Finanziato dall'Unione europea NextGenerationEU

Ministero dell'Università e della Ricerca

Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA

Finanziato nell'ambito del Piano Nazionale di Ripresa e Resilienza PNRR. Missione 4, Componente 2, Investimento 1.3 Creazione di "Partenariati estesi alle università, ai centri di ricerca, alle aziende per il finanziamento di progetti di ricerca di base"

GRINS FOUNDATION

WP 4 – Public debt and the financial system

Predicting Tail-Risks for the Italian Economy

M. Boeck‡, M. Marcellino‡, M. Pfarrhoffer†, T. Tornese‡

‡ Università Bocconi

† WU Vienna

Document data	
Title	Spoke 4 Work Package 4 Public debt and the financial system under compounding risks
Last version date	24/05/2024

Executive summary

This paper investigates the empirical performance of various econometrics methods for predicting tail risks for the Italian economy. It provides an overview of the recent econometric methods to assess tail risks, including Bayesian VARs with stochastic volatility (BVAR-SV), Bayesian additive regression trees (BART), and Gaussian processes (GP). In an out-of-sample forecasting exercise for the Italian economy, the paper assesses the point, density, and tail predictive performance for GDP growth, inflation, debt-to-GDP, and deficit-to-GDP ratios. It turns out that BVAR-SV performs particularly well for Italy, in particular for the tails. It is then used to also predict expected shortfalls and longrises for the variables of interest, and the probability of specific interesting events, such as negative growth, inflation above the 2% target, an increase in debt-to-GDP ratio, or a deficit-to-GDP ratio above 3%.

TABLE OF CONTENTS

1. Introduction **2**

2. Econometric Framework **4**

2.1. Nested model variants 5

2.2. Priors and posteriors 7

2.3. Predictive density 8

2.4. Forecast evaluation metrics 9

3. Data **10**

4. Results **11**

4.1. Quarterly forecasting results 11

4.2. Monthly forecasting results 20

5. Concluding Remarks **22**

A. References 25

B. Data 29

C. Additional Results 31

Predicting Tail-Risks for the Italian Economy*

Maximilian Boeck[1], Massimiliano Marcellino[2], Michael Pfarrhofer[3], and Tommaso Tornese[1]

[1] Bocconi University
[2] Bocconi University, IGIER, CEPR, Baffi-Carefin and BIDSA
[3] Vienna University of Economics and Business

May 2024

Abstract

This paper investigates the empirical performance of various econometrics methods for predicting tail risks for the Italian economy. It provides an overview of the recent econometric methods to assess tail risks, including Bayesian VARs with stochastic volatility (BVAR-SV), Bayesian additive regression trees (BART), and Gaussian processes (GP). In an out-of-sample forecasting exercise for the Italian economy, the paper assesses the point, density, and tail predictive performance for GDP growth, inflation, debt-to-GDP, and deficit-to-GDP ratios. It turns out that BVAR-SV performs particularly well for Italy, in particular for the tails. It is then used to also predict expected shortfalls and longrises for the variables of interest, and the probability of specific interesting events, such as negative growth, inflation above the 2% target, an increase in debt-to-GDP ratio, or a deficit-to-GDP ratio above 3%.

Keywords: Density forecasts, tail forecasts, Bayesian VAR, BART, Gaussian Process, Debt, Deficit, Italy

JEL Codes: C11, C32, C53.

*Maximilian Boeck, Università Bocconi, Department of Economics, Via Roentgen 1, 20136 Milano, Italy. E-mail: maximilian.boeck@unibocconi.it. Massimiliano Marcellino, Università Bocconi, Department of Economics, Via Roentgen 1, 20136 Milano, Italy. massimiliano.marcellino@unibocconi.it. Michael Pfarrhofer, Vienna University of Economics and Business, Department of Economics, Welthandelsplatz 1, 1020 Vienna, Austria. michael.pfarrhofer@wu.ac.at. Tommaso Tornese, Università Bocconi, Department of Economics, Via Roentgen 1, 20136 Milano, Italy. tommaso.tornese@unibocconi.it. We would like to thank Valentina Aprigliano and Francesco Corsello for helpful conversations about data and Todd Clark for useful comments on a previous version. This study was funded by the European Union - *NextGenerationEU*, in the framework of the *GRINS – Growing Resilient, INclusive and Sustainable project* (GRINS PE00000018 – CUP B43C22000760006). The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them.

1. Introduction

Both central banks and governments are increasingly concerned about tail risk. In a recent speech, for instance, ECB executive board member Isabel Schnabel highlighted the increased attention towards inflation risk (Schnabel, 2023). Governments faced with increasing public debts or risks thereof, on the other hand, are confronted with different policy tradeoffs (Cochrane, 2022; Reis, 2022). Projections of short- to medium-term macroeconomic and fiscal dynamics are thus key inputs in economic and monetary policymaking. To monitor potential risks and implement forward-looking measures, officials in governments and central banks require timely information about the effects of policy and non-policy shocks and subsequently induced business cycle fluctuations, and potential paths towards recovery in the aftermath of adverse shocks. This paper is concerned with providing projections on potential risks for the Italian economy, though the methods considered are of course of general applicability.

We evaluate a battery of multivariate econometric specifications to flexibly model and forecast interrelated dynamics in relevant macroeconomic, financial, and fiscal variables. Based on data for the Italian economy, we assess the usefulness of a range of different Bayesian linear and nonlinear econometric models to improve predictive accuracy for key variables such as GDP growth, inflation, unemployment, interest rates, deficit-to-GDP ratio, and debt-to-GDP ratio. We are interested both in traditional point forecast metrics, but also consider density and tail forecast performance.

The focus on density predictions, and especially tail forecast metrics, is due to the recent emphasis on predicting macroeconomic risk in the wake of Adrian, Boyarchenko and Giannone (2019). The most popular targeted objects, often with single-equation variants of quantile regression (QR), are the quantiles of GDP growth, which are commonly referred to as Growth-at-Risk (see, e.g., Adrian et al., 2022; Clark et al., 2024b; Plagborg-Møller et al., 2020). But there are other econometric approaches that have been shown to perform equally well or better than QR (see Carriero, Clark and Marcellino, 2022; Clark et al., 2023; Delle Monache, De Polis and Petrella, 2023, among others), and many papers have considered a wider set of economic indicators, such as inflation and unemployment (see, e.g., Adams et al., 2021; Galbraith and van Norden, 2019; Manzan, 2015; Pfarrhofer, 2022, among others).

A second important aspect that we consider is to address nonlinearities of potentially unknown form; nonlinearities are particularly relevant to capture the effects of extreme events such as the financial crisis or the Covid pandemic, but possibly also the time variation in economic relationships. Traditional implementations of nonlinearities (such as with discrete regime switches or drifting parameters in otherwise conditionally linear models, and stochastic volatilities) often offer gains in predictive accuracy (see, e.g., D'Agostino, Gambetti and Giannone, 2013; Hauzenberger et al., 2022, among many others). More recently, more flexible specifications that treat the conditional mean functions as unknown have been shown to work well in time series applications. For instance, Bayesian Additive Regression Tree (BART) models in multivariate contexts have been used successfully in Clark et al. (2023), Huber et al. (2023) and Huber and Rossini (2022); and Gaussian process (GP) regression is useful when applied to macroeconomic data (see, e.g., Clark et al., 2024a; Hauzenberger et al., 2024a,b), particularly to model tail events.

These considerations yield our full set of competing specifications, which is comprised of established workhorse models and also inspired by the recent literature on Bayesian machine learning methods in macroeconometrics. We model the joint dynamics of macroeconomic, financial, and fiscal variables in variants of parametric and nonparametric vector autoregressions. All variants of our models are considered with and without heteroskedastic errors, which have been shown to be a crucial feature when the aim is to improve density forecasts (see, e.g., Carriero, Clark and Marcellino, 2016; Chan, 2022; Clark, 2011). Moreover, Carriero, Clark and Marcellino (2024) show that BVAR-SVs generally produce quantile forecasts that are very similar to those resulting from quantile regressions, making the latter overall redundant. Hence, in our evaluation we consider standard Bayesian VARs, VARs with time-varying parameters (TVP-VARs), BART-VARs, and GP-VARs, where an -SV suffix indicates the same models complemented with stochastic volatility. These specifications are applied to model small and larger sets of variables (ranging from 5 to 25) taken from two datasets for Italy, one monthly and one quarterly, in order to assess whether including more information in the models and a higher sampling frequency matter, covering the period 2005–2023. We then compute and evaluate point, density and tail forecasts from all the models and for all variables, for forecast horizons from one-month/quarter up to two-year ahead.

Our empirical results indicate that none of the more sophisticated models we consider beats systematically and/or by substantial amounts the *medium*-size BVAR-SV in point, density and tail predictions, though some models are better for specific variables, metrics, and horizons. For example, left tail forecasts for growth produced by BART and GP at the two-year horizon lead to about 6% gains according to CRPS-L, but the same models produce losses at the other horizons and in overall density forecasts.

It is also worth noting that SV is very helpful for the linear BVAR models, as the forecasting performance of a homoskedastic BVAR is generally much worse, in line with the cited literature. Instead, SV is much less important for the nonlinear specifications, due to their more flexible specification for the conditional mean. Also, for both the BVAR-SV and the non-parametric specifications, working with larger models does not systematically lead to better performance for our variables.

We also inspect expected shortfall (ES) and longrise (LR) as risk measures. They are used to characterize downside and upside risks and are also called conditional value at risk, as they provide the expected value of a variable conditioning on the event that its realization is in the left (ES) or right (LR) tail. Finally, we use the models to compute predictions of the probability of specific risk events, such as negative growth, inflation above the ECB target, an increase in the debt-to-GDP ratio, or a deficit-to-GDP ratio above 3%.

When we consider the outcomes of the ES/LR and event probability analyses, we find that overall the outcomes supports the use of the BVAR-SV for monitoring economic and fiscal variables for Italy, also when interest focuses on tail risks, but we have provided a set of alternative and more sophisticated econometric tools that can be used for comparison with the BVAR-SV, in particular when tail risks are of interest.

The remainder of the paper is structured as follows. In Section 2, we provide an overview of the econometric models, estimation procedures and evaluation criteria used in the forecasting exercise. In Section 3, we describe the quarterly and monthly datasets involved in the forecasting exercise. In Section 4 we discuss the results, and in Section 5 we summarize the main findings and conclude. The Appendix contains additional details.

2. Econometric Framework

Our set of competing models is comprised of established multivariate workhorse models and flexible extensions of them. These extensions are inspired by the recent literature on Bayesian machine learning methods in macroeconometrics. Below, we provide an overview and refer the technically-inclined reader to Marcellino and Pfarrhofer (2024) for details and examples. Let $\boldsymbol{y}_t = (y_{1t}, \dots, y_{nt})'$ be a de-meaned $n \times 1$ vector of target variables, and $\boldsymbol{x}_t = (\boldsymbol{y}'_{t-1}, \dots, \boldsymbol{y}'_{t-p})'$ a

$k \times 1$ vector which collects lagged endogenous variables with $k = np$:

$$\boldsymbol{y}_t = F(\boldsymbol{x}_t) + \boldsymbol{\epsilon}_t. \tag{2.1}$$

We assume equation-specific unknown conditional mean functions $f_i(\boldsymbol{x}_t) : \mathbb{R}^k \to \mathbb{R}$ for $i = 1, \ldots, n$, that comprise $F(\boldsymbol{x}_t) = (f_1(\boldsymbol{x}_t), \ldots, f_n(\boldsymbol{x}_t))'$. The reduced form errors are stacked in $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \ldots, \epsilon_{nt})'$. To enable computationally and statistically efficient estimation procedure, we impose a particular structure on these errors:

$$\begin{aligned}
&\boldsymbol{\epsilon}_t = \boldsymbol{L}\boldsymbol{\mathfrak{F}}_t + \boldsymbol{\eta}_t, \quad \boldsymbol{\mathfrak{F}}_t \sim \mathcal{N}(\boldsymbol{0}_q, \boldsymbol{\Omega}_t), \quad \boldsymbol{\eta}_t \sim \mathcal{N}(\boldsymbol{0}_n, \boldsymbol{H}_t), \qquad (2.2)\\
&\boldsymbol{\Omega}_t = \operatorname{diag}\left(\exp(\omega_{1t}), \ldots, \exp(\omega_{qt})\right), \quad \boldsymbol{H}_t = \operatorname{diag}\left(\exp(h_{1t}), \ldots, \exp(h_{nt})\right).
\end{aligned}$$

This partitions each error into a common and an idiosyncratic component, and establishes a factor stochastic volatility (FSV, see, e.g., Kastner and Huber, 2020, in a VAR-context) model, which offers two key advantages. First, the common component determined by the $n \times q$ loadings matrix $\boldsymbol{L}$ and the $q \times 1$ vector of common static factors $\boldsymbol{\mathfrak{F}}_t = (\mathfrak{f}_{1t}, \ldots, \mathfrak{f}_{qt})'$ reflects the feature that economic variables and their volatilities exhibit significant co-movement over the business cycle, while the idiosyncratic component captures variable specific shocks $\boldsymbol{\eta}_t = (\eta_{1t}, \ldots, \eta_{nt})'$. Second, conditional on the common factors, this error specification implies independent equations when computing $\tilde{\boldsymbol{y}}_t \equiv \boldsymbol{y}_t - \boldsymbol{L}\boldsymbol{\mathfrak{F}}_t = F(\boldsymbol{x}_t) + \boldsymbol{\eta}_t$, which we leverage to set up an order-invariant equation-by-equation sampling algorithm. Notice that the covariance matrix is given by $\operatorname{Var}(\boldsymbol{\epsilon}_t) \equiv \boldsymbol{\Sigma}_t = \boldsymbol{L}\boldsymbol{\Omega}_t\boldsymbol{L}' + \boldsymbol{H}_t$, and the diagonal structure of $\boldsymbol{H}_t$ imposes the said independence.

2.1 Nested model variants

Established multivariate time series models can easily be obtained by assuming linear functional relationships (e.g., with constant or time varying parameters). This collapses the potentially high-dimensional and general model of Eq. (2.1) into plain linear regressions:

$$f_i(\boldsymbol{x}_t) = \boldsymbol{x}_t'\boldsymbol{a}_{it} \tag{2.3}$$

where $\boldsymbol{a}_{it}$ is a $k \times 1$-vector of equation-specific VAR coefficients. We stack the coefficients column-wise in a vector $\boldsymbol{a}_t$ and assume a random walk state equation:

$$\boldsymbol{a}_t = \boldsymbol{a}_{t-1} + \boldsymbol{\Theta}^{1/2}\boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\boldsymbol{0}_{nk}, \boldsymbol{I}_{nk}), \quad \boldsymbol{\Theta}^{1/2} = \operatorname{diag}\left(\theta_1^{1/2}, \ldots, \theta_{nk}^{1/2}\right).$$

This notation refers to the TVP-VAR case, but assuming the parameters to be constant (by setting $\boldsymbol{\Theta}^{1/2} = \boldsymbol{0}_{nk}$) results in an otherwise conventional BVAR.

