{\frac{1}{2}}$$

for a sequence of S nowcasts, where $\hat{y}_{i,T+s} | \Omega_v$ is the mean of the predictive density of variable

i at time $T + s$, conditional on the information set Ω_{v_s} available in real-time at date v_s when the forecast for period $T + s$ was made. The corresponding realization, given by the first release, is denoted by $y'_{i,T+s}$. Density forecasts are evaluated by the continuous ranked probability score (CRPS) which is given by (see Krüger et al., 2016):

$$\text{CRPS} = \frac{1}{S} \sum_{s=1}^S \left(\frac{1}{G} \sum_{k=1}^G \left| y_{i,T+s}^{(k)} | \Omega_{v_s} - y'_{i,T+s} \right| - \sum_{k=1}^G \sum_{j=1}^G \left| y_{i,T+s}^{(k)} | \Omega_{v_s} - y_{i,T+s}^{(j)} | \Omega_{v_s} \right| \right).$$

To calculate the CRPS given draws from the predictive density as well a realization, we use the R package `scoringRules` (Jordan et al., 2019).

Finally, to compare the accuracy of the conditional and unconditional forecasts more formally, we follow Ganics and Odendahl (2021) and run Diebold-Mariano tests (Diebold and Mariano, 1995). Let $L_t(y_{i,t+h}^c)$ denote the loss associated with the conditional forecasts made at time t for variable i and horizon h . In the case of point forecasts, this is simply given by the squared forecast error; for density forecasts by the corresponding CRPS. Similarly, define $L_t(y_{i,t+h}^u)$ as the loss for the unconditional forecast and $d_t = L_t(y_{i,t+h}^c) - L_t(y_{i,t+h}^u)$ the loss differential. The test statistic can then be written as

$$DM_{i,h} = \frac{\frac{1}{T} \sum_{t=1}^T d_t}{\sqrt{\hat{\sigma}_d^2}}$$

where $\hat{\sigma}_d^2$ is an estimate of the variance of d_t . We test the null hypothesis of equal predictive ability against the one-sided alternative that the conditional forecasts yield a lower loss.

4 Results

In this section we present the results of the forecast evaluation. We begin with point forecast performance before looking at how the evaluation changes when we focus on the entire predictive distribution. Lastly, we consider several robustness checks like the evaluation period as well as the model size (i.e. the number of factors).

4.1 Point forecast accuracy

Turning to the performance of the unconditional forecasts first, we find that for all horizons and the majority of variables, the factor model produces forecasts that perform similar to the autoregressive benchmark with relative RMSFE between 0.9 and 1.1 (Figure 3). Outliers in this context are the headline and "core", i.e. excluding energy, producer price indices at $h = 0$ where the performance is considerably worse than the benchmark. Conversely, for a few series there are noteworthy gains over the simple univariate AR(2) model and these persist or even increase for larger horizons. For example, for $h = 0$ the relative RMSFE for the construction price index and gross value added in the industrial sector is around 0.6. Smaller but still sizeable improvements can also be found for two series from the

expenditure side of the national accounts: equipment investment as well as the deflator for residential investment. For larger h , these relative gains even increase: the relative RMSFE for the construction subcomponent of the PPI reaches 0.35; for gross value added in the industrial sector is 0.4. In addition to these variables, at $h = 2$ there also noteworthy improvements for series that are likely highly correlated to gross industrial value added like production in the industrial and construction sector as well as industrial turnover. Note, however, that these improvements are partly due to the fact that the performance of the benchmark generally worsens as the forecast horizon increases.

For comparison, the model's unconditional forecasts for GDP growth lead to substantial relative gains of up to 20 percent relative to the autoregressive benchmark. For $h = 0$, the professional forecasters are considerably more accurate (relative RMSFE of 0.65) but for larger horizons the performance is similar to the unconditional forecasts. The picture for CPI inflation is similar: the professional forecasters are more accurate than the model's unconditional forecast at $h = 0$ but by a much smaller margin; for $h = \{1, 2\}$, the forecasts from the Reuters Poll are considerably (slightly) worse than the autoregressive benchmark (the model's unconditional) forecasts.

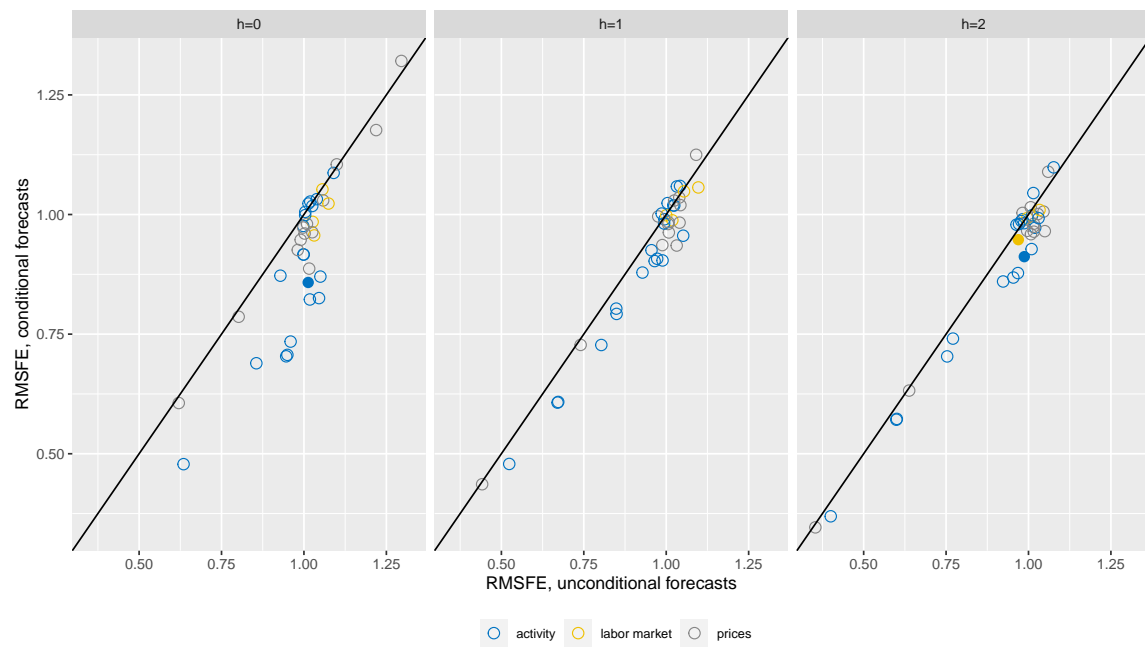
Focusing on how the forecast performance increases when we condition on professional forecasters' views on GDP growth and CPI inflation, we find that there are only large differences in the relative RMSFE between the unconditional and conditional forecasts for $h = 0$. Perhaps not surprisingly gains to conditioning arise for those series for which the models already produce decent forecasts (relative to autoregressive benchmarks) and are mainly concentrated to series from the national accounts group and - to some extent - indicators capturing real activity (but not part of the national accounts). For the labor market indicators and CPI and PPI inflation, we find much smaller gains from conditioning. Filled entries in the above plot correspond to those variables for which the null hypothesis of the Diebold-Mariano test can be rejected at the 5 percent level in favor of the alternative that the loss of the conditional forecasts is lower. This is almost never the case for any of the 37 variables we consider irrespective of the forecast horizon (Table 1). At $h = 0$, i.e. forecasts that are made in the reference quarter, we reject the null for the gross value added in the trade, transport and hospitality sector; at $h = 2$ for gross value added in the trade, transport and hospitality sector and employment.