The first nonparametric variant arises when we approximate the functions $f_i(\boldsymbol{x}_t)$ with additive regression trees. This approach is labeled BART, following Chipman, George and McCulloch (2010), and for tree functions $\ell_{is}(\boldsymbol{x}_t|\mathcal{T}_{is}, \boldsymbol{\mu}_{is})$ we have:

$$f_i(\boldsymbol{x}_t) \approx \sum_{s=1}^{S} \ell_{is}(\boldsymbol{x}_t|\mathcal{T}_{is}, \boldsymbol{\mu}_{is}), \tag{2.4}$$

with $\mathcal{T}_{is}$ being tree structures with associated terminal node parameters (the "leaves" or fitted values) of the tree collected in $\boldsymbol{\mu}_{is}$. And we have a sum of S of these trees, where each individual tree is pruned with our prior setup. In the machine learning jargon, the individual trees are "weak learners," but their ensemble may approximate very general functional relationships. The resulting multivariate model is labeled BART-VAR.

Another approach is to use Gaussian processes to infer the unknown functions. This can be achieved by imposing priors directly on the functional relationship:

$$f_i(\boldsymbol{x}_t) \sim \mathcal{GP}\left(0, \mathcal{K}_{\boldsymbol{\vartheta}_i}(\boldsymbol{x}_t, \boldsymbol{x}_t)\right). \tag{2.5}$$

This prior is defined with a suitable distance-based covariance function $\mathcal{K}_{\boldsymbol{\vartheta}_i}$ called the kernel, governed by a low-dimensional set of tuning parameters $\boldsymbol{\vartheta}_i$. When we condition on a finite sample, this specification translates to a multivariate Gaussian distribution on the function values, which has several favorable statistical and computational properties. We rely on equation-specific squared exponential kernels with two hyperparameters, the unconditional variance and the inverse length-scale. The resulting model is labeled GP or GP-VAR.

In terms of the variance components, we consider both homoskedastic and heteroskedastic specifications. The homoskedastic ones trivially arise when setting $\omega_{jt} = 0$ for $j = 1, \dots, q$, and $h_{it} = h_i$ for $i = 1, \dots, n$, and all $t = 1, \dots, T$. For heteroskedastic ones, we use independent stochastic volatility specifications:

$$\begin{aligned} h_{it} &= \mu_{ih} + \phi_{ih}(h_{it-1} - \mu_{ih}) + \varsigma_{ih}\zeta_{it,h}, \quad \zeta_{it,h} \sim \mathcal{N}(0,1), \quad \text{for } i = 1, \dots, n, \\ \omega_{jt} &= \phi_{j\omega}\omega_{jt-1} + \varsigma_{j\omega}\zeta_{jt,\omega}, \quad \zeta_{jt,\omega} \sim \mathcal{N}(0,1), \quad \text{for } j = 1, \dots, q. \end{aligned} \tag{2.6}$$

In some cases, we also consider t-distributed idiosyncratic errors with ν_i degrees of freedom which we estimate from the data, $\eta_{it} \sim t_{\nu_i}(0, \exp(h_{it}))$, as this can help in particular the linear models to better fit outliers and other data irregularities, see e.g., Carriero et al. (2022).

We refer to the three alternative specifications for the variance of the errors as hom, SV, and SV-t, and these terms are added to the name of the specification of the model for the conditional mean, e.g. GP-SV or BART-SV.

Table 1: Model Overview.

Conditional mean models	
`BVAR`	$f_i(\boldsymbol{x}_t) = \boldsymbol{x}_t'\boldsymbol{a}_i$
`TVP-VAR`	$f_i(\boldsymbol{x}_t) = \boldsymbol{x}_t'\boldsymbol{a}_{it}$
`BART-VAR`	$f_i(\boldsymbol{x}_t) \approx \sum_{s=1}^{S} \ell_{is}(\boldsymbol{x}_t \vert \mathcal{T}_{is}, \boldsymbol{\mu}_{is})$ with BART
`GP-VAR`	$f_i(\boldsymbol{x}_t) \sim \mathcal{GP}\left(0, \mathcal{K}_{\boldsymbol{\vartheta}_i}(\boldsymbol{x}_t, \boldsymbol{x}_t)\right)$ with GP prior
Conditional variance models	
`hom`	homoskedastic model $\boldsymbol{H}_t = \boldsymbol{H}$ and $\boldsymbol{\Omega}_t = \boldsymbol{I}_q$
`sv`	stochastic volatility, AR(1) with Gaussian errors
`sv-t`	stochastic volatility, AR(1) with t-distributed errors

Note: In the general from, the model is $\boldsymbol{y}_t = F(\boldsymbol{x}_t) + \boldsymbol{\epsilon}_t$, $\boldsymbol{\epsilon}_t \sim \mathcal{N}(\boldsymbol{0}_n, \boldsymbol{\Sigma}_t)$ where $F(\boldsymbol{x}_t) = (f_1(\boldsymbol{x}_t), \dots, f_n(\boldsymbol{x}_t))'$. The covariance matrix is decomposed as $\boldsymbol{\Sigma}_t = \boldsymbol{L}\boldsymbol{\Omega}_t\boldsymbol{L}' + \boldsymbol{H}_t$.

2.2 Priors and Posteriors

For the linear versions of our model, we exploit the non-centered parameterization of Frühwirth-Schnatter and Wagner (2010) to define the following equation-specific priors on the regression coefficients:

$$\boldsymbol{a}_{it} = \boldsymbol{a}_i + \text{diag}(\boldsymbol{\theta}_i^{1/2})\tilde{\boldsymbol{a}}_{it}, \quad \tilde{\boldsymbol{a}}_{it} \sim \mathcal{N}(\tilde{\boldsymbol{a}}_{it-1}, \boldsymbol{I}_k), \quad \tilde{\boldsymbol{a}}_{i0} = \boldsymbol{0}_k,$$
$$\boldsymbol{a}_i \sim \mathcal{N}(\boldsymbol{0}_k, \tau_{ai}^2 \cdot \text{diag}(\lambda_{a,i1}^2, \dots, \lambda_{a,ik}^2)), \quad \boldsymbol{\theta}_i^{1/2} \sim \mathcal{N}(\boldsymbol{0}_k, \tau_{\theta i}^2 \cdot \text{diag}(\lambda_{\theta,i1}^2, \dots, \lambda_{\theta,ik}^2)).$$

The τ's are equation-specific global scaling parameters, while the λ's are local scalings, each of them equipped with half-Cauchy priors. This establishes the horseshoe (HS) prior, see Carvalho, Polson and Scott (2010). Note that the constant parameter version arises when setting $\boldsymbol{\theta}_i = \boldsymbol{0}_k$. This prior setup gives rise to several well-known conditional Gaussian posterior distributions, and forward-filtering backward sampling can be used to update the latent states. The hyperparameters of the horseshoe can be updated as in Makalic and Schmidt (2015). We use the same HS prior on the elements of the loadings matrix $\boldsymbol{L}$.

For the nonparametric methods BART and GP, we rely on the following prior setup. For BART we use on the default setup as suggested in Chipman, George and McCulloch (2010). We set $S = 250$ in line with the previous literature, and use symmetric Gaussian priors on the terminal node parameters. In addition, we impose uniform priors both over the splitting variables and the thresholds which determine the tree structures. It is worth noting that the trees are sampled marginally of the terminal node parameters using Metropolis-Hastings updates, which provides significant computational advantages. On the hyperparameters of the GP prior we assume independent Gamma distributed priors, which can be updated using Metropolis-Hastings steps. The GP posteriors take a standard, conditionally multivariate Gaussian, form.

For the SV state equation parameters, we use default priors as in Kastner and Huber (2020), which involves a Gaussian distribution for the unconditional mean, a Beta distributed prior on the autoregressive parameter and a Gamma prior on the variance. Further details on priors and posteriors can be found in Marcellino and Pfarrhofer (2024).

2.3 Predictive Density

The distribution of the one-step ahead forecast is given by:

$$p(\boldsymbol{y}_{T+1} \mid \boldsymbol{y}_{1:T}) = \int p(\boldsymbol{y}_{T+1} \mid \boldsymbol{y}_{1:T}, \boldsymbol{\Xi}) p(\boldsymbol{\Xi} \mid \boldsymbol{y}_{1:T}) d\boldsymbol{\Xi},$$

where $\boldsymbol{y}_{1:T}$ denotes the information set up to time T and $\boldsymbol{\Xi}$ contains all parameters and latent states. Due to the nonlinear/nonparametric setup, this distribution has no well-known form and we thus rely on simulation-based methods to generate draws.

Let m denote the iteration of the sampling algorithm such that $x^{(m)}$ indicates the mth draw for some random variable X. In each sweep of our MCMC algorithm, we may simulate from the known density:

$$\hat{\boldsymbol{y}}_{T+1|T}^{(m)} \sim p\left(\boldsymbol{y}_{T+1} \mid \boldsymbol{y}_{1:T}, \boldsymbol{\Xi}^{(m)}\right) = \mathcal{N}\left(F^{(m)}(\boldsymbol{x}_{T+1}), \boldsymbol{\Sigma}_{T+1|T}^{(m)}\right),$$

where $\boldsymbol{\Sigma}_{T+1|T}^{(m)}$ is the conditional volatility forecast based on Eq. (2.6). Higher-order predictions for $h \geq 2$ are then computed recursively by setting the predictors to $\boldsymbol{x}_{T+h|T}^{(m)} = (\hat{\boldsymbol{y}}_{T+h-1|T}^{(m)}{}', \hat{\boldsymbol{y}}_{T+h-2|T}^{(m)}{}', \ldots)'$, iterating forward in time:

$$\hat{\boldsymbol{y}}_{T+h|T}^{(m)} \sim p\left(\boldsymbol{y}_{T+h} \mid \boldsymbol{y}_{1:T}, \boldsymbol{\Xi}^{(m)}\right) = \mathcal{N}\left(F^{(m)}\left(\boldsymbol{x}_{T+h|T}^{(m)}\right), \boldsymbol{\Sigma}_{T+h|T}^{(m)}\right).$$

This procedure exploits the known Gaussian form of the conditional one-step ahead predictive distribution, while taking the recursive structure and nonlinearities into account, and integrates out the future forecast errors. This delivers draws from the desired predictive distributions $p(\boldsymbol{y}_{T+h} \mid \boldsymbol{y}_{1:T})$ for horizons $h = 1, 2, \ldots$, which may exhibit non-Gaussian features such as asymmetries, multiple modes or heavy tails. Predictive inference is then based on $M_{\text{save}} = M_{\text{total}} - M_{\text{burn}}$ draws $\hat{\boldsymbol{y}}_{T+h|T}^{(m)}$ where $m = M_{\text{burn}} + 1, \ldots, M_{\text{total}}$, and we discard M_{burn} draws as burn-in.

2.4 Forecast Evaluation Metrics

We now briefly discuss how we evaluate the predictive accuracy of forecasts, taking particularly the evaluation of tail risks into account. We consider a broad range of evaluation metrics: quantile scores (QS), expected shortfall (ES), longrise (LR), and quantile-weighted continuous ranked probability scores (qwCRPS, see also Gneiting and Raftery, 2007) with three weighting schemes, i.e., the conventional `CRPS`, a version to target the left tail, labeled as `qwCRPS-L`, and to target the right tail, labeled as `qwCRPS-R`.

We start by defining the quantile score (`QS`) because it can be used to construct several of our more general predictive loss metrics. The QS is based on the tick loss function (see, e.g., Giacomini and Komunjer, 2005):

$$\texttt{QS}_{\tau,it} = 2(y_{it+h|T} - \mathcal{Q}_{\tau,it})(\tau - \mathbb{1}\{y_{it+h|T} \leq \mathcal{Q}_{\tau,it}\}), \tag{2.7}$$

where $\mathcal{Q}_{\tau,it}$ is the forecast quantile of the ith variable at quantile τ, and the indicator function $\mathbb{1}\{y_{it+h|T} \leq \mathcal{Q}_{\tau,it}\}$ has a value of 1 if the outcome is at or below the forecast quantile and 0 otherwise. For $\tau = 0.5$ we obtain the absolute error; averaging these over the holdout period yields the mean absolute error (`MAE`) which is our preferred metric of point forecast accuracy.

We also evaluate density and tail forecast accuracy using several implementations of the (quantile-weighted) CRPS (`qwCRPS`, see Gneiting and Ranjan, 2011). Following Gneiting and Raftery (2007), the plain `CRPS` is defined as:

$$\texttt{CRPS}(y_{it+h}) = \int_{-\infty}^{\infty} \left(\mathcal{F}(y_{it+h|T}) - \mathbb{1}\{y_{it+h|T} \leq z\}\right)^2 dy, \tag{2.8}$$

where $\mathcal{F}$ denotes the cumulative distribution function associated with the predictive density $p(y_{it+h}|\boldsymbol{y}_{1:T})$ for $i = 1, \ldots, n$; $\mathbb{1}\{y_{it+h|T} \leq z\}$ is an indicator function taking value 1 if $y_{it+h|T} \leq z$ and 0 otherwise. We may compute the CRPS as a special case of the qwCRPS. In this context, the `qwCRPS` can be approximated as a weighted sum of `QS`s over a grid of quantiles:

$$\texttt{qwCRPS}_{it} = \frac{1}{J-1} \sum_{j=1}^{J-1} \omega(\tau_j) \texttt{QS}_{\tau_j,it}, \tag{2.9}$$

with $\tau_j = j/J$. We rely on a grid of $J - 1 = 91$ quantiles $\tau \in \{0.05, 0.06, \ldots, 0.94, 0.95\}$ to compute these scores. Using equal weights, i.e., setting $\omega(\tau_j) = 1$, results in the plain `CRPS` metric. For targeting tail forecast performance, in another implementation (denoted `qwCRPS-L`), we set the weights to $\omega(\tau_j) = (1 - \tau_j)^2$ in order to target the left tail (downside risk), and in the other (denoted `qwCRPS-R`), we set the weights as $\omega(\tau_j) = \tau_j^2$ to target the right tail (upside risk) of the predictive distribution.

Lastly, we also inspect expected shortfall (ES) and longrise (LR) as risk measures. These measures are used to characterize downside and upside risks and are also called conditional value at risk. For a chosen target probability π, the ES and LR are defined as

$$\mathtt{ES}_{it+h} = \frac{1}{\pi} \int_0^{\pi} \hat{\mathcal{F}}^{-1}_{it+h|x_t}(\tau \mid \boldsymbol{x}_t) d\tau, \tag{2.10}$$

$$\mathtt{LR}_{it+h} = \frac{1}{\pi} \int_{1-\pi}^{1} \hat{\mathcal{F}}^{-1}_{it+h|x_t}(\tau \mid \boldsymbol{x}_t) d\tau, \tag{2.11}$$

where $\hat{\mathcal{F}}^{-1}_{it+h|x_t}$ denotes the inverse of the cumulative density function, the quantile function, of variable i at horizon $t+h$ given the information in $\boldsymbol{x}_t$. Expected shortfall and longrise summarize the tail behavior of the conditional distribution in absolute terms.

3. Data

In this section, we provide an overview of the datasets alongside some descriptive statistics. We use two datasets with different frequencies: monthly and quarterly. The datasets comprise both country-specific and external variables, which come from various data sources, such as the Federal Reserve Economic Database (FRED) compiled by the St. Louis Federal Reserve, Refinitiv, Eurostat, the Istituto Nazionale di Statistica (Istat), and the Statistical Data Warehouse (SDW) of the European Central Bank (ECB). Besides the core macroeconomic and fiscal variables of interest, the datasets capture information on trade, the labor market, prices, financing conditions, and soft information in terms of sentiment indicators, which are identified by Aprigliano (2020) as valuable predictors for the Italian economy. In addition, we include variables to reflect the global business cycle and the exposure of Italy to international trade, which cover Euro area developments and specific important trading partner economies. An overview on the exact variables, data definitions, and data sources is given in Appendix A.

The quarterly dataset runs from 2000Q4 to 2023Q2. 2023Q2 is the latest data point we have from the vintage of 2024Q1, particularly for the government debt and deficit-to-GDP ratio (with positive values for this variable indicating a surplus).[1] Similarly, the monthly dataset runs from 2001M1 to 2023M10, where the end is again determined through the lens of the vintage of 2024M1. For the quarterly dataset, we design three information sets: a *small*-sized (**S**) information set, $\mathcal{M}_S$, which consists of 6 variables, a *medium*-sized (**M**) information set, $\mathcal{M}_M$,

[1] The variable refers to total deficit including interest payments.

which comprises of 11 variables, and a *large*-sized (**L**) information set, $\mathcal{M}_L$, which captures 25 variables, information set. For the monthly dataset, we only have a *small*-sized (**S**) information set, $\mathcal{M}_S$, which consists of 5 variables, and a *large*-sized (**L**) information set, which comprises 12 variables.

Unfortunately, most variables are not available in real time, so that we conduct a pseudo-out-of-sample evaluation.

4. Results

In this section we assess the performance of the econometric models described in Section 2 in terms of the metrics discussed above. We focus on the ability of the competing models to provide accurate point and density forecasts for the variables that are of paramount interest in this paper, i.e., fiscal aggregates and key macroeconomic indicators. We organize the comparison in two subsections, the first one focuses on the results obtained for the quarterly dataset, the second one is based on monthly data. All the metrics reported are computed in relative terms with respect to the ones attained by the *medium* (**M**) BVAR-SV model.

4.1 Quarterly Forecasting Results

The key target variables in the quarterly dataset are real GDP growth, inflation measured as the rate of percent change in the HICP, and government debt and deficit, both computed as a share of GDP. These latter two indicators provide crucial information about the sustainability of a country's public finances, therefore understanding their likely future path and the amount of risk embedded is of primary importance for public institutions and investors.[2] The first estimation sample is 2001Q4 to 2010Q4, which is then recursively expanded, so that the forecast evaluation period includes about 90 observations. The first evaluation sample thus ranges from 2011Q1 to 2012Q4.

Table 2 reports the CRPSs and MAEs attained by the competing models relative to those of the *medium* BVAR-SV, for which the metrics are reported in absolute terms. The four blocks on the left of the Table (one for each variable), denoted by **S**, refer to the *small* versions of the models, while the last four blocks on the right, denoted by **M**, refer to the *medium* specifications. In the appendix, Tables B1 and B2 report also the relative comparison of the medium-sized model to the *small* (**S**) and *large* (**L**) model. The *large* (**L**) model specification does not lead to any

[2] Results for the other variables are available upon request.

Table 2: Evaluation based on variants of CRPS/MAE.

S

Variable	Model	CRPS 1	CRPS 4	CRPS 8	CRPS−L 1	CRPS−L 4	CRPS−L 8	CRPS−R 1	CRPS−R 4	CRPS−R 8	MAE 1	MAE 4	MAE 8
Government Debt Ratio	BVAR SV	1.00	1.02	1.23	1.05	1.07	1.22	0.97	**0.98**	1.22	1.00	1.06	1.22
	BVAR SV−t	1.20	1.23	1.38	1.20	1.36	1.46	1.22	1.16	1.31	1.07	1.29	1.39
	TVP SV	1.04	*0.99*	*1.21*	1.10	*1.00*	*1.21*	0.99	0.98	*1.19*	1.01	*1.02*	*1.22*
	BART hom	0.96	1.24	1.37	0.99	1.35	1.41	0.95	1.15	1.31	0.89	1.24	1.43
	BART SV	*0.94*	1.23	1.39	*0.96*	1.30	1.40	*0.93*	1.17	1.35	*0.88*	1.26	1.47
	GP hom	1.01	1.27	1.42	1.04	1.37	1.41	0.99	1.19	1.41	0.96	1.30	1.46
	GP SV	1.00	1.25	1.43	1.01	1.30	1.40	1.00	1.21	1.42	0.96	1.31	1.46
Deficit Ratio	BVAR SV	1.11	1.02	0.98	1.15	1.00	0.97°	1.06	1.04	0.99	1.09	1.03	0.98
	BVAR SV−t	1.21	1.02	1.01	1.34	1.01	0.99	1.10	1.04	1.02	1.14	1.03	1.03
	TVP SV	*0.98*	*1.01*	0.99	*0.97*	*0.99*	0.99	0.99	1.04	0.99	*0.96*	*0.98*	0.96'
	BART hom	1.05	1.04	0.95°	1.15	1.08	0.97	0.95	0.99	0.91*	0.98	1.04	0.93°
	BART SV	0.99	1.03	0.96°	1.06	1.07	0.98	*0.91*	0.98	0.92*	0.98	1.04	0.95'
	GP hom	1.10	1.05	*0.92**	1.23	1.08	**0.94°**	0.96	0.99	*0.88°*	1.02	1.05	0.92'
	GP SV	1.07	1.03	0.92°	1.18	1.09	0.95'	0.97	*0.95*	0.88°	1.01	1.04	*0.92'*
Growth	BVAR SV	1.01	0.96	0.95	0.92	**0.94**	**0.93'**	1.12	0.96	0.95	1.02	0.98	0.98
	BVAR SV−t	1.08	**0.95**	0.95	0.97	0.94	0.93	1.22	**0.96**	0.96	1.12	0.99	0.98
	TVP SV	**0.97**	0.97	0.95	**0.92'**	0.95	0.94	1.04	0.98	0.96	0.98	*0.98*	0.98
	BART hom	1.12	1.00	0.96'	1.08	0.99	0.95'	1.14	1.01	0.96'	1.11	1.07	1.00
	BART SV	1.00	1.01	0.97'	0.99	1.01	0.96'	1.01	1.00	0.96'	1.02	1.06	0.98
	GP hom	1.03	1.01	0.95'	0.97	0.99	0.94'	1.08	1.03	**0.94**	*0.98*	1.05	0.99
	GP SV	1.00	1.01	**0.95'**	0.98	1.00	0.94'	*1.01*	1.01	0.94	0.99	1.05	**0.98**
Inflation	BVAR SV	0.98	0.96'	*0.98*	**0.98**	0.96	0.99	0.98	0.96°	0.97°	0.99	0.96'	0.97
	BVAR SV−t	0.98	1.00	1.00	0.99	1.03	1.02	**0.97**	0.99	0.98'	**0.98**	0.98	0.98
	TVP SV	**0.97**	**0.94°**	0.98	0.98	0.93	0.99	0.97	**0.94°**	0.97	0.99	**0.93°**	0.99
	BART hom	1.09	1.00	1.02	1.07	0.95	1.00	1.09	1.02	1.02	1.09	0.97	0.97'
	BART SV	1.13	0.97	1.01	1.10	**0.92'**	*0.98*	1.15	1.01	1.02	1.14	0.95'	*0.97°*
	GP hom	1.29	1.02	1.02	1.21	0.97	1.00	1.34	1.03	1.03	1.31	0.99	0.99
	GP SV	1.28	1.01	1.03	1.26	0.98	1.01	1.29	1.03	1.03	1.34	1.00	0.99

M

Variable	Model	CRPS 1	CRPS 4	CRPS 8	CRPS−L 1	CRPS−L 4	CRPS−L 8	CRPS−R 1	CRPS−R 4	CRPS−R 8	MAE 1	MAE 4	MAE 8
Government Debt Ratio	BVAR SV	1.07	2.7	3.86	0.3	0.66	0.94	0.33	0.93	1.35	1.54	3.36	4.8
	BVAR SV−t	1.23	1.03	1.09	1.24	1.01	1.15	1.22	1.03	1.04	1.08	1.02	1.12
	TVP SV	1.02	*1.00*	1.02	1.00	*0.98*	1.04	1.03	1.01	1.01	1.01	*1.00*	**0.99**
	BART hom	0.96	1.21	1.28	1.00	1.33	1.39	0.93	1.11	1.18	*0.89*	1.22	1.36
	BART SV	*0.94*	1.15	1.24	*0.96*	1.18	1.29	*0.93*	1.12	1.19	0.90	1.20	1.36
	GP hom	1.04	1.29	1.40	1.06	1.39	1.42	1.02	1.21	1.36	0.98	1.32	1.47
	GP SV	1.03	1.26	1.38	1.06	1.31	1.38	1.01	1.21	1.36	0.98	1.31	1.46
Deficit Ratio	BVAR SV	0.31	0.96	1.7	0.09	0.33	0.62	0.09	0.23	0.38	0.42	1.16	1.96
	BVAR SV−t	1.01	**0.99**	1.03	1.03	**0.99**	1.02	0.99	0.99	1.05	1.01	**0.97**	1.03
	TVP SV	*0.95'*	1.02	1.05	*0.94'*	1.00	1.03	0.97	1.05	1.08	*0.94*	0.99	1.06
	BART hom	1.04	1.05	0.95°	1.16	1.08	0.97	0.93	1.00	0.91*	0.98	1.04	0.95
	BART SV	0.98	1.03	0.96'	1.06	1.06	0.99	*0.90*	0.97	0.91°	0.97	1.05	0.96
	GP hom	1.08	1.04	**0.92°**	1.22	1.09	*0.95*	0.94	0.98	**0.86***	1.01	1.02	**0.90°**
	GP SV	1.05	1.03	0.92°	1.15	1.08	0.96	0.95	*0.96*	0.86*	1.00	1.02	0.90'
Growth	BVAR SV	3.81	4.3	4.86	1.25	1.37	1.48	1	1.14	1.36	4.61	4.65	5.16
	BVAR SV−t	1.04	*0.99*	0.99	1.02	*0.98*	0.99	1.06	1.00	1.00	1.00	0.99	1.01
	TVP SV	1.05	1.01	1.01	1.02	1.01	1.01	1.10	1.01	1.01	1.03	*0.98'*	1.00
	BART hom	1.06	1.02	*0.95*	1.01	0.99	*0.95'*	1.08	1.04	0.95	1.01	1.07	0.99
	BART SV	*0.98*	1.02	0.99	1.02	1.03	0.98	**0.95'**	1.01	0.99	**0.95**	1.06	1.01
	GP hom	1.06	1.02	0.96	1.01	1.00	0.96	1.08	1.03	*0.95*	0.99	1.06	0.99
	GP SV	1.03	1.03	0.97	1.03	1.01	0.97	1.03	1.04	0.97	0.99	1.05	*0.99*
Inflation	BVAR SV	1.04	1.8	1.97	0.27	0.46	0.51	0.34	0.6	0.66	1.32	2.34	2.58
	BVAR SV−t	1.01	1.02	1.03	1.05	1.08	1.03	*0.97*	*0.98*	1.02	1.03	1.01	1.03
	TVP SV	1.00	*0.97*	*0.98*	1.01	0.95	0.99	1.00	0.99	**0.96**	1.03	0.98	1.03
	BART hom	1.17	1.03	1.02	1.13	0.99	0.99	1.19	1.05	1.03	1.18	0.98	0.96°
	BART SV	1.22	0.98	0.99	1.20	*0.94'*	*0.95'*	1.25	1.02	1.01	1.26	*0.97'*	*0.96°*
	GP hom	1.37	1.01	1.00	1.27	0.95	0.96'	1.43	1.04	1.02	1.39	0.99	0.97'
	GP SV	1.33	1.00	1.00	1.30	0.95	0.96	1.35	1.03	1.02	1.41	0.99	0.97

Horizon

Notes: The table shows the continuous ranked probability score (CRPS) and quantile-weighted version for the left (qwCRPS-L) and right (qwCRPS-R) tail, together with the mean absolute error (MAE) scores for key variables and different models for horizon $h = \{1, 4, 8\}$. Results in the left panel are based on the *small* (**S**) information set; on the right panel on the *medium* (**M**) information set. Grey shaded rows indicate raw losses for the benchmark, all others are shown as ratios relative to these values (red indicates worse and blue superior performance).

gains, thus we restrict the discussion on the two smaller specifications. Specifically for the deficit ratio, we find that the MAE and CRPS show no gains in the small or large model. However, the CRPS-R as well as the QS75/90 point to predictive gains at the right tail of the distribution for the nonlinear models at longer horizons.

Starting with the comparison of the point forecasts measured in terms of MAE, for the debt-to-GDP ratio, the *medium* BVAR-SV appears to be a very tough benchmark at all horizons. The absolute errors made on average by the other models are in most cases significantly larger than the ones made by the *medium* BVAR-SV. For the deficit-to-GDP ratio, the benchmark is hard to beat at short horizons. However, at longer horizons we find predictive gains particularly for the nonparametric models. For GDP growth, there remains the remarkably good performance of the *medium* BVAR-SV, echoing the results on the deficit-to-GDP ratio. For inflation, the *medium* TVP-BVAR and homoskedastic BART-VAR are only slightly better than the benchmark at the

Table 3: Evaluation based on Quantile Scores.