Table 1: Point forecast evaluation

| category/horizon | h=0 | h=1 | h=2 |
|------------------|-------------|-------------|-------------|
| activity | 1/20 | 0/20 | 1/20 |
| labor market | 0/4 | 0/4 | 1/4 |
| prices | 0/13 | 0/13 | 0/13 |
| total | 1/37 | 0/37 | 2/37 |

The table shows the fraction of times the null hypothesis of the Diebold-Mariano test is rejected at the 5 percent significance level against the one-sided alternative that the conditional forecasts yield a lower loss. For details, see the main text.

Figure 3: Point forecast evaluation

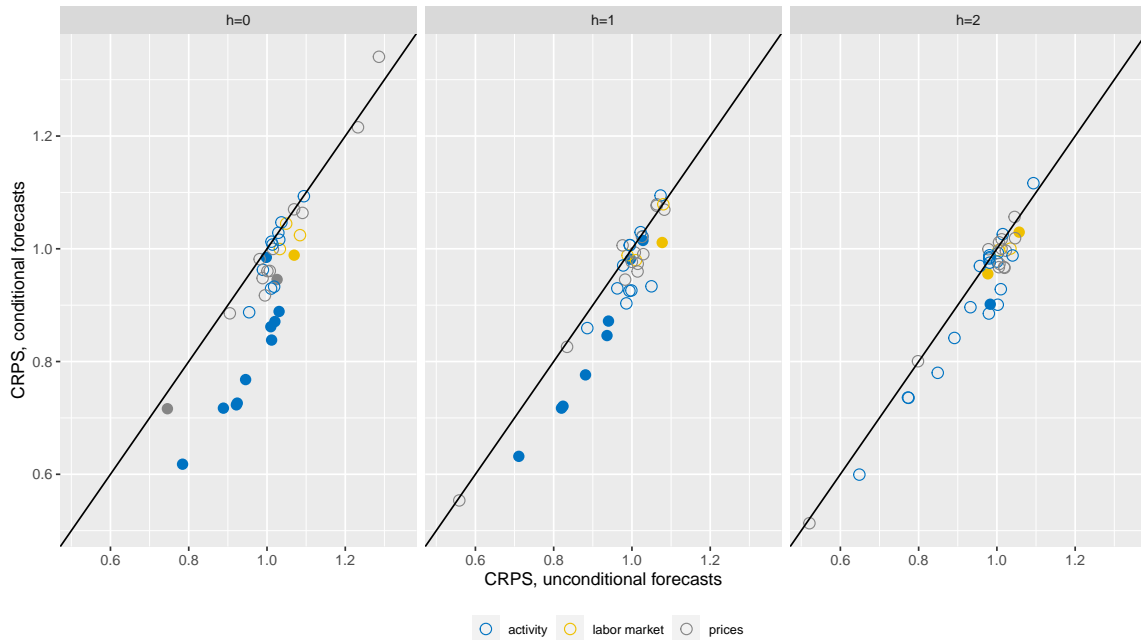


Note: The figure shows the root mean squared forecast errors (RMSFE) corresponding to unconditional forecasts (x-axis) and forecasts conditional on professional forecasters' view on GDP growth and CPI inflation (y-axis) for different time series. For each series, the RMSFE is relative to an autoregressive benchmark. Entries above (below) the 45-degree line indicate that conditional forecasts perform worse (better) than the unconditional ones. Filled points correspond to those variables for which the null hypothesis of the Diebold-Mariano test can be rejected at the 5 percent level. For details, see the main text.

4.2 Density forecast accuracy

By and large, the results obtained in terms of point forecast accuracy carry over when we evaluate the entire predictive density. In particular, the performance relative to the benchmark as well as the gains from conditioning as measured by the average CRPS are very similar (Figure 4). Contrary to the findings for point forecasts, however, we do find significant differences between the conditional and unconditional forecast performance, at least for $h = \{0, 1\}$, again highlighted in the figure by the filled dots. For example, at $h = 0$ the null of equal predictive accuracy is rejected for a third of all the variables that we consider, mainly concentrated in the activity group. For one-quarter ahead forecasts the number decreases slightly but for many activity series we still find significant gains (Table 2).

Figure 4: Density forecast evaluation



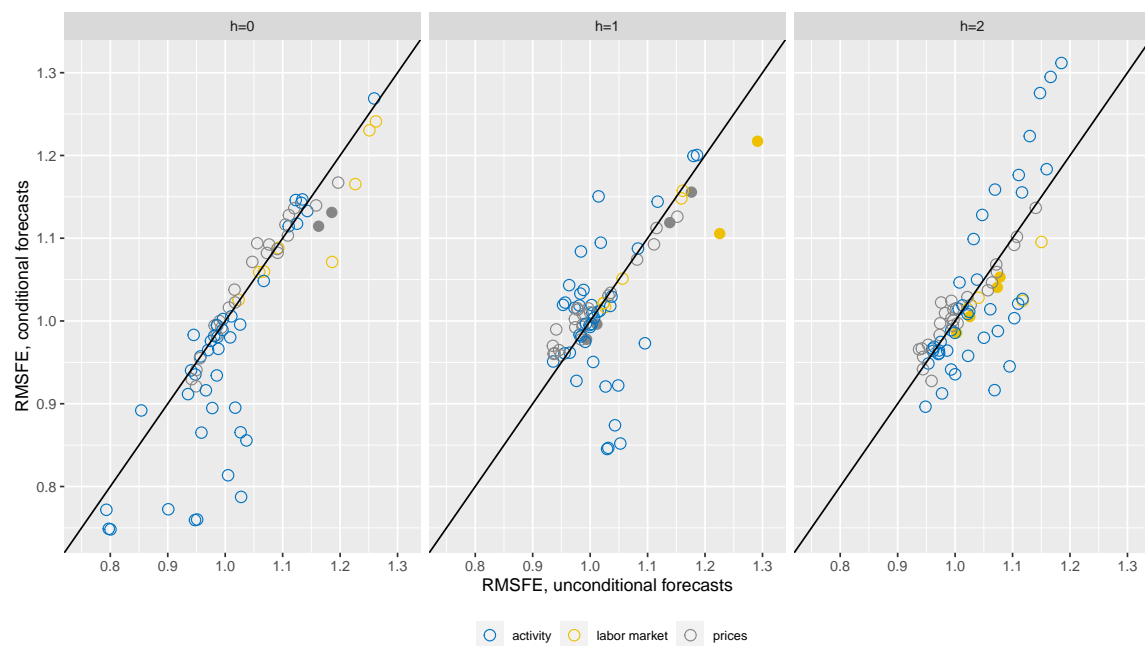
Note: The figure shows the average continuous ranked probability score (CRPS) corresponding to unconditional forecasts (x-axis) and forecasts conditional on professional forecasters' view on GDP growth and CPI inflation (y-axis) for different time series. For each series, the CRPS is relative to an autoregressive benchmark. Entries above (below) the 45-degree line indicate that conditional forecasts perform worse (better) than the unconditional ones. Filled points correspond to those variables for which the null hypothesis of the Diebold-Mariano test can be rejected at the 5 percent level. For details, see the main text.

4.3 Robustness checks

As a first robustness check, we drop the global financial crisis of 2008-09 from the evaluation sample, to see if the gains in predictive accuracy for point and density forecasts are also apparent in more tranquil times. Figure 5 presents the results for the evaluation subsample ranging from 2011Q1 to 2017Q4.

Overall, the forecast performance declines markedly relative to the autoregressive benchmarks but for some series, gains from conditioning still arise. However, for $h = \{1, 2\}$ we

Figure 5: Robustness check: post-crisis evaluation sample



Note: The figure shows the root mean squared forecast errors (RMSFE) over the evaluation sample 2011Q1-2017Q4 corresponding to unconditional forecasts (x-axis) and forecasts conditional on professional forecasters' view on GDP growth and CPI inflation (y-axis) for different time series. For each series, the RMSFE is relative to an autoregressive benchmark. Entries above (below) the 45-degree line indicate that conditional forecasts perform worse (better) than the unconditional ones. Filled points correspond to those variables for which the null hypothesis of the Diebold-Mariano test can be rejected at the 5 percent level. For details, see the main text.

Table 2: Density forecast evaluation

| category/horizon | h=0 | h=1 | h=2 |
|------------------|--------------|-------------|-------------|
| activity | 10/20 | 8/20 | 1/20 |
| labor market | 1/4 | 1/4 | 2/4 |
| prices | 2/13 | 0/13 | 0/13 |
| total | 13/37 | 9/37 | 3/37 |

The table shows the fraction of times the null hypothesis of the Diebold-Mariano test is rejected at the 5 percent significance level against the one-sided alternative that the conditional forecasts yield a lower loss. For details, see the main text.

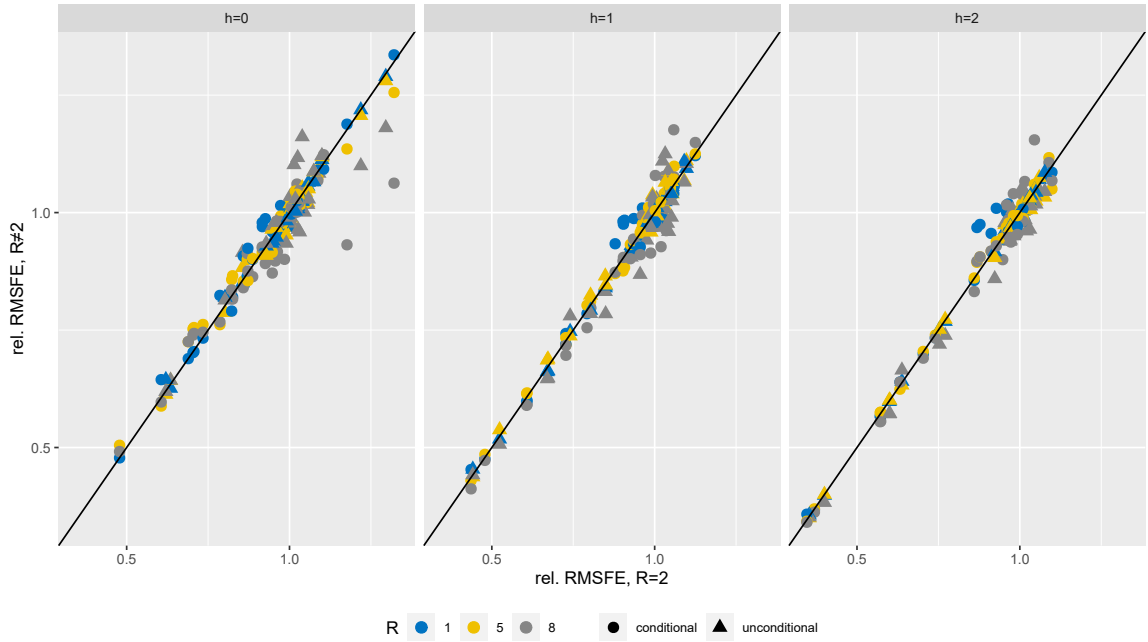
also find that for many series - particularly from the activity group - conditional forecasts are on average less accurate. Rejection of the null of the Diebold-Mariano test still occurs but with a few exceptions, this occurs in cases where the autoregressive benchmark actually performs better than the factor model. As such, it is fair to say that developments during the financial crisis exert some influence on the results presented above for the entire evaluation sample.

Besides the evaluation period, the results presented above may also be sensitive to the model specification. In particular, they were obtained under a given number of factors, $R = 2$. However, differently specified models yield similar forecast performance. Figure 6 shows the relative RMSFE of the model with $R = 2$ on which the results above are based as well as both smaller ($R = 1$) and larger ($R = 5, 8$) models. There is little indication that model specification systematically alters the forecast performance. Conditional forecasts at $h = 0$ for the change in the producer index improve substantially when $R = 8$ with the relative RMSFE decreasing by 0.25. To a lesser extent this also holds for the unconditional forecasts with the relative RMSFE dropping by around 0.1. However, this is the exception and for other horizons and variables we find much smaller differences in either way. Consequently, the points corresponding to the relative RMSFE for $R = 1, 5, 8$ in Figure 6 all hug the 45-degree line, which indicates identical forecast performance, quite closely.

5 Conclusion

In this paper we assess the forecast performance of factor models in real-time for Germany covering the period 2006-2017. We contribute to the literature by broadening the horizon of the forecast evaluation to include a large number of the variables that are used to estimate the model. In addition, we also investigate to what extent the forecast performance of the model improves when we incorporate external information in the form of professional forecasters' views on GDP growth and CPI inflation. Conditioning on forecasts rather than actual realized values gives an accurate assessment of the uncertainty around conditional forecasts that is relevant to policy-makers by capturing both how accurate the model can summarize the comovement of the time series as well as the accuracy of the forecasts on

Figure 6: Robustness check: different R



Note: The figure shows the root mean squared forecast error (RMSFE) relative to the autoregressive benchmarks for different number of factors. On the x-axis are the values of the RMSFE for the model with 2 factors on which the main results are based. On the y-axis are the RMSFE for alternative model specifications, differentiated by colors. Entries above (below) the 45-degree line indicate that the model with 2 factors performs better (worse) than the alternative models.

which we condition.

Our results show that the factor model produces forecasts for most of the series and horizons that are as accurate as those from autoregressive benchmarks. In some cases they outperform the benchmark by a large margin. When conditioning on the forecasts of professional forecasters for GDP growth and CPI inflation, we generally find some improvements in forecast accuracy for the variables of interest from categories such as real activity, prices and the labor market. Only for a smaller number of series, however, do we find that conditioning produces large gains. Moreover, for point forecasts the differences are statistically insignificant for virtually all series. When evaluating the predictive densities with the CRPS, we do find significant gains at $h = 0$ ($h = 1$) for around a third (a quarter) of the variables under consideration. To some extent the results appear to be driven by the large drop in output in the wake of the Global Financial Crisis. as the forecast performance deteriorates relative to the benchmark when we shorten the evaluation sample to 2011-2017. On the other hand, the results are robust to model specification in the sense that increasing or decreasing the number of factors from $R = 2$ as in the baseline specification does not systematically alter the forecast performance.