S

		QS10			QS25			QS75			QS90		
Government Debt Ratio	BVAR SV	1.24	1.04	1.15	1.05	1.13	1.23	0.97	**0.94**	1.16	0.93	**0.94**	1.26
	BVAR SV−t	1.65	1.54	1.47	1.15	1.42	1.50	1.21	1.13	1.27	1.54	1.05	1.25
	TVP SV	1.41	*1.03*	*1.11*	1.09	*1.01*	*1.22*	1.00	0.97	*1.13*	**0.88**	0.97	*1.24*
	BART hom	1.25	1.64	1.36	1.01	1.40	1.41	0.89	1.11	1.26	1.18	1.07	1.26
	BART SV	*1.09*	1.42	1.25	*1.00*	1.35	1.39	*0.88*	1.14	1.30	1.11	1.11	1.29
	GP hom	1.24	1.60	1.30	1.05	1.41	1.39	0.91	1.16	1.35	1.18	1.12	1.44
	GP SV	1.12	1.29	1.26	1.02	1.34	1.40	0.93	1.18	1.36	1.19	1.16	1.47
Deficit Ratio	BVAR SV	1.06	0.97	0.97'	1.28	1.00	0.96'	1.04	1.05	1.02	1.05	1.05	0.95'
	BVAR SV−t	1.55	**0.95'**	0.97°	1.44	1.04	0.97'	1.04	1.05	1.03	1.05	1.01	0.97
	TVP SV	*1.01*	0.98	0.99	*0.96*	*1.00*	0.98	1.01	1.09	1.01	1.04	1.08	0.96
	BART hom	1.33	1.09	0.96	1.23	1.10	0.98	0.90	1.00	0.89°	0.93	0.86	0.80°
	BART SV	1.08	1.06	0.98	1.14	1.09	1.00	*0.87'*	0.99	0.90°	**0.83°**	0.81°	0.81'
	GP hom	1.42	1.10	**0.95°**	1.33	1.11	**0.94'**	0.91	0.97	0.85°	0.88	0.86	**0.79'**
	GP SV	1.30	1.11	0.96'	1.29	1.12	0.95	0.89'	*0.92*	*0.83°*	0.98	**0.79'**	0.82'
Growth	BVAR SV	**0.84°**	*0.92*	0.91	**0.90'**	**0.93**	**0.92'**	1.14	1.01	0.97	1.26	0.90	0.92
	BVAR SV−t	0.86°	0.92	0.90	0.92	0.93	0.92	1.23	*1.00*	0.97	1.38	**0.87'**	0.92
	TVP SV	0.86°	0.93	0.92	0.91°	0.95	0.93	1.04	1.01	0.97	1.13	0.92	0.93
	BART hom	1.04	0.92	0.91'	1.09	0.99	0.95	1.05	1.09	0.96	1.37	0.88	0.91
	BART SV	0.96	0.96	0.92	0.99	1.01	0.96	*1.00*	1.03	0.98	1.07	0.91	0.92
	GP hom	0.94	0.92	**0.90'**	0.97	0.99	0.95	1.04	1.12	*0.95*	1.30	0.91	**0.90**
	GP SV	0.95°	0.94	0.91'	0.99	1.00	0.94	1.02	1.08	0.96	*1.04*	0.90	0.91
Inflation	BVAR SV	0.97	0.92	1.06	*0.97*	0.96	1.00	0.98	0.97°	0.96°	0.98	0.96°	0.99
	BVAR SV−t	1.01	1.10	1.13	0.99	1.05	1.03	*0.97*	0.98	0.97°	0.94	0.99	0.99
	TVP SV	0.96	0.90	*1.00*	0.98	0.94	1.00	0.97	**0.95°**	0.98	**0.94°**	**0.92**	*0.95*
	BART hom	1.10	0.93	1.16	1.05	0.94	0.98	1.10	1.02	1.00	1.08	1.09	1.11
	BART SV	1.08	*0.83'*	1.04	1.08	**0.91'**	*0.96*	1.17	1.00	0.99	1.13	1.07	1.11
	GP hom	1.09	0.90	1.08	1.16	0.97	0.99	1.35	1.04	1.00	1.38	1.08	1.10
	GP SV	1.20	0.91	1.11	1.22	0.98	1.01	1.32	1.03	1.01	1.23	1.06	1.11
	Horizon	1	4	8	1	4	8	1	4	8	1	4	8

M

		QS10			QS25			QS75			QS90		
Government Debt Ratio	BVAR SV	0.53	1.24	1.94	1.13	2.31	3.3	1.27	3.57	5.43	0.71	2.97	4.15
	BVAR SV−t	1.68	*0.99*	1.13	1.23	1.01	1.21	1.20	1.03	1.00	1.54	1.06	1.00
	TVP SV	**0.98**	1.06	1.16	1.01	*0.96*	1.05	1.06	1.01	1.02	1.00	1.01	**1.00**
	BART hom	1.27	1.65	1.42	1.02	1.36	1.41	*0.87*	1.09	1.14	1.12	1.03	1.06
	BART SV	1.09	1.15	1.08	*1.00*	1.20	1.33	0.88	1.11	1.16	1.12	1.06	1.07
	GP hom	1.27	1.61	1.30	1.09	1.41	1.43	0.96	1.20	1.31	1.22	1.11	1.29
	GP SV	1.22	1.35	1.17	1.08	1.36	1.40	0.95	1.19	1.31	1.19	1.13	1.32
Deficit Ratio	BVAR SV	0.23	1.12	2.29	0.34	1.23	2.36	0.33	0.81	1.32	0.18	0.48	0.7
	BVAR SV−t	1.01	1.01	1.01	1.08	**0.99**	1.01	0.98	1.04	1.05	1.01	0.97	1.10
	TVP SV	1.00	1.01	1.00	**0.90**	1.00	1.04	0.99	1.10	1.10	1.03	1.08	1.11
	BART hom	1.36	1.09	0.96	1.24	1.10	0.99	0.88	1.00	0.87°	0.87	0.91	0.81°
	BART SV	1.09	1.05	0.97	1.15	1.09	1.01	**0.84°**	0.94	0.86°	0.85°	*0.81°*	0.81'
	GP hom	1.45	1.12	*0.95*	1.32	1.12	*0.96*	0.90	0.95	0.81*	*0.84'*	0.90	*0.80'*
	GP SV	1.26	1.11	0.96	1.26	1.11	0.97	0.90'	*0.93*	**0.80°**	0.91	0.84'	0.82'
Growth	BVAR SV	4.02	4.62	4.81	4.69	5.16	5.51	3.74	3.74	4.71	2.26	3.58	4.39
	BVAR SV−t	1.01	0.97	0.97	1.02	*0.97*	0.98	1.08	1.04	1.01	1.12	0.97	0.98
	TVP SV	1.00	1.02	1.02	1.02	1.01	1.01	1.07	1.03	1.03	1.22	1.02	1.01
	BART hom	*0.98*	*0.90*	*0.91*	1.03	0.99	*0.95*	1.02	1.11	*0.97*	1.28	0.93	0.90
	BART SV	1.01	0.99	0.96°	1.05	1.03	0.98	**0.90°**	1.00	0.99	**0.94**	0.97	0.98
	GP hom	0.99	0.91	0.92	1.03	1.00	0.96	1.03	1.12	0.97	1.32	0.92	*0.90*
	GP SV	1.02	0.96	0.96	1.05	1.02	0.97	1.00	1.11	0.98	1.13	0.97	0.93
Inflation	BVAR SV	0.51	0.84	0.89	1.01	1.69	1.89	1.3	2.25	2.5	1.03	1.8	1.92
	BVAR SV−t	1.14	1.24	1.07	1.06	1.09	1.03	**0.94**	*0.98*	1.03	*0.95*	*0.94*	1.01
	TVP SV	1.02	0.92	0.99	1.00	*0.92*	0.96	0.99	1.00	*0.98*	0.98	0.98	*0.85*
	BART hom	1.13	1.00	1.14	1.09	0.97	0.97	1.20	1.04	1.00	1.20	1.12	1.14
	BART SV	1.16	*0.86'*	*0.96*	1.14	0.93'	*0.94°*	1.29	1.02	1.00	1.18	1.06	1.10
	GP hom	1.10	0.86°	1.00	1.22	0.93'	0.95'	1.44	1.03	1.00	1.47	1.09	1.10
	GP SV	1.19	0.88'	0.99	1.26	0.94'	0.95'	1.39	1.03	1.00	1.22	1.08	1.10
	Horizon	1	4	8	1	4	8	1	4	8	1	4	8

Notes: The table shows the quantile scores for quantiles $\tau = \{0.10, 0.25, 0.75, 0.90\}$ for key variables and different models for horizon $h = \{1, 4, 8\}$. Results in the left panel are based on the *small* (**S**) information set; on the right panel on the *medium* (**M**) information set. Grey shaded rows indicate raw losses for the benchmark, all others are shown as ratios relative to these values (red indicates worse and blue superior performance).

one quarter horizon. For inflation, BART and GP are slightly better at the four- and eight-quarter ahead horizons, but much worse at the one-quarter horizon.

Also for density forecasts, according to the CRPS there is no model that exhibits a clear and systematic advantage relative to the *medium* BVAR-SV. For the deficit-to-GDP ratio, BART and GP are slightly better only one quarter ahead, with comparable outcomes for the small and large versions of the models. For growth, adding TVP helps a bit at all horizons, but only in the *small* specification. For inflation, there are also some slight gains from TVP-BVAR, but mainly at short horizons.

As the risks for policy makers and investors mainly come from events in the tails of the distributions, we now consider CRPS-R for future debt- and inflation, and CRPS-L for deficit and GDP growth. Also in this case, the non-parametric models do not generally outperform the more standard BVAR-SV model.

It is worth noting that SV is very helpful for the linear models, as the forecasting performance of a homoskedastic BVAR is generally much worse. Instead, SV is much less important for the nonlinear specifications, due to their more flexible specification for the conditional mean. Also, for both the BVAR-SV and the nonparametric specifications, working with larger models does not systematically lead to better performance.

Table 3, which reports quantile scores computed for the 10th, 25th, 75th, and 90th percentiles, provides a slightly more positive picture for the nonparametric models, though without a clear and systematic winner emerging. In particular, focusing on the 90th percentile for deficit-to-GDP, and on the 10th and 90th ones for growth, it turns out that BART and GP are often helpful at the one- and two-year horizons. For the debt-to-GDP ratio, we do not find predictive gains for the nonparametric models at longer horizons but some at the one-step ahead predictions for the 75th percentile.

For deficit,the gains from employing non-parametric models are concentrated in the right tail of the distributions, which contains favourable events from the perspective of the policymaker. For growth, we find remarkable gains for the *small* specification in the BVAR-SV but also for the BVAR-SV-t and TVP-BVAR-SV, whereas predictive gains are mostly visible at longer horizons for the nonparametric models. For inflation, the *small* TVP-BVAR-SV produces gains between 5% and 7%.

Hence, if the focus is on specific percentiles, variables, and horizons, time-varying and/or nonlinear models applied to different information sets can yield some gains with respect to the *medium* BVAR-SV benchmark, but no specification can be selected as systematically best for the four key indicators we are studying. The medium BVAR-SV, on the other hand, seems overall a robust tool for investigating tail risks.

Next, we evaluate whether this finding holds for the entire evaluation period or the ranking of the models changed over time.

Time Variation in Predictive Accuracy

To assess stability of the (absolute and relative) predictive accuracy, we plot in Figure 1 the cumulative CRPS. The cumulation is over the hold-out period, so that the value corresponding to 2022Q4 reflects the overall score attained by each model, but the time evolution of the measure allows to shed light on specific dates in which some models provided better density forecasts. Figures 2 and 3 report the cumulative left- and right-weighte CRPS, respectively. Since in general we did not encounter benefits in considering *large* specifications, in this section we only focus on the *medium* versions of the models. Moreover, since we did not observe appreciable gains from introducing SV in the nonparametric models, we only consider homoskedastic BART and GP in the figures below.

As for government debt-to-GDP ratio, it can be seen that the BVAR-SV (solid black line) remains amongst the best performing models throughout the hold-out period. Almost all models perform similarly for the whole time span, with the exception of the BVAR SV-t model (dashed blue line), which accumulates a large disadvantage for short-term forecasts after the Covid-19 pandemics, and the nonparametric models (dotted grey and solid orange lines), which appear less suitable at longer horizons.

All the considered models perform similarly throughout the hold-out sample also when forecasting deficit-to-GDP ratios. Nonparametric models show some forecasting advantages since the beginning of the hold-out period for one-year and two-year ahead predictions. This advantage, however, mainly comes from the right part of the distribution, which does not contain unfavourable events from a policymaker's point of view.

For growth, as for the other variables, there is a deterioration in the absolute performance with the onset of the Covid-19 pandemic, but the differences between alternative models remain small over the entire evaluation period.