Lastly, this paper proposes precision-based sampling algorithms for conditional and unconditional forecasts in factor models and more broadly state space models with unobserved components. Treating forecasts as missing observations, the insights in Hauber and Schumacher (2021) can be obtained directly to obtain draws from the relevant predictive

densities. Simulations document that these algorithms can be more efficient from a computational point of view than simulation smoothers based on the Kalman filter. While the exact quantitative results of the exercise may vary depending on platform and implementation, the qualitative results indicate that precision-based sampling algorithms are a viable alternative in such applications.

Appendix

6 Data

This Appendix provides detail on the real-time dataset used in the empirical application. The 57 time series can be grouped in to the following categories: **activity** which includes the gross domestic product as well as expenditure- and production-side components of output. In addition, series like industrial production, turnover and orders also form part of this category. The **prices** category contains the (implicit) deflators corresponding to the chained volume indices of the national accounts mentioned above as well as consumer and producer price series. The **labor market** is covered by series on total employment, wages as well as hours worked in the industrial and construction sector. These are the series which we focus on in the forecast evaluation and their vintages are all sourced from the Deutsche Bundesbank's Real-Time Database (RTD). Moreover, to capture the **financial** side of the economy, we include series such as interest rates, commodity prices and the CDAX stock market index. These series are also sourced from the Deutsche Bundesbank's website. Lastly, **survey** indicators are taken from the European Commission's Business and Consumer Surveys. We focus on the components of the Economic Sentiment Indicator (ESI) which capture confidence in the industrial, consumer, services, retail trade and construction sectors. In addition, we also include the employment expectations index (EEI).⁷ Table 3 lists the data series used in the empirical application, the group the indicators belong to, their source as well as any transformation applied to the series prior to estimation, e.g. whether series are in logs and/or differenced (diff).

⁷The series were downloaded on April 27th, 2021 from https://ec.europa.eu/economy_finance/db_indicators/surveys/documents/series/nace2_ecfin_2106/main_indicators_sa_nace2.zip

Table 3: Description of the dataset

| Series | Mnem. | Group | Transf. | Source | Notes |
|---|---------------|--------------|---------------------------|----------------|---|
| gross domestic product, chain index ^{s,c} | gdp | activity | log, diff | Bundesbank RTD | Code: BBKRTQ.DE.YA.AG1.CA010.A.I |
| gross domestic product, deflator ^{s,c} | p_gdp | prices | log, 1 st diff | Bundesbank RTD | implicit deflator, i.e. nominal series divided by chain index |
| private consumption expenditure, chain index ^{s,c} | c_priv | activity | log, diff | Bundesbank RTD | Code: BBKRTQ.DE.YA.CA1.BA100.A.I |
| private consumption expenditure, deflator ^{s,c} | p_c_priv | prices | log, diff | Bundesbank RTD | implicit deflator, i.e. nominal series divided by chain index |
| government consumption expenditure, chain index ^{s,c} | c_gov | activity | log, diff | Bundesbank RTD | Code: BBKRTQ.DE.YA.CA1.BA100.A.I |
| government consumption expenditure, implicit deflator ^{s,c} | p_c_gov | prices | log, diff | Bundesbank RTD | nominal series divided by chain index |
| equipment investment, chain index ^{s,c} | gfcf equip | activity | log, diff | Bundesbank RTD | Code: BBKRTQ.DE.YA.CEI.CA010.A.I |
| equipment investment, implicit deflator ^{s,c} | p_gfcf equip | prices | log, diff | Bundesbank RTD | implicit deflator, i.e. nominal series divided by chain index |
| construction investment, chain index ^{s,c} | gfcf constr | activity | log, diff | Bundesbank RTD | Code: BBKRTQ.DE.YA.CFI.CA010.A.I |
| construction investment, implicit deflator ^{s,c} | p_gfcf constr | prices | log, diff | Bundesbank RTD | implicit deflator, i.e. nominal series divided by chain index |
| other investment, chain index ^{s,c} | gfcf other | activity | log, diff | Bundesbank RTD | Code: BBKRTQ.DE.YA.CI1.CA010.A.I |
| other investment, implicit deflator ^{s,c} | p_gfcf other | prices | log, diff | Bundesbank RTD | nominal series (BBKRTQ.DE.YA.CI1.CA010.VA) divided by chain index |
| exports (of goods and services), chain index ^{s,c} | x | activity | log, diff | Bundesbank RTD | Code: BBKRTQ.DE.YA.CX1.CA010.A.I |
| exports (of goods and services), implicit deflator ^{s,c} | p_x | prices | log, diff | Bundesbank RTD | nominal series (BBKRTQ.DE.YA.CX1.CA010.VA) divided by chain index |
| imports (of goods and services), chain index ^{s,c} | m | activity | log, diff | Bundesbank RTD | Code: BBKRTQ.DE.YA.CMI.CA010.A.I |
| imports (of goods and services), implicit deflator ^{s,c} | p_m | prices | log, diff | Bundesbank RTD | nominal series (BBKRTQ.DE.YA.CMI.CA010.VA) divided by chain index |
| gross value added, industry, chain index ^{s,c} | gva_ind | activity | log, diff | Bundesbank RTD | Code: BBKRTQ.DE.YA.AU1.AA020.A.I |
| gross value added, industry, implicit deflator ^{s,c} | p_gva_ind | prices | log, diff | Bundesbank RTD | nominal series (BBKRTQ.DE.YA.CMI.CA010.VA) divided by chain index |
| gross value added, construction, chain index ^{s,c} | gva constr | activity | log, diff | Bundesbank RTD | Code: BBKRTQ.DE.YA.AU1.AA030.A.I |
| gross value added, construction, implicit deflator ^{s,c} | p_gva constr | prices | log, diff | Bundesbank RTD | nominal series (BBKRTQ.DE.YA.AU1.AA030.VA) divided by chain index |
| gross value added, trade, transport & hospitality, chain index ^{s,c} | gva_th | activity | log, diff | Bundesbank RTD | Code: BBKRTQ.DE.YA.AU1.AA040.A.I |
| gross value added, trade, transport & hospitality, implicit deflator ^{s,c} | p_gva_th | prices | log, diff | Bundesbank RTD | nominal series (BBKRTQ.DE.YA.AU1.AA040.VA) divided by chain index |
| consumer prices, index ^{s,c} | cpi | prices | log, diff | Bundesbank RTD | Code: BBKRTM.DE.YPPC1.PC100.R.I |
| consumer prices excl. food & energy, index ^{s,c} | cpi_core | prices | log, diff | Bundesbank RTD | Code: BBKRTM.DE.S.PPC1.PC110.R.I |
| producer prices, index ^s | ppl | prices | log, diff | Bundesbank RTD | Code: BBKRTM.DE.S.PPP1.PP100.R.I |
| producer prices excl. energy, index ^s | ppl_core | prices | log, diff | Bundesbank RTD | Code: BBKRTM.DE.S.PPP1.PP200.R.I |
| producer prices, construction, index | ppl constr | prices | log, diff | Bundesbank RTD | Code: BBKRTQ.DE.N.PPP1.PP300.R.I |
| producer prices, agriculture, index ^s | ppl_agri | prices | log, diff | Bundesbank RTD | Code: BBKRTQ.DE.N.PPP1.PP400.R.I |
| production, industry, index ^{s,c} | prod | activity | log, diff | Bundesbank RTD | Code: BBKRTM.DE.YI.IP1.ACM01.C.I |
| orders, industry, index ^{s,c} | ord | activity | log, diff | Bundesbank RTD | Code: BBKRTM.DE.YI.IO1.ACM01.C.I |
| turnover, industry, index ^{s,c} | to | activity | log, diff | Bundesbank RTD | Code: BBKRTM.DE.YI.IT1.ACM01.VI |
| production, construction, index ^{s,c} | prod | activity | log, diff | Bundesbank RTD | Code: BBKRTM.DE.YI.IP1.AA020.C.I |
| orders, construction, index ^{s,c} | ord | activity | log, diff | Bundesbank RTD | Code: BBKRTM.DE.YI.IO1.AA031.C.I |
| turnover, construction, index ^{s,c} | to | activity | log, diff | Bundesbank RTD | Code: BBKRTM.DE.YI.IT1.AA031.VA |
| turnover, retail, index ^{s,c} | to | activity | log, diff | Bundesbank RTD | price adjusted, Code: BBKRTM.DE.YI.IT1.AGA01.C.I |
| employment % ^c | emp | labor market | log, diff | Bundesbank RTD | Code: BBKRTQ.DE.SA.BFI.CA010.PA |
| hours worked, industry ^{s,c} | h_ind | labor market | log, diff | Bundesbank RTD | Code: BBKRTM.DE.YI.BE2.AA022.H.I |
| hours worked, construction ^{s,c} | h_constr | labor market | log, diff | Bundesbank RTD | Code: BBKRTM.DE.YI.BE2.AA031.H.A |
| wages and salaries ^{s,c} | w | labor market | log, diff | Bundesbank RTD | Code: BBKRTQ.DE.SA.DE2.CA010.VA |

| CDAX_index | cdax | financial | log_diff | Bundesbank | Code: BBK01.WU001A |
|--|---------------|-----------|----------|------------------|--|
| Nominal effective exchange rate, narrow | neer19 | financial | log_diff | Bundesbank | Vis-a-vis 19 trading partners, code: BBEE1.M.18.AAA.XZE012.A.AABAN.M00 |
| Nominal effective exchange rate, broad | neer38 | financial | log_diff | Bundesbank | Vis-a-vis 38 trading partners, code: BBEE1.M.18.AAA.XZE021.A.AABAN.M00 |
| Government bond yields (1-year) | ily | financial | diff | Bundesbank | Code: BBK01.WZ9808 |
| Government bond yields (5-year) | ily | financial | diff | Bundesbank | Code: BBK01.WZ9816 |
| Government bond yields (10-year) | ily | financial | diff | Bundesbank | Code: BBK01.WZ9826 |
| Commodity prices, energy, index | comm_en | financial | log_diff | HWWI | in €, sourced via Deutsche Bundesbank (BBDG1.M.HWWI.N.EURO.ENERGY00.I15.EUR.A) |
| Commodity prices, excl. energy, index | comm_exen | financial | log_diff | HWWI | in €, sourced via Deutsche Bundesbank (BBDG1.M.HWWI.N.EURO.TOTNXNGY.I15.EUR.A) |
| Industrial confidence indicator ^s | survey_ind | surveys | diff | Eurp. Commission | — |
| services confidence indicator ^s | survey_serv | surveys | diff | Eurp. Commission | — |
| consumer confidence indicator ^s | survey_cond | surveys | diff | Eurp. Commission | — |
| retail trade confidence indicator ^s | survey_retail | surveys | diff | Eurp. Commission | — |
| construction confidence indicator ^s | survey_constr | surveys | diff | Eurp. Commission | — |
| employment expectation index ^s | survey_eei | surveys | diff | Eurp. Commission | — |

^s seasonally adjusted

^c calendar adjusted

7 Gibbs Sampler

This Appendix details how to sample from the posterior distribution of the parameters of the factor model, $p(\Theta|\mathbf{y})$. For the reader's convenience, I repeat the model description before discussing the blocks of the Gibbs Sampler. Setting $J = 1$, $P = 2$ and $S = 0$, the following equations constitute the model:

$$\mathbf{y}_t = \lambda \mathbf{f}_t + \mathbf{e}_t \quad (7a)$$

$$\mathbf{f}_t = \phi_1 \mathbf{f}_{t-1} + \phi_2 \mathbf{f}_{t-2} + \mathbf{v}_t \quad (7b)$$

$$\mathbf{e}_t = \psi \mathbf{e}_{t-1} + \varepsilon_t \quad (7c)$$

The Gibbs Sampler then cycles through the following blocks to draw from the conditional posterior distributions⁸:

Step 1: $p(\mathbf{f}|\mathbf{y}, \Theta)$

Conditional on the parameters, the factors are drawn from their posterior distribution following the precision sampler outlined in Section 2.1 of the main text.

Step 2: $p(\lambda|\mathbf{y}, \mathbf{f}, \psi, \Sigma)$

Given the factors, the observation equations of the model are independent and the loadings can be sampled variable by variable. Let the prior on the loadings be given by $\mathcal{N}(0, \mathbf{P}_\lambda^{-1})$ and let $\mathbf{P}_{e_i} = (\mathbf{H}_\psi^\top \sigma_i^{-2} \mathbf{H}_\psi)^{-1}$ denote the $T \times T$ precision matrix of \mathbf{e}_i , where

$$\begin{bmatrix} I_N & & & & \\ -\psi_i & I_N & & & \\ & -\psi_i & I_N & & \\ & & & \dots & \dots & -\psi_i & I_N \\ & & & & & \dots & \dots & -\psi_i & I_N \end{bmatrix}$$

Using standard linear regression results (see Kroese and Chan, 2013), the conditional posterior for the i -th equation is given by

$$\lambda_i | \cdot \sim \mathcal{N}(\mathbf{m}_i^\lambda, \mathbf{M}_i^\lambda) \quad (8)$$

$$\mathbf{M}_i^\lambda = (\mathbf{P}_\lambda + \mathbf{f}^\top \mathbf{P}_{e_i} \mathbf{f})^{-1} \quad (9)$$

$$\mathbf{m}_i^\lambda = \mathbf{M}_i^\lambda (\mathbf{f}^\top \mathbf{P}_{e_i} \mathbf{y}_i). \quad (10)$$

We set $\mathbf{P}_\psi = \mathbf{I}_R$ for all variables.

⁸Initial values for the coefficients Θ are obtained from principal component estimates of the factors.