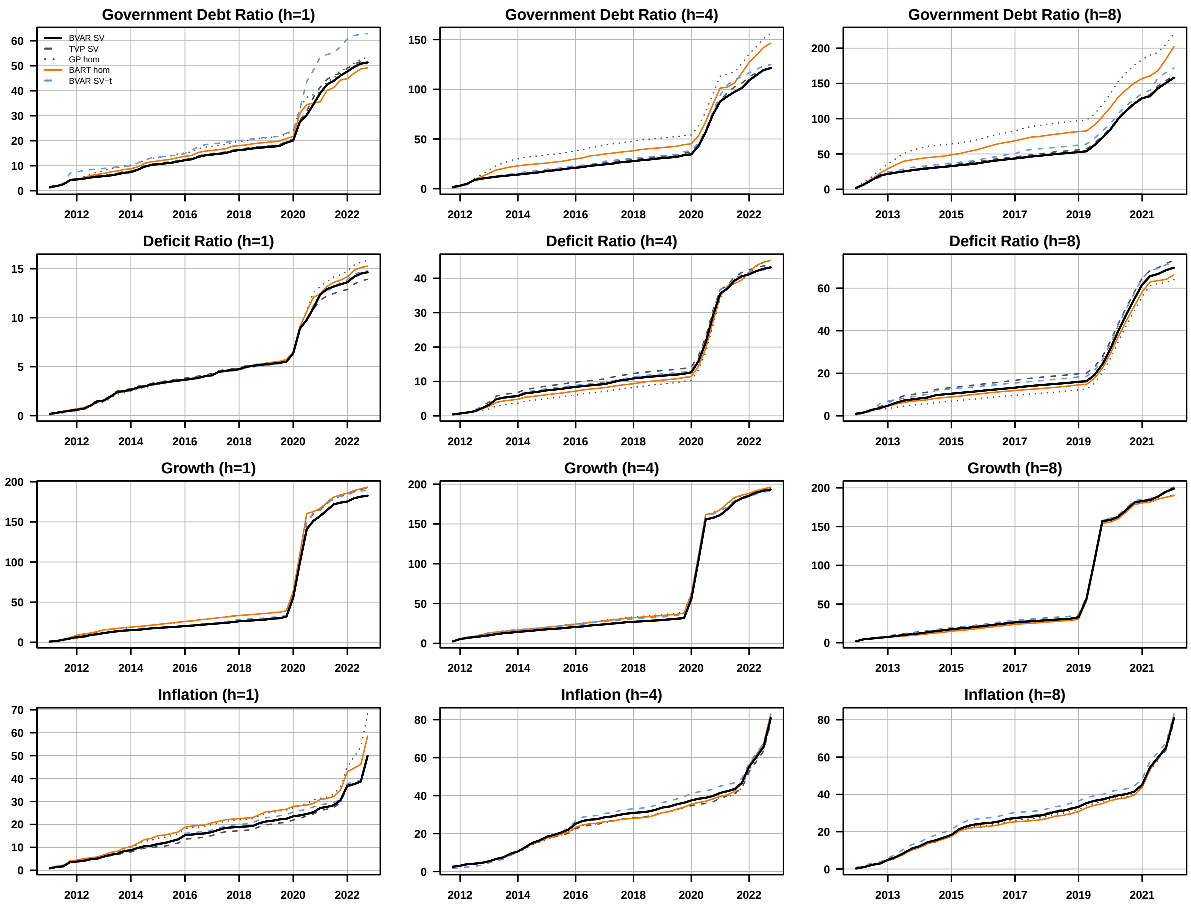
Also for inflation, the performance of all models are relatively similar throughout the period, with a sensible deterioration triggered by the surge in inflation started at the end of 2021 and exacerbated by the Russian invasion of Ukraine. Prior to these events, nonparametric models retained some slight advantage at one- and two-year horizons, but after the inflation spike the edge was reabsorbed.

Measusres of Tail Risk

Tail risk can be measured in many ways. Some of the most useful and straightforward among these measures are expected shortfall (ES) and longrise (LR), which can be thought as point predictions conditional on being in an extreme scenario. More specifically, we focus on 5% expected shortfall and longrise, which define the expected future value of the target variable, conditional on landing respectively on the left or on the right 5% tail of the distribution.

Figure 4 plots the two measures (shades of blue for shortfall, shades of red for longrise) computed using 1-quarter ahead and 1-year ahead predictive densities, together with the actual

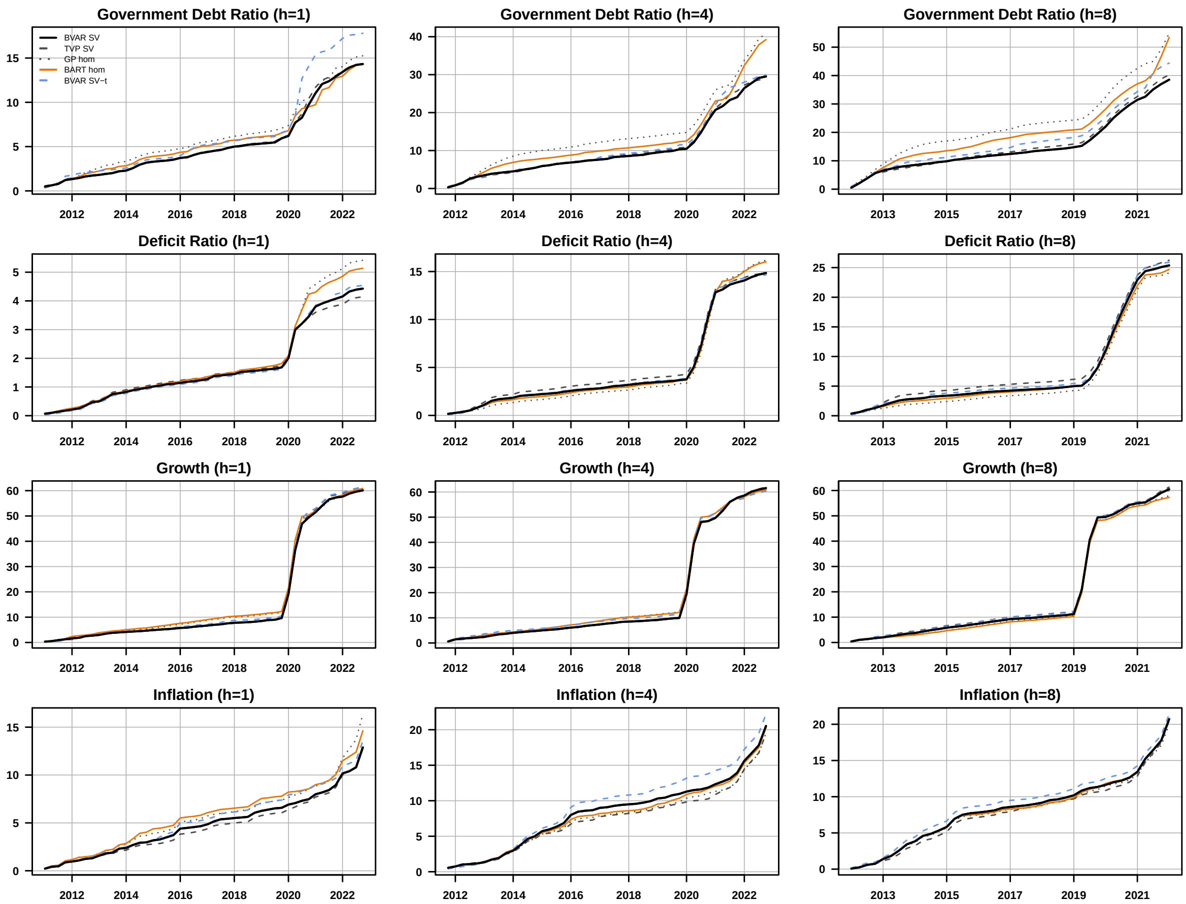
Figure 1: Cumulative Scores Over the Hold-Out Sample: `CRPS`.



Notes: The figure shows the time series evolution of continuous ranked probability score (`CRPS`) for different variables and models at horizon $h = \{1, 4, 8\}$. Results are based on the *medium* (**M**) information set.

realization (solid black). Before the Covid-19 pandemics, the pictures provided by the different models were almost indistinguishable. In fact, before 2020, the actual realizations of debt and deficit ratios never exceeded the shortfall or longrise levels predicted by all models. As the outbreak of Covid-19 was not predictable, all models failed to assign enough probability, either one year or one month in advance, to the large debt increase of 2020Q2. Following the large deficit run by the Italian government to face the initial phase of the pandemics, the volatility of both debt- and deficit-to-GDP ratios increased remarkably and generated substantial differences between the risk measures provided by different models. Both shortfall and longrise become larger in absolute value for all models, but much more so for models that feature stochastic volatility, which were able to take the increased uncertainty into account. A similar comment applies for the evolution of ES and LR for growth, while for inflation the most challenging event for the models is its recent increase. The LR of all models also increased, but typically not fast enough.

Figure 2: Cumulative Scores Over the Hold-Out Sample: `qwCRPS-L`.

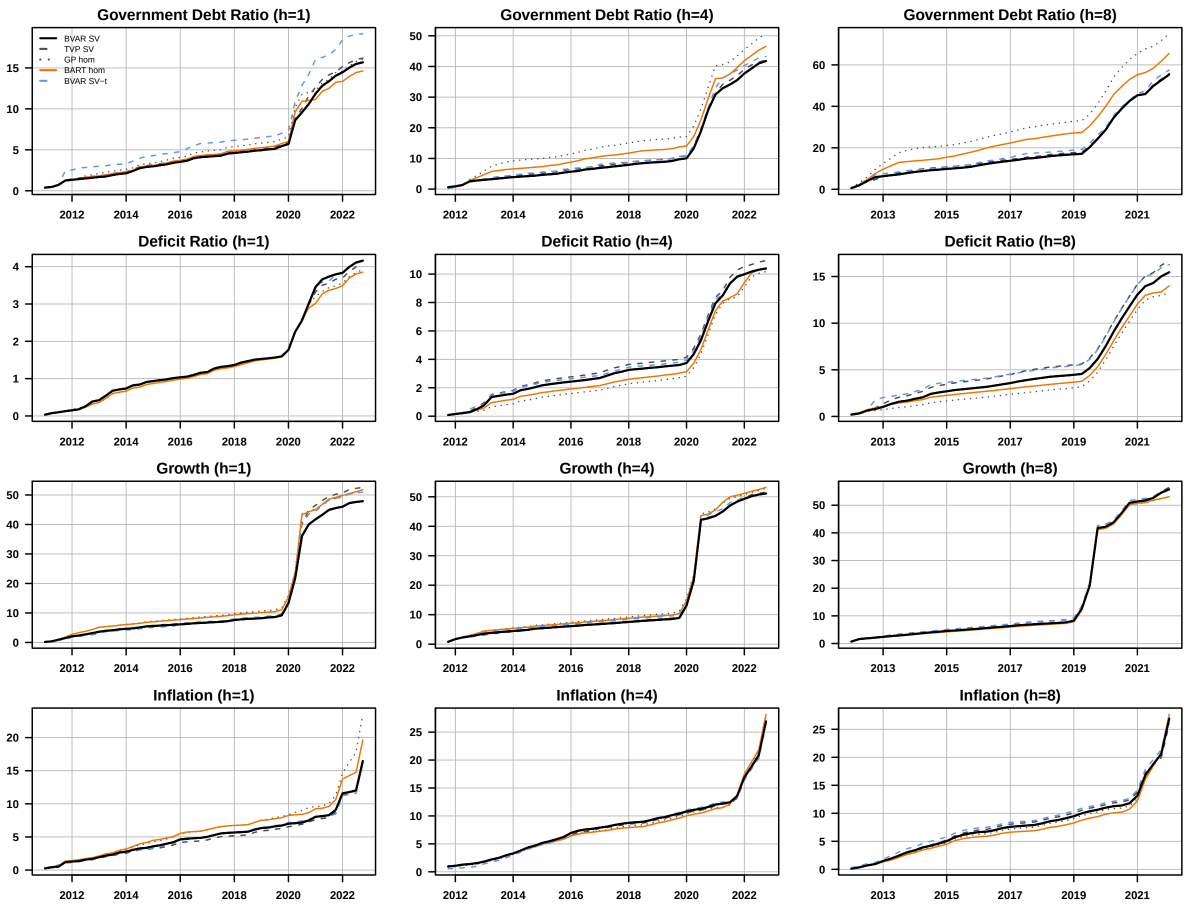


Notes: The figure shows the time series evolution of quantile-weighted continuous ranked probability score for the left tail (`qwCRPS-L`) scores for different variables and models at horizon $h = \{1, 4, 8\}$. Results are based on the *medium* (**M**) information set.

As an additional way to measure tail risk, Figure 5 plots the one-quarter and one-year ahead predicted probabilities of events of particular interest: a further increase in the debt/GDP ratio, a deficit/GDP ratio above 3%, negative growth, and inflation above 2%.[3] A few noticeable features are the increases of the probability of negative growth during the sovereign debt crisis and the Covid-19 pandemic, deficit/GDP above 3%, as well as the increase in the probability of inflation higher than 2% after Covid, with the increasing happening first for the `GP-hom` model, already around 2020.

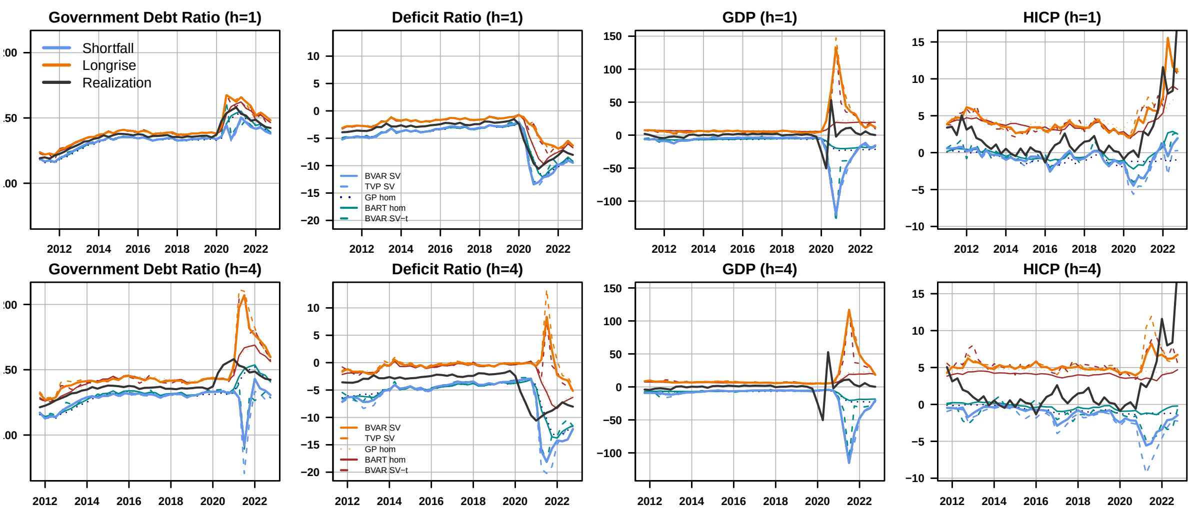
[3] These are just examples of specific events of typical interest for policy makers, of course our methodology permits the computation of the probability of occurrence of all sorts of events related to the modelled variables.

Figure 3: Cumulative Scores Over the Hold-Out Sample: `qwCRPS-R`.