Step 3: $p(\psi|\mathbf{y}, \mathbf{f}, \lambda, \Sigma)$

Conditional on a draw of the loadings, for each series we can back out the idiosyncratic components \mathbf{e} and update ψ . Given a Normal prior for ψ_i with mean 0 and precision matrix P_ψ , the resulting posterior is given by

$$\psi_{i,\cdot}|\cdot \sim \mathcal{N}(m_i^\psi, M_i^\psi) \quad (11)$$

$$M_i^\psi = \left(P_\psi + \frac{1}{\sigma_i^2} \sum_{t=1}^{T-1} e_{i,t}^2 \right)^{-1} \quad (12)$$

$$m_i^\psi = M_i^\psi \left(\frac{1}{\sigma_i^2} \sum_{t=2}^T e'_{i,t-1} e_{i,t} \right). \quad (13)$$

For all variables, we set $P_\psi = 1$.

Step 4: $p(\Sigma|\mathbf{y}, \mathbf{f}, \lambda, \psi)$

Each diagonal elements of Σ can be sampled by drawing from

$$\sigma_i^2 \sim \mathcal{G}^{-1} \left(\frac{u + T - 1}{2}, \frac{U + \sum_{t=2}^T \varepsilon_{i,t}^2}{2} \right)$$

where u and U are the prior shape and rate. We set $u = 3$, $U = 0.5$ so that the prior mean and standard deviation equal 0.25.

Step 5: $p(\phi_1, \phi_2|\mathbf{f}, \Omega)$

Let $\mathbf{X}_t = I_R \otimes [\mathbf{f}_{t-1}^\top, \mathbf{f}_{t-2}^\top]$ denote the lags of the factors and $\mathbf{X} = [\mathbf{X}_1, \dots, \mathbf{X}_T]$. Then given a Normal prior $p(\phi) \sim \mathcal{N}(0, \mathbf{P}_\phi^{-1})$, the coefficients of the factor VAR can be sampled by drawing from $\phi|\cdot \sim \mathcal{N}(\mathbf{m}^\phi, \mathbf{M}^\phi)$ where

$$\mathbf{M}^\phi = (\mathbf{P}_\phi + \mathbf{X}^\top (\mathbf{I}_T \otimes \Omega^{-1}) \mathbf{X})^{-1}, \quad \mathbf{m}^\phi = \mathbf{M}^\phi (\mathbf{X}^\top (\mathbf{I}_T \otimes \Omega^{-1}) \mathbf{f})$$

In line with the literature on Bayesian vectorautoregressions, we impose a Minnesota-type prior on the VAR coefficients. The prior mean is set equal to 0 and the prior precision increases in the lag length. Specifically, the prior variance of the elements of ϕ for $p = 1 : P$ is given by:

$$\text{Var}(\phi_{p,ij}) = \begin{cases} \frac{\pi_0}{p^2}, & \text{if } i = j \\ \frac{\pi_1}{p^2}, & \text{otherwise.} \end{cases}$$

Common values in the literature are $\pi_0 = 0.2$ and $\pi_1 = 0.1$, thus shrinking coefficients on the lags of other factors stronger towards 0.

8 Computational efficiency analysis

This Appendix compares the computational efficiency of the precision-sampling algorithms outlined in the main text to simulation smoothers such as Carter and Kohn (1994) and Durbin and Koopman (2002) that rely on the Kalman-filter and are commonly used in the literature. Previous work in this area has analyzed the computational advantages of such simulation smoothers compared to the algorithms for conditional forecasting given in Waggoner and Zha (1999) in the context of a vector autoregressions (Banbura et al., 2015); McCausland et al. (2011) analyzes the performance of precision-based samplers in the context of a time-varying parameter regression and dynamic factor models and compares the performance to the Durbin and Koopman (2002) simulation smoother.

The focus lies on both small factor models with $N = 20$ variables and $R = 2$ factors as well as two variants of larger factor models with $N = 100, R = 2$ (*large N factor model*) and $N = 100, R = 2$ (*large factor model*). We evaluate the runtime of obtaining 100 draws from the predictive density $p(y^f | y^o, y^c)$ given a sample size $T = 100$ and forecast horizons $H = \{5, 20, 50\}$ conditional on a share $\kappa = 0.1$ of the observables (*hard conditioning*).⁹ Furthermore, we also consider the costs of producing repeated samples from the predictive density given the parameters as would be the case when on top of the conditioning set

$$y^c = [y_{1,T+1}, \dots, y_{\kappa \cdot N, T+1}, \dots, y_{1, T+H}, \dots, y_{\kappa \cdot N, T+H}]$$

restrictions are imposed on the remaining observables to lie within a prespecified range (*soft conditioning*).

The dynamics of the factor model outlined in Section 2 are given by a VAR(1) with parameters $\phi_1 = 0.7 \cdot I_R$ and $\Sigma = I_R$. The elements of the loadings on the contemporaneous factors λ_0 are drawn from independent Normal distributions with mean 0.5 and variance 0.1; all remaining loadings are set to 0, i.e. $K = 0$. Setting $J = 0$, the idiosyncratic components are i.i.d and ω_i is chosen such that the common component - $\lambda_{0,i} \mathbf{f}_t$ - explains $\frac{2}{3}$ of the total variation in the i -th variable.

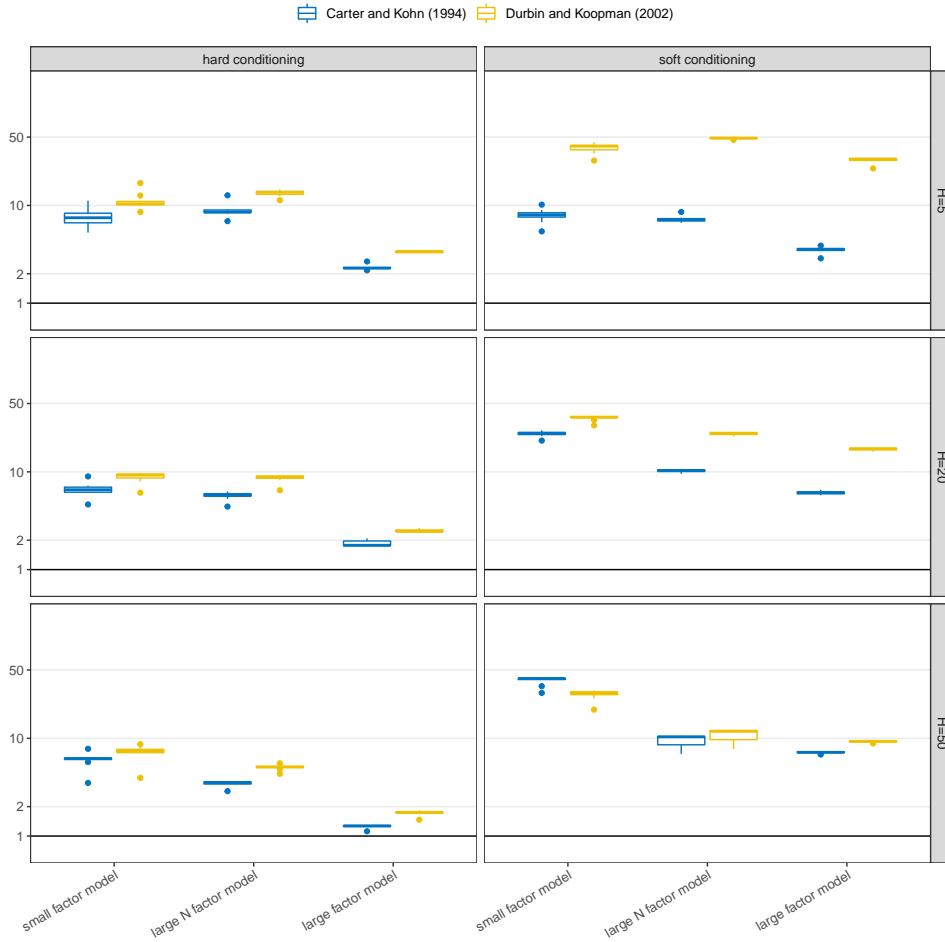
Details on the implementation of the Carter and Kohn (1994) and Durbin and Koopman (2002) simulation smoothers are provided in Appendix 9.

Each model is simulated 10 times.¹⁰ Figure 7 shows the relative runtime for hard (left column) and soft (right column) conditioning as boxplots. The precision-based algorithms performs compared to the Kalman-filter based simulation smoothers for all horizons, with gains in computational efficiency up to a factor of 10. When repeated draws from the predictive density are required as in the case of soft conditioning, the gains in computational efficiency are much larger.

⁹The size of the conditioning set does not have a large impact on the relative runtime and in the interest of space I do not report results for $\kappa = \{0.5, 0.75\}$. They are available upon request.

¹⁰The simulations were run using Matlab2020a on an Intel Core i5 2.3 Ghz with 8 GB of RAM. To time the execution of the code, we used the function `timeit()`.

Figure 7: Computational efficiency analysis of different simulation smoothers



Note: The figure shows the time it takes to generate draws from the predictive density $p(\mathbf{y}^f | \mathbf{y}^o, \mathbf{y}^c)$ for different forecast horizons H using the Carter and Kohn (1994) and Durbin and Koopman (2002) simulation smoothers relative to the precision sampler outlined in the main text. The conditioning set \mathbf{y}^c consists of the first $0.1 \cdot N$ variables (left column). For the small factor model, the number of variables is $N = 20$ and the number of static factors is $R = 2$; for the large N factor model, the number of variables is $N = 100$, while the large factor model also has $R = 10$. The runtime measures the entire procedure of generating a draw from the predictive density, including building the system matrices. For the precision-sampler this includes the costs of setting up the permutation matrices which is incurred only once. In the case of soft conditioning (right column), the runtime of producing 100 draws given the same parameters is reported. For details, see the main text.

9 Kalman-filter based simulation smoothers

This appendix provides details on the Carter and Kohn (1994) and Durbin and Koopman (2002) simulation smoothers that are used in the simulations in Appendix 8. Although both the Carter and Kohn (1994) and Durbin and Koopman (2002) simulation smoothers rely on the Kalman filter and smoother to produce a draw from the conditional distribution of the states, there are important conceptual differences between the two.

Let $a_{t|s} = E[\alpha_t | \mathbf{y}_1, \dots, \mathbf{y}_s]$ and $P_{t|s} = \text{Var}[\alpha_t | \mathbf{y}_1, \dots, \mathbf{y}_s]$ denote the conditional mean and variance of the state at time t conditional on information up to time s which can be obtained from the Kalman filter or smoother. In addition, let a_t denote a draw from the conditional distribution $p(\alpha_t | \mathbf{y}_1, \dots, \mathbf{y}_{T+H})$.

Carter and Kohn (1994) simulation smoother

Given the parameters of the state space model Θ and $a_{t|t}, P_{t|t} \forall t = 1, \dots, T+H$, the algorithm in Carter and Kohn (1994) generates a_{T+H} from

$$p(\alpha_{T+H} | \mathbf{y}_1, \dots, \mathbf{y}_{T+H})$$

with

$$\begin{aligned} E[\alpha_{T+H} | \mathbf{y}_1, \dots, \mathbf{y}_{T+H}] &= a_{T+H|T+H} \\ \text{Var}[\alpha_{T+H} | \mathbf{y}_1, \dots, \mathbf{y}_{T+H}] &= P_{T+H|T+H}. \end{aligned}$$

For $t = T+H-1, \dots, 1$ a_t is generated from $p(\alpha_t | \mathbf{y}_t, a_{t+1})$. For details on how to derive the moments of the conditional distributions given the output of the Kalman filter, see the original paper as well as the textbook treatment in Kim and Nelson (1999).

Durbin and Koopman (2002) simulation smoother

The Durbin and Koopman (2002) simulation smoother produces a draw from the conditional distribution of the states by first simulating the state space model in Equation 5, producing a draw from the *joint* distribution of states and observables

$$p(\alpha_1, \dots, \alpha_{T+H}, \mathbf{y}_1, \dots, \mathbf{y}_{T+H}).$$

Denote this joint draw by $a_1^+, \dots, a_{T+H}^+, y_1^+, \dots, y_{T+H}^+$. Running the Kalman smoother recursions yields $a_{t|T+H}$ and $a_{t|T+H}^+$ where $a_{t|T+H}^+ = E[\alpha_t^+ | y_1^+, \dots, y_{T+H}^+]$. By setting $a_t = a_{t|T+H} - a_{t|T+H}^+ + a_t^+$ adjusts the mean of a_t^+ to yield a draw from the desired conditional distribution. In practice, it is more efficient to first construct $y^* = y - y^+$ and run the Kalman smoother only once, yielding $a_{t|T+H}^* = E[\alpha_t^* | y^*]$. A draw from $p(\alpha_1, \dots, \alpha_{T+H} | \mathbf{y}_1, \dots, \mathbf{y}_{T+H})$ is then obtained by setting $a_t = a_{t|T+H}^* + a_t^+$. Note that since only an estimate of the smoothed

mean of the states is required, more efficient smoothing recursions can be employed which do not require the storage of $a_{t|t}$ and $P_{t|t}$ (see Durbin and Koopman, 2002, ch. 4.4.2).

Drawing from the (conditional) predictive density $p(y^f | I, \Theta)$

Under the common assumption of uncorrelated measurement errors, draws from the predictive density $p(y_t^f | a_t, y^o, y^c, \Theta)$ can be obtained by sampling the measurement errors independently from $\mathcal{N}(0, \omega_i) \forall i$ such that $y_{i,t} \notin y_{c,t}$ and adding them to the draw of the state vector (see Banbura et al., 2015, for the general case of a full covariance matrix Ω). For the Durbin and Koopman (2002) simulation smoother the generated artificial observations y^+ can be re-used, since $Fa_{t|T+H}^* + y_t^+$ is a draw of the observations conditional on the sampled states.

Multiple draws from $p(y^f | I, \Theta)$

Additional draws from the (conditional) predictive density given the same set of parameters Θ can be produced recursively from the Carter and Kohn (1994) simulation smoother by independently sampling a_t as many times as required and conditioning on the draw to generate y_t^f . In the case of the Durbin and Koopman (2002) simulation smoother, each additional draw requires new simulated values of the states and observations, $a_1^+, \dots, a_{T+H}^+, y_1^+, \dots, y_{T+H}^+$. Given these, the entire filtering and smoothing recursions need to be re-run. However, computational savings still arise as the state covariances $P_{t|t}$ and functions thereof that are calculated during the recursions do not depend on the observations. As such they only need be calculated once and can be passed on to subsequent runs of the Kalman filter and smoother (Durbin and Koopman, 2002, p. 606).