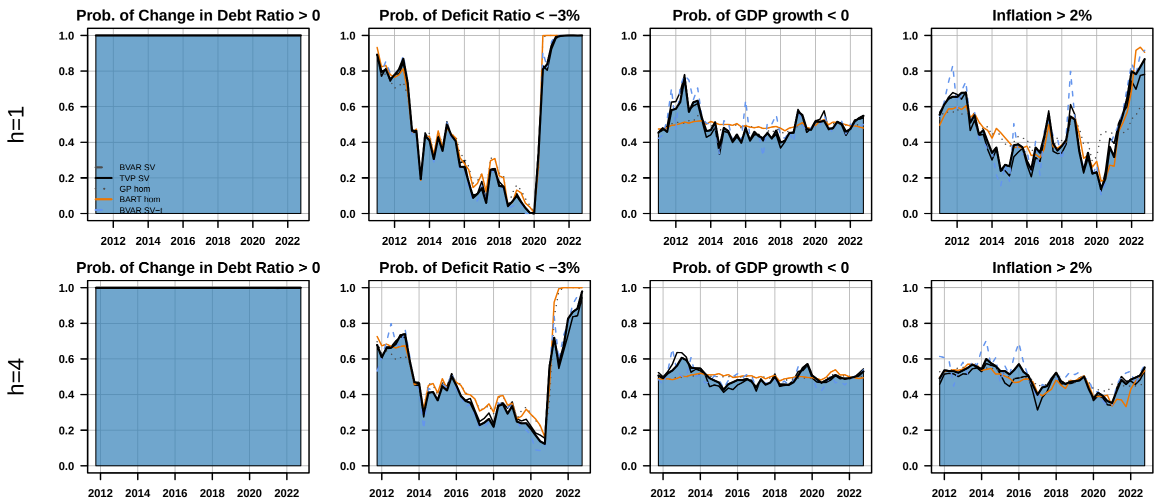
Notes: quantile-weighted continuous ranked probability score for the right tail (`qwCRPS-R`) scores for different variables and models at horizon $h = \{1, 4, 8\}$. Results are based on the *medium* (**M**) information set.

Figure 4: Expected Shortfall and Longrise Over Time.



Notes: The figure shows the time series evolution of the 5 percent expected shortfall (blue lines) and longrise (orange lines) for different variables and models at horizon $h = \{1, 4\}$. Results are based on the *medium* (**M**) information set. Black bold line denotes realization of outcome variable.

Figure 5: Probabilities of Scenarios.



Notes: The figure shows the time series evolution of the following scenarios: change in government debt > 0, deficit ratio $<$ -3%, GDP growth < 0, and inflation $> 2\%$ at horizon $h = \{1, 4\}$. Results are based on the *medium* (**M**) information set. Black bold line denotes BVAR-SV.

Table 4: Evaluation Based on Variants of CRPS/MAE: Growth rate of Monthly Indicators.

S

		CRPS				CRPS−L				CRPS−R				MAE			
	Horizon	1	3	12	24	1	3	12	24	1	3	12	24	1	3	12	24
Industrial Production	BVAR SV	1.54	1.58	1.63	1.72	0.51	0.48	0.48	0.51	0.4	0.44	0.47	0.49	2.02	1.95	1.97	2.11
	BVAR SV−t	1.02	**0.99**	**0.99**	1	1	**0.98**	**1.00**	1.01	1.05	**0.99**	**0.99**	**1.00**	1.05	1	1.01	1.01
	TVP SV	1.03	1.01	1.03	1.04	1.01	1.01	1.03	1.05	1.06	1	1.02	1.03	1.04	**1.00**	1	1.01
	BART hom	1.49	1.28	1.04	1.03	1.2	1.25	1.04	1.02	1.84	1.32	1.05	1.03	1.39	1.3	1.03	1.02
	BART SV	*0.97*	1.02	1.01	1	0.94	1.01	1.01	1	1.02	1.04	1.01	1.01	0.96	1.02	1.01	1.01
	GP hom	1.02	1.02	1.01	*1.00*	0.94	1	1.01	*1.00*	1.11	1.04	1.02	1	0.96	1.01	1.01	1
	GP SV	0.97	1.02	1.02	1	**0.90**	1.01	1.02	1	1.07	1.02	1.03	1	*0.96*	1.01	1	*1.00*
Consumer Prices	BVAR SV	0.13	0.2	0.38	0.49	0.04	0.05	0.1	0.12	0.04	0.06	0.13	0.17	0.17	0.26	0.51	0.65
	BVAR SV−t	1.01	1.01	1.03	1.04	1.01	1.01	1.03	1.04	1	1.01	1.02	1.03	1.01	1.01	1.02	1.03
	TVP SV	**0.97**	**0.94**	1.07	1.18	**0.97**	**0.98**	1.12	1.24	**0.98**	**0.91**	1.03	1.14	**0.95'**	**0.94**	1.09	1.11
	BART hom	1.27	1.27	1.26	1.3	1.24	1.3	1.37	1.53	1.31	1.24	1.16	1.11	1.32	1.29	1.22	1.26
	BART SV	1.21	1.12	**1.00**	1.23	1.19	1.13	1.1	1.4	1.24	1.1	**0.91**	1.1	1.22	1.13	**0.97**	1.16
	GP hom	1.7	1.39	1.4	1.42	1.65	1.42	1.51	1.74	1.76	1.35	1.3	1.18	1.83	1.47	1.38	1.4
	GP SV	1.54	1.28	1.34	1.42	1.53	1.34	1.43	1.72	1.57	1.22	1.26	1.19	1.65	1.35	1.33	1.43

L

		CRPS				CRPS−L				CRPS−R				MAE			
	Horizon	1	3	12	24	1	3	12	24	1	3	12	24	1	3	12	24
Industrial Production	BVAR SV	1.03	1.05	1.07	1.04	1.03	1.06	1.07	1.03	1.03	1.04	1.07	1.05	1.04	1.01	**1.00**	**0.99'**
	BVAR SV−t	1.04	1.04	1.05	1.04	1.04	1.04	1.05	1.03	1.04	1.05	1.06	1.05	1.03	*1.00*	1.01	1.01
	TVP SV	1.06	1.06	1.1	1.13	1.05	1.06	1.1	1.13	1.06	1.06	1.11	1.15	1.06	1.01	1.02	1.02
	BART hom	1	1.03	1.02	1	*0.93*	*1.01*	1.02	1.01	1.09	1.05	*1.01*	*1.00*	**0.95**	1.02	1.02	1
	BART SV	**0.96**	1.03	1.04	1.02	0.93	1.04	1.04	1.02	*1.02*	*1.02*	1.03	1.02	0.97	1.03	1.02	1
	GP hom	1.03	*1.02*	*1.01*	**1.00**	0.95	1.01	*1.02*	**0.99**	1.13	1.04	1.02	1	0.99	1.01	1.01	1
	GP SV	0.98	1.03	1.05	1.02	0.94	1.04	1.05	1.02	1.05	1.02	1.05	1.02	0.99	1.01	1.01	1
Consumer Prices	BVAR SV	*1.00*	1.03	1.06	*1.15*	1	1.04	*1.09*	*1.23*	*1.01*	1.02	1.02	1.11	1.02	1.05	1.09	1.14
	BVAR SV−t	1.01	1.04	1.09	1.22	*0.99*	1.04	1.11	1.29	1.03	1.03	1.07	1.17	1.02	1.06	1.11	1.2
	TVP SV	1.02	*0.98*	1.09	1.29	1.01	*1.02*	1.19	1.5	1.03	*0.95*	1.02	1.15	*1.02*	*0.99*	1.11	1.18
	BART hom	1.25	1.2	1.12	1.32	1.18	1.18	1.2	1.46	1.32	1.21	1.04	1.19	1.24	1.18	1.07	1.23
	BART SV	1.3	1.14	*1.02*	1.18	1.26	1.15	1.12	1.35	1.35	1.12	*0.94*	*1.05*	1.29	1.1	*0.98*	*1.10*
	GP hom	1.78	1.57	1.53	1.54	1.64	1.54	1.72	1.95	1.92	1.58	1.36	1.22	1.89	1.67	1.5	1.47
	GP SV	1.7	1.55	1.49	1.48	1.6	1.54	1.68	1.89	1.8	1.54	1.33	1.17	1.81	1.63	1.47	1.42

Notes: The table shows the continuous ranked probability score (`CRPS`) and quantile-weighted version for the left (`qwCRPS-L`) and right (`qwCRPS-R`) tail, together with the mean absolute error (`MAE`) scores for key variables and different models for horizon $h = \{1, 3, 12, 24\}$. Results in the left panel are based on the *small* (**S**) information set; on the right panel on the *large* (**L**) information set. Grey shaded rows indicate raw losses for the benchmark, all others are shown as ratios relative to these values (red indicates worse and blue superior performance).

4.2 Monthly Forecasting Results

The monthly variables we forecast are the key macroeconomic variables constantly monitored by policy makers and market participants. Specifically, our target indicators are industrial production growth (year-on-year) and consumer prices inflation (year-on-year). A timely analysis of the path of these two indicators is also central in the assessment of the sustainability of a country's fiscal situation. In fact, Industrial Production growth is commonly considered as a higher frequency proxy for GDP growth, which appears at the denominator of both fiscal ratios considered in the previous sub-section. Consumer prices, on the other hand, influence the real public cost of servicing the existing debt.

Tables 2 and 3 report CRPSs/MAEs and quantile scores, respectively, for the various models we consider and forecast horizons from 1- to 24-month ahead. Also in this case, the more sophisticated time-varying and nonparametric models are in general unable of over-perform the *small* BVAR-SV. For industrial production, both the *small* and the *large* versions of the non-parametric BART and GP-VAR models provide accurate point forecasts and a precise quantification of left-tail risk. The superiority, however, is only apparent for one-moth ahead predictions, while at all other horizons the metrics are virtually identical to those attained by the benchmark model.

Most of the cells in the bottom blocks of the tables are red, which indicates that for inflation the metrics computed for most models/horizons are significantly inferior to those computed for

Table 5: Evaluation based on Quantile Scores: Growth rate of Monthly Indicators.

S

		QS10				QS25				QS75				QS90			
	Horizon	1	3	12	24	1	3	12	24	1	3	12	24	1	3	12	24
Industrial Production	BVAR SV	1.44	1.39	1.37	1.45	1.93	1.82	1.79	1.9	1.5	1.6	1.7	1.79	0.67	1.15	1.3	1.34
	BVAR SV−t		**0.95°**	**0.99**	**0.98**	1	**0.99**	**0.99**	1.02	1.04	**0.99**	**0.99**	1	1.2	**0.96'**	**0.97'**	**0.97**
	TVP SV	0.97	1.01	1.05	1.08	1.02	1.02	1.05	1.07	1.08	1	1.03	1.02	1.08	1	1.04	1.06
	BART hom	0.99	1.12	1.08	1.04	1.15	1.25	1.03	1.01	1.81	1.35	1.06	1.04	3.32	1.3	1.06	1.02
	BART SV	0.93	1	1.01	1	0.94	1.01	1.01	**0.99**	*0.99*	1.06	1.01	1.01	1.38	1.01	1	1
	GP hom	0.93	0.99	1.02	0.99	0.93	1.01	1.01	0.99	1.08	1.07	1.03	1	1.62	1.04	1.01	1
	GP SV	**0.82**	1.01	1.03	1.01	**0.90**	1.02	1.02	1	1.04	1.03	1.04	1.01	1.57	1.01	1.04	1.01
Consumer Prices	BVAR SV	0.09	0.12	0.19	0.21	0.14	0.2	0.37	0.44	0.14	0.24	0.49	0.66	0.1	0.17	0.35	0.47
	BVAR SV−t	1.01	0.98	1.03	1.12	1.01	1.02	1.04	1.04	1	1.01	1.01	1.02	**0.99**	1	1.05	1.07
	TVP SV	1	1	1.21	1.63	0.97	1.01	1.13	1.25	**0.98**	**0.89**	1.03	1.11	1.03	**0.88**	0.94	1.23
	BART hom	1.15	1.29	1.8	2.31	1.22	1.33	1.37	1.58	1.33	1.22	1.12	1.06	1.3	1.21	1.1	*0.96*
	BART SV	1.18	1.12	1.44	2.08	1.17	1.16	1.12	1.42	1.24	1.08	**0.88**	1.04	1.27	1.09	**0.86**	1.08
	GP hom	1.33	1.23	1.88	2.61	1.67	1.5	1.54	1.85	1.78	1.33	1.26	1.1	1.64	1.23	1.22	0.97
	GP SV	1.29	1.22	1.72	2.46	1.56	1.41	1.44	1.84	1.58	1.18	1.23	1.12	1.44	1.1	1.18	0.98

L

		QS10				QS25				QS75				QS90			
	Horizon	1	3	12	24	1	3	12	24	1	3	12	24	1	3	12	24
Industrial Production	BVAR SV	0.99	1.08	1.13	1.09	1.05	1.07	1.08	1.03	0.98	1.03	1.09	1.06	*1.14*	1.08	1.14	1.12
	BVAR SV−t	1.03	1.06	1.09	1.07	1.04	1.06	1.06	1.03	1	1.07	1.07	1.06	1.17	1.07	1.1	1.09
	TVP SV	1.02	1.09	1.17	1.24	1.07	1.07	1.1	1.13	1.01	1.06	1.13	1.18	1.22	1.09	1.19	1.29
	BART hom		*0.98*	*1.01*	1	*0.92*	*1.00*	*1.01*	1.01	1.04	1.06	*1.02*	**1.00**	1.67	1.09	1	*0.99*
	BART SV	0.88	1.05	1.07	1.05	0.95	1.05	1.04	1	**0.97**	*1.00*	1.04	1.02	1.38	*1.03*	1.05	1.04
	GP hom	0.94	1	1.02	*0.99*	0.93	1	1.01	*0.99*	1.1	1.06	1.03	1.01	1.7	1.06	1.01	1
	GP SV	*0.88*	1.07	1.09	1.05	0.95	1.04	1.05	1.02	1	1	1.06	1.02	1.45	1.07	1.09	1.06
Consumer Prices	BVAR SV	0.98	*0.99*	*1.13*	*1.40*	0.99	1.06	*1.10*	*1.28*	*1.00*	1	1.02	1.08	*1.01*	0.99	0.94	1.1
	BVAR SV−t	**0.96**	0.99	1.15	1.46	0.98	1.06	1.12	1.35	1.05	1.03	1.06	1.13	1.01	1	1.02	1.18
	TVP SV	1.04	1.06	1.39	2.48	0.99	*1.03*	1.21	1.51	1.03	*0.94*	1.03	1.15	1.06	*0.89'*	*0.88*	1.12
	BART hom	1.13	1.15	1.56	2.21	1.15	1.19	1.2	1.44	1.35	1.18	1.01	1.14	1.42	1.29	1.03	1.2
	BART SV	1.25	1.19	1.51	2.11	1.24	1.17	1.12	1.36	1.37	1.09	*0.91*	*1.01*	1.45	1.18		1.03
	GP hom	1.35	1.26	2.19	3.2	1.58	1.58	1.79	2.1	2	1.56	1.3	1.12	1.88	1.47	1.23	1.01
	GP SV	1.33	1.26	2.16	3.1	1.55	1.59	1.74	2.02	1.88	1.52	1.27	1.08	1.71	1.45	1.2	**0.95**

Notes: The table shows the quantile scores for quantiles $\tau = \{0.10, 0.25, 0.75, 0.90\}$ for key variables and different models for horizon $h = \{1, 3, 12, 24\}$. Results in the left panel are based on the *small* (**S**) information set; on the right panel on the *large* (**L**) information set. Grey shaded rows indicate raw losses for the benchmark, all others are shown as ratios relative to these values (red indicates worse and blue superior performance).

the *small* BVAR-SV. The advantage of the most simple model in this case is clear, especially when compared to non-parametric models. The only model that can rival the benchmark is the *small* BVAR-SV with the addition of time varying coefficients, but also in this case the its performance deteriorates as the forecasting horizon increases. As before, considering larger information sets generates often poorer results.

To summarize, also when the focus is on the monthly frequency, and the target variables are industrial production growth and consumer price inflation, the simple BVAR-SV tends to be the most reliable forecasting model. Furthermore, there is no evidence of benefits when enlarging the set of predictors included in the model.

Time Variation in Prediction Accuracy

Looking at the upper rows of all panels of Figure 6, it is clear that the predictive densities resulting from all models, at all horizons, achieve virtually the same level of accuracy for the whole hold-out period preceding the Covid-19 crisis. During 2020 and 2021, the homoskedastic BVAR and BART-VAR models fared significantly worse than the competitors at shorter horizons, while GP-VAR, both homoskedastic and heteroskedastic, and BART-VAR-SV provided more reliable left-tail forecasts. For 1-year ahead or longer horizon forecasts, the differences between different models are minimal.

When focusing on inflation forecasts, the superiority of the BVAR-SV is apparent at all horizons. The advantage of the BVAR-SV emerges in the close-to-zero inflation period after

2013, and keeps widening at least until 2020. During the Covid-19 pandemics, the gap is partially reabsorbed only for very long horizons, thanks to the ability of more flexible models to provide better right tail forecasts.

Figure 7 reports the shortfall and longrise produced by the different models and the actual realized values, for both Industrial Production growth and Consumer Price inflation.

As in the case of the quarterly fiscal variables, the differences in the risk measures provided by different models before 2020 are not substantial, especially for Industrial Production growth. Clearly, all models were unable to foresee the possibility of a drop in production as large as the one observed in 2020Q1, but some models, such as the BVAR-SV, were able to anticipate one quarter in advance the equally large increase that followed. Specifications featuring stochastic volatility adapted to the more risky environment maintaining large absolute values of both shortfall and longrise for the rest of the hold-out period.

The plots showing the time variation of longrise and shortfall for inflation are interesting because the last decade has been abundant of both positive and negative tail events. Focusing on one-year ahead predictions, it can be seen that, for many of the models, the actual realization between 2013 and 2016 was too low for too long when compared to the 5 % shortfall inflation. The shortfalls provided by the BVAR-SV and the BVAR-SV-TVP in that period appear more sensible, which justifies the good performance of these models in terms of `qwCRPS-L` seen in 6.

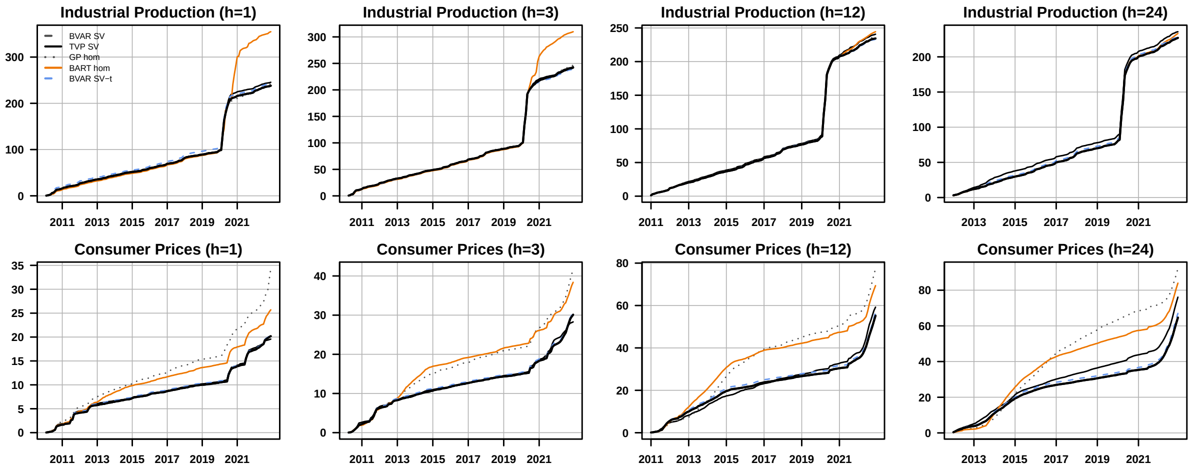
Another crucial event on which it is worth focusing is the surge in inflation observed at the end of 2022. The BVAR-SV clearly fared well compared to many other models, when looking at 1-quarter ahead predictions. On the other hand, more flexible models, such as BART-VAR-SV were able to adapt more timely to the extreme inflation rates experienced in 2022.

5. Concluding Remarks

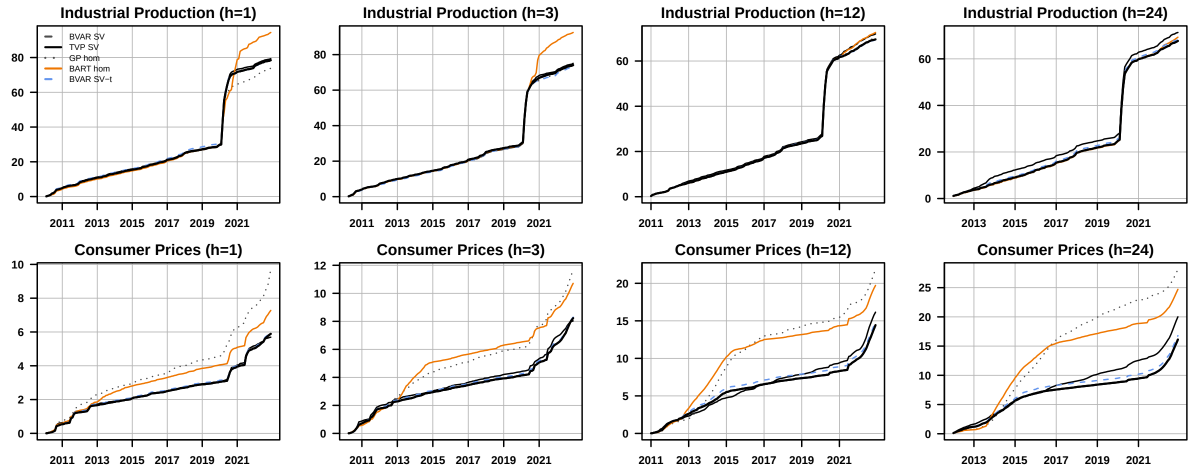
We have investigated the empirical performance of various econometrics methods for predicting tail risks for the Italian economy.

After an overview of Bayesian VARs with stochastic volatility, Bayesian additive regression trees, and Gaussian processes, which are among the most suited methods according to the recent econometric literature, we have assessed their point, density and, specifically, tail predictive performance for GDP growth, inflation, debt to GDP and deficit-to-GDP ratios, key economic and fiscal indicators for Italy.

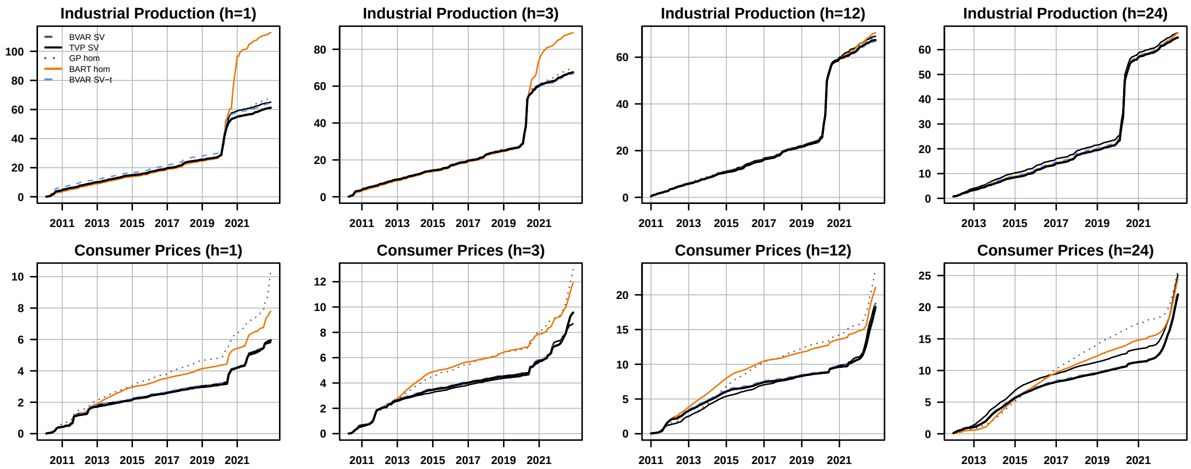
Figure 6: Cumulative Scores Over the Hold-Out Sample.



(a) CRPS



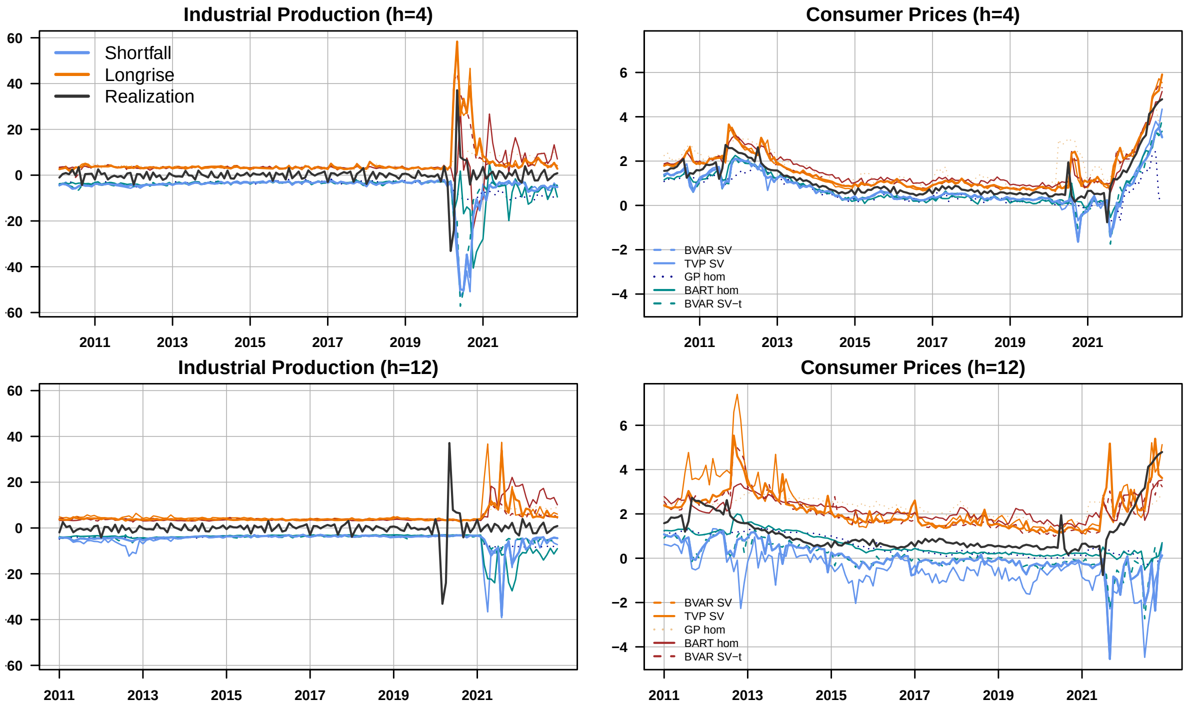
(b) qwCRPS-L



(c) qwCRPS-R

Notes: The figure shows the time series evolution of continuous ranked probability score (CRPS) and quantile-weighted version for the left (qwCRPS-L) and right (qwCRPS-R) tail for different variables and models at horizon $h = \{1, 3, 12, 24\}$. Results are based on the *small* information set.

Figure 7: Expected Shortfall and Longrise Over Time.



Notes: The figure shows the time series evolution of the 5 percent expected shortfall (blue lines) and longrise (orange lines) at horizon $h = \{1, 4\}$. Results are based on the *small* information set. Black bold line denotes realization of outcome variable.

It turns out that BVAR-SV performs overall particularly well, though for some variables and horizons some of the more sophisticated models do better for the 10th or 90th percentiles, but not systematically so across variables and forecast horizons.

We have then used the models to also predict expected shortfalls and longrises for the variables of interest, and the probability of policy relevant events, such as negative growth or deficit-to-GDP ratios above 3%, providing an additional tool for monitoring risks to the Italian economy.

Finally, we have repeated the analysis for a set of monthly indicators, finding results broadly in line with those for the quarterly variables.

References

Adams Patrick A, Adrian Tobias, Boyarchenko Nina and Giannone Domenico (2021) 'Forecasting macroeconomic risks'. *International Journal of Forecasting* Vol. 37(3), pp. 1173–1191.

Adrian Tobias, Boyarchenko Nina and Giannone Domenico (2019) 'Vulnerable growth'. *American Economic Review* Vol. 109(4), pp. 1263–1289.

Adrian Tobias, Grinberg Federico, Liang Nellie, Malik Sheheryar and Yu Jie (2022) 'The term structure of growth-at-risk'. *American Economic Journal: Macroeconomics* Vol. 14, pp. 283–323.

Aprigliano Valentina (2020) 'A large Bayesian VAR with a block-specific shrinkage: A forecasting application for Italian industrial production'. *Journal of Forecasting* Vol. 39(8), pp. 1291–1304.

Carriero Andrea, Clark Todd E and Marcellino Massimiliano (2016) 'Common drifting volatility in large Bayesian VARs'. *Journal of Business & Economic Statistics* Vol. 34(3), pp. 375–390.

Carriero Andrea, Clark Todd E and Marcellino Massimiliano (2022) 'Nowcasting tail risks to economic activity with many indicators'. *Journal of Applied Econometrics* Vol. 37(5), pp. 843–866.

Carriero Andrea, Clark Todd E and Marcellino Massimiliano (2024) 'Capturing Macroeconomic Tail Risks with Bayesian Vector Autoregressions'. *Journal of Money, Credit and Banking* Vol. forthcoming.

Carriero Andrea, Clark Todd E, Marcellino Massimiliano and Mertens Elmar (2022) 'Addressing COVID-19 outliers in BVARs with stochastic volatility'. *Review of Economics and Statistics*, pp. 1–38.