Initialisation

The Kalman filter is initialized by setting $a_{1|0}$ and $P_{1|0}$ equal to the unconditional mean and variance, respectively, of the factor VAR.

References

- Banbura, Marta, Domenico Giannone, and Michele Lenza, “Conditional forecasts and scenario analysis with vector autoregressions for large cross-sections,” *International Journal of Forecasting*, 2015, 31 (3), 739–756.
- Bishop, Christopher M., *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Berlin, Heidelberg: Springer-Verlag, 2006.
- Bok, Brandyn, Daniele Caratelli, Domenico Giannone, Argia Sbordon, and Andrea Tambalotti, “Macroeconomic Nowcasting and Forecasting with Big Data,” *Annual Review of Economics*, 08 2018, 10.
- Carter, C. K. and R. Kohn, “On Gibbs sampling for state space models,” *Biometrika*, 1994, 81 (3), 541–553.
- Chan, Joshua C.C. and Ivan Jeliazkov, “Efficient simulation and integrated likelihood estimation in state space models,” *International Journal of Mathematical Modelling and Numerical Optimisation*, 2009, 1 (1-2), 101–120.
- Clark, Todd E. and Michael W. McCracken, “Tests of Predictive Ability for Vector Autoregressions Used for Conditional Forecasting,” *Journal of Applied Econometrics*, 2017, 32 (3), 533–553.
- Del Negro, Marco and Frank Schorfheide, “DSGE Model-Based Forecasting,” in Graham Elliott and Allan Timmermann, eds., *Handbook of Economic Forecasting*, Vol. 2 of *Handbook of Economic Forecasting*, Elsevier, 2013, pp. 57–140.
- Diebold, Francis X. and Roberto S. Mariano, “Comparing Predictive Accuracy,” *Journal of Business & Economic Statistics*, July 1995, 13 (3), 253–263.
- Durbin, James and Siem Jan Koopman, “A simple and efficient simulation smoother for state space time series analysis,” *Biometrika*, 2002, 89 (3), 603–616.
- Faust, Jon and Jonathan H. Wright, “Comparing Greenbook and Reduced Form Forecasts Using a Large Realtime Dataset,” *Journal of Business & Economic Statistics*, 2009, 27 (4), 468–479.
- Ganics, Gergely and Florens Odendahl, “Bayesian VAR forecasts, survey information, and structural change in the euro area,” *International Journal of Forecasting*, 2021, 37 (2), 971–999.
- Geweke, John and Charles Whiteman, “Bayesian Forecasting,” in G. Elliott, C.W.J. Granger, and A. Timmermann, eds., *Handbook of Economic Forecasting*, Vol. 1, Elsevier, 2006, pp. 3–80.

- Giannone, Domenico, Lucrezia Reichlin, and David Small, “Nowcasting: The real-time informational content of macroeconomic data,” *Journal of Monetary Economics*, 2008, 55 (4), 665–676.
- , Michele Lenza, Daphne Momferatou, and Luca Onorante, “Short-term inflation projections: A Bayesian vector autoregressive approach,” *International Journal of Forecasting*, 2014, 30 (3), 635–644.
- Golub, Gene H. and Charles F. Van Loan, *Matrix Computations* Johns Hopkins Studies in the Mathematical Sciences, 4th ed., Johns Hopkins University Press, 2013.
- Hauber, Philipp and Christian Schumacher, “Precision-based sampling with missing observations: A factor model application,” Discussion Papers 11/2021, Deutsche Bundesbank 2021.
- Jordan, Alexander, Fabian Krüger, and Sebastian Lerch, “Evaluating Probabilistic Forecasts with scoringRules,” *Journal of Statistical Software, Articles*, 2019, 90 (12), 1–37.
- Kaufmann, Sylvia and Christian Schumacher, “Identifying relevant and irrelevant variables in sparse factor models,” *Journal of Applied Econometrics*, 2017, 32 (6), 1123–1144.
- and —, “Bayesian estimation of sparse dynamic factor models with order-independent and ex-post mode identification,” *Journal of Econometrics*, 2019, 210 (1), 116 – 134.
- Kim, Chang-Jin and Charles R. Nelson, *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*, Vol. 1 of *MIT Press Books*, The MIT Press, September 1999.
- Knotek, Edward S. and Saeed Zaman, “Financial nowcasts and their usefulness in macroeconomic forecasting,” *International Journal of Forecasting*, 2019, 35 (4), 1708–1724.
- Krüger, Fabian, Todd E. Clark, and Francesco Ravazzolo, “Using Entropic Tilting to Combine BVAR Forecasts With External Nowcasts,” *Journal of Business & Economic Statistics*, 2017, 35 (3), 470–485.
- Kroese, Dirk P. and Joshua C.C. Chan, *Statistical Modeling and Computation*, Springer Publishing Company, Incorporated, 2013.
- Krüger, Fabian, Sebastian Lerch, Thordis L. Thorarinsdottir, and Tilmann Gneiting, “Predictive Inference Based on Markov Chain Monte Carlo Output,” *arXiv e-prints*, Aug 2016, p. arXiv:1608.06802.
- Kuzin, Vladimir, Massimiliano Marcellino, and Christian Schumacher, “Pooling versus model selection for nowcasting GDP with many predictors: Empirical evidence for six industrialized countries,” *Journal of Applied Econometrics*, 2013, 28 (3), 392–411.

- McCausland, William J., “The HESSIAN method: Highly efficient simulation smoothing, in a nutshell,” *Journal of Econometrics*, 2012, 168 (2), 189–206.
- , “Dynamic Factor Models with stochastic volatility,” 2015. University of Montreal, Department of Economics, mimeo.
- , Shirley Miller, and Denis Pelletier, “Simulation smoothing for state-space models: A computational efficiency analysis,” *Computational Statistics & Data Analysis*, 2011, 55 (1), 199–212.
- Rue, H. and L. Held, *Gaussian Markov Random Fields: Theory and Applications* Chapman & Hall/CRC Monographs on Statistics & Applied Probability, CRC Press, 2005.
- Rue, Håvard, “Fast sampling of Gaussian Markov random fields,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2001, 63 (2), 325–338.
- Schumacher, Christian and Jörg Breitung, “Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data,” *International Journal of Forecasting*, 2008, 24 (3), 386–398.
- Stock, James H. and Mark W. Watson, “Macroeconomic forecasting using diffusion indexes,” *Journal of Business & Economic Statistics*, 2002, 20 (2), 147–162.
- and —, “Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics,” in John. B. Taylor and Harald Uhlig, eds., *Handbook of Macroeconomics*, Vol. 2 of *Handbook of Macroeconomics*, Elsevier, 2016, chapter 0, pp. 415–525.
- Tallman, Ellis W. and Saeed Zaman, “Combining survey long-run forecasts and nowcasts with BVAR forecasts using relative entropy,” *International Journal of Forecasting*, 2020, 36 (2), 373–398.
- Waggoner, Daniel F. and Tao Zha, “Conditional Forecasts in Dynamic Multivariate Models,” *The Review of Economics and Statistics*, 1999, 81 (4), 639–651.
- Wolters, Maik H., “Evaluating Point and Density Forecasts of DSGE Models,” *Journal of Applied Econometrics*, 2015, 30 (1), 74–96.