Carvalho Carlos M, Polson Nicholas G and Scott James G (2010) 'The horseshoe estimator for sparse signals'. *Biometrika* Vol. 97(2), pp. 465–480.

Chan Joshua CC (2022) 'Comparing stochastic volatility specifications for large Bayesian VARs'. *Journal of Econometrics* Vol. forthcoming.

Chipman Hugh A., George Edward I. and McCulloch Robert E. (2010) 'BART: Bayesian additive regression trees'. *The Annals of Applied Statistics* Vol. 4(1), pp. 266–298.

Clark Todd E (2011) 'Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility'. *Journal of Business & Economic Statistics* Vol. 29(3), pp. 327–341.

Clark Todd E, Huber Florian, Koop Gary and Marcellino Massimiliano (2024a) 'Forecasting US inflation using Bayesian nonparametric models'. *Annals of Applied Statistics* Vol. in-press.

Clark Todd E, Huber Florian, Koop Gary, Marcellino Massimiliano and Pfarrhofer Michael (2023) 'Tail forecasting with multivariate bayesian additive regression trees'. *International Economic Review* Vol. 64(3), pp. 979–1022.

Clark Todd E, Huber Florian, Koop Gary, Marcellino Massimiliano and Pfarrhofer Michael (2024b) 'Investigating Growth-at-Risk Using a Multicountry Non-parametric Quantile Factor Model'. *Journal of Business & Economic Statistics*(forthcoming).

Cochrane John H (2022) 'Fiscal histories'. *Journal of Economic Perspectives* Vol. 36(4), pp. 125–146.

D'Agostino Antonello, Gambetti Luca and Giannone Domenico (2013) 'Macroeconomic forecasting and structural change'. *Journal of Applied Econometrics* Vol. 28(1), pp. 82–101.

Delle Monache Davide, De Polis Andrea and Petrella Ivan (2023) 'Modeling and forecasting macroeconomic downside risk'. *Journal of Business & Economic Statistics*, pp. 1–27.

Frühwirth-Schnatter Sylvia and Wagner Helga (2010) 'Stochastic model specification search for Gaussian and partial non-Gaussian state space models'. *Journal of Econometrics* Vol. 154(1), pp. 85–100.

Galbraith John W. and van Norden Simon (2019) 'Asymmetry in unemployment rate forecast errors'. *International Journal of Forecasting* Vol. 35(4), pp. 1613–1626.

Giacomini Raffaella and Komunjer Ivana (2005) 'Evaluation and combination of conditional quantile forecasts'. *Journal of Business & Economic Statistics* Vol. 23(4), pp. 416–431.

Gneiting Tilmann and Raftery Adrian E (2007) 'Strictly proper scoring rules, prediction, and estimation'. *Journal of the American statistical Association* Vol. 102(477), pp. 359–378.

Gneiting Tilmann and Ranjan Roopesh (2011) 'Comparing density forecasts using threshold- and quantile-weighted scoring rules'. *Journal of Business & Economic Statistics* Vol. 29(3), pp. 411–422.

Hauzenberger Niko, Huber Florian, Koop Gary and Onorante Luca (2022) 'Fast and flexible Bayesian inference in time-varying parameter regression models'. *Journal of Business & Economic Statistics* Vol. 40(4), pp. 1904–1918.

Hauzenberger Niko, Huber Florian, Marcellino Massimiliano and Petz Nico (2024a) 'Gaussian process vector autoregressions and macroeconomic uncertainty'. *Journal of Business & Economic Statistics*(forthcoming).

Hauzenberger Niko, Marcellino Massimiliano, Pfarrhofer Michael and Stelzer Anna (2024b) 'Nowcasting with mixed frequency data using Gaussian processes'. *arXiv* Vol. 2402.10574.

Huber Florian, Koop Gary, Onorante Luca, Pfarrhofer Michael and Schreiner Josef (2023) 'Nowcasting in a pandemic using non-parametric mixed frequency VARs'. *Journal of Econometrics* Vol. 232(1), pp. 52–69.

Huber Florian and Rossini Luca (2022) 'Inference in Bayesian additive vector autoregressive tree models'. *The Annals of Applied Statistics* Vol. 16(1), pp. 104–123.

Kastner Gregor and Huber Florian (2020) 'Sparse Bayesian vector autoregressions in huge dimensions'. *Journal of Forecasting* Vol. 39(7), pp. 1142–1165.

Makalic Enes and Schmidt Daniel F (2015) 'A simple sampler for the horseshoe estimator'. *IEEE Signal Processing Letters* Vol. 23(1), pp. 179–182.

Manzan Sebastiano (2015) 'Forecasting the distribution of economic variables in a data-rich environment'. *Journal of Business and Economic Statistics* Vol. 33(1), pp. 144–164.

Marcellino Massimiliano and Pfarrhofer Michael (2024) 'Bayesian Nonparametric Methods for Macroeconomic Forecasting'. *Handbook of Macroeconomic Forecasting.* Ed. by Michael P. Clements and Ana Beatriz Galvao. Vol. forthcoming. Edward Elgar Publishing Ltd.

McCracken Michael W and Ng Serena (2016) 'FRED-MD: A Monthly Database for Macroeconomic Research'. *Journal of Business & Economic Statistics* Vol. 34(4), pp. 574–589.

Pfarrhofer Michael (2022) 'Modeling tail risks of inflation using unobserved component quantile regressions'. *Journal of Economic Dynamics and Control* Vol. 143, p. 104493.

Plagborg-Møller Mikkel, Reichlin Lucrezia, Ricco Giovanni and Hasenzagl Thomas (2020) 'When is growth at risk?' *Brookings Papers on Economic Activity*, pp. 167–229.

Reis Ricardo (2022) 'Debt revenue and the sustainability of public debt'. *Journal of Economic Perspectives* Vol. 36(4), pp. 103–124.

Sax Christoph and Eddelbuettel Dirk (2018) 'Seasonal Adjustment by X-13ARIMA-SEATS in R'. *Journal of Statistical Software* Vol. 87(11), pp. 1–17.

Schnabel Isabel (2023) *The risks of stubborn inflation.* Speech at the Euro50 Group conference on "New challenges for the Economic and Monetary Union in the post-crisis environment". https://www.ecb.europa.eu/press/key/date/2023/html/ecb.sp230619_1~2c0bdf2422.en.html.

A. Data

All series were gathered from the sources listed below, including the Federal Reserve Economic Database (FRED) compiled by the St. Louis Federal Reserve (McCracken and Ng, 2016), Refinitiv, Eurostat, the Istituto Nazionale di Statistica (Istat), and the Statistical Data Warehouse (SDW) of the European Central Bank (ECB). If necessary, series are seasonally adjusted with the X-13ARIMA-SEATS model (Sax and Eddelbuettel, 2018). All series are approximately stationary.

Table A1: Quarterly Dataset.

Variable	Details	Source	$\mathcal{M}_S$	$\mathcal{M}_M$	$\mathcal{M}_L$	*tcode*
$\texttt{gov\_debt\_ratio}_t$	government debt (consolidated, as % of GDP)	SDW ECB	✓	✓	✓	2
$\texttt{gov\_deficit\_ratio}_t$	government primary deficit or surplus (as % of GDP)	SDW ECB	✓	✓	✓	2
$\texttt{rgdp}_t$	real gross domestic product	FRED	✓	✓	✓	4
$\texttt{hicp}_t$	harmonized consumer price index, overall index	SDW ECB	✓	✓	✓	4
$\texttt{unrate}_t$	unemployment rate	SDW ECB	✓	✓	✓	2
$\texttt{ltir}_t$	long-term interest rate for convergence purpuses, 10 years maturity	SDW ECB	✓	✓	✓	2
$\texttt{sovciss}_t$	composite indicator of systematic stress	SDW ECB		✓	✓	0
$\texttt{baltic}_t$	Baltic Dry Index	Refinitiv		✓	✓	0
$\texttt{ip\_constr}_t$	index of production construction	Istat		✓	✓	1
$\texttt{prodnxt3m}_t$	production, next 3 months, balance	Istat		✓	✓	0
$\texttt{oilprice}_t$	crude oil prices, Brent, Europe	FRED		✓	✓	0
$\texttt{gfc}_t$	gross fixed capital transformation	SDW ECB			✓	0
$\texttt{imp}_t$	imports of goods and services	SDW ECB			✓	0
$\texttt{exp}_t$	exports of goods and services	SDW ECB			✓	0
$\texttt{wage}_t$	wage and salries	SDW ECB			✓	0
$\texttt{hours}_t$	hours worked, total employment	SDW ECB			✓	0
$\texttt{labprod\_ph}_t$	labor productivity, per hours worked	SDW ECB			✓	0
$\texttt{ea\_indpro}_t$	euro area 19 industrial production (including construction)	SDW ECB			✓	0
$\texttt{hicp\_core}_t$	harmonized consumer price index, excluding energy and food	SDW ECB			✓	0
$\texttt{ea\_indpro\_exconstr}_t$	euro area 19 industrial production (excluding construction)	SDW ECB			✓	0
$\texttt{ea\_hicp}_t$	euro area 19 harmonized consumer price index, overall index	SDW ECB			✓	0
$\texttt{ip\_ger}_t$	industrial production of Germany (excluding construction)	FRED			✓	0
$\texttt{ip\_us}_t$	industrial production of the US (excluding construction)	FRED			✓	0
$\texttt{ea\_stir}_t$	euro area 19 short-term interest rates, 3 months maturity	FRED			✓	2
$\texttt{cos\_it}_t$	consumer sentiment index for Italy	FRED			✓	3

Notes: $\mathcal{M}_S$ denotes the *small* information set, $\mathcal{M}_M$ is the *medium* information set, and $\mathcal{M}_L$ is the *large* information set. Transformation codes (*tcode*): 0 = level, 1 = 100× log-differences, 2 = differences, 3 = log-level, 4 = annualized differences.

Table A2: Monthly Dataset.

Variable	Details	Source	$\mathcal{M}_S$	$\mathcal{M}_L$	*tcode*
$\texttt{ip\_exclconstr}_t$		Istat	✓	✓	1
$\texttt{unrate}_t$	unemployment rate	SDW ECB	✓	✓	0
$\texttt{hicp}_t$	harmonized consumer price index, overall index	SDW ECB	✓	✓	0
$\texttt{ltir\_spread}_t$	$\texttt{ltir}_t - \texttt{ltir\_ger}_t$	FRED	✓	✓	0
$\texttt{ciss}_t$	composite indicator of systematic stress (Italy)	SDW ECB	✓	✓	0
$\texttt{exp}_t$	exports (goods)	FRED		✓	2
$\texttt{hours}_t$	hours worked per employee (excluding short-time work allowance)	Istat		✓	1
$\texttt{hicp\_core}_t$	harmonized consumer price index, excluding energy and food	SDW ECB		✓	0
$\texttt{ltir}_t$	long-term government bond yields: 10-year benchmark for Italy	FRED		✓	0
$\texttt{prodnxt3m}_t$	production, next 3 months, balance	Istat		✓	0
$\texttt{esi}_t$	economic sentiment indicator	Eurostat		✓	3
$\texttt{baltic}_t$	Baltic Dry Index	Refinitiv		✓	3
$\texttt{ordnxt3m}_t$	order books, next 3 months	Istat			0
$\texttt{totord\_manuf}_t$	total order books, balance	Istat			0
$\texttt{toteng}_t$	total energy consumption	Istat			0
$\texttt{ip\_constr}$	index of production construction	Istat			1
$\texttt{ip\_ger}_t$	industrial production of Germany (excluding construction)	FRED			0
$\texttt{ltir\_ger}_t$	long-term government bond yields: 10-year benchmark for Germany	FRED			0
$\texttt{ip\_us}_t$	industrial production of the US (excluding construction)	FRED			0
$\texttt{sovciss}_t$	sovereign composite indicator of systematic stress (Italy)	SDW ECB			0
$\texttt{usdeur}_t$	US Dollar to Euro exchange rate	SDW ECB			0

Notes: $\mathcal{M}_S$ denotes the *small* information set and $\mathcal{M}_L$ is the *large* information set. Transformation codes (*tcode*): 0 = level, 1 = log-differences, 2 = differences, 3 = log-level.

B. Additional Results

Table B1: Evaluation Based on Variants of CRPS/MAE: Small and Large Information Sets.

S

Variable	Model	CRPS 1	CRPS 4	CRPS 8	CRPS−L 1	CRPS−L 4	CRPS−L 8	CRPS−R 1	CRPS−R 4	CRPS−R 8	MAE 1	MAE 4	MAE 8
Government Debt Ratio	BVAR SV	1.00	1.02	1.23	1.05	1.07	1.22	0.97	**0.98**	1.22	1.00	1.06	1.22
	BVAR SV−t	1.20	1.23	1.38	1.20	1.36	1.46	1.22	1.16	1.31	1.07	1.29	1.39
	TVP SV	1.04	*0.99*	*1.21*	1.10	*1.00*	*1.21*	0.99	0.98	*1.19*	1.01	*1.02*	*1.22*
	BART hom	0.96	1.24	1.37	0.99	1.35	1.41	0.95	1.15	1.31	0.89	1.24	1.43
	BART SV	*0.94*	1.23	1.39	*0.96*	1.30	1.40	*0.93*	1.17	1.35	*0.88*	1.26	1.47
	GP hom	1.01	1.27	1.42	1.04	1.37	1.41	0.99	1.19	1.41	0.96	1.30	1.46
	GP SV	1.00	1.25	1.43	1.01	1.30	1.40	1.00	1.21	1.42	0.96	1.31	1.46
Deficit Ratio	BVAR SV	1.11	1.02	0.98	1.15	1.00	0.97°	1.06	1.04	0.99	1.09	1.03	0.98
	BVAR SV−t	1.21	1.02	1.01	1.34	1.01	0.99	1.10	1.04	1.02	1.14	1.03	1.03
	TVP SV	*0.98*	*1.01*	0.99	*0.97*	*0.99*	0.99	0.99	1.04	0.99	*0.96*	*0.98*	0.96'
	BART hom	1.05	1.04	0.95°	1.15	1.08	0.97	0.95	0.99	0.91*	0.98	1.04	0.93°
	BART SV	0.99	1.03	0.96°	1.06	1.07	0.98	*0.91*	0.98	0.92*	0.98	1.04	0.95'
	GP hom	1.10	1.05	*0.92**	1.23	1.08	**0.94°**	0.96	0.99	*0.88°*	1.02	1.05	0.92'
	GP SV	1.07	1.03	0.92°	1.18	1.09	0.95'	0.97	*0.95*	0.88°	1.01	1.04	*0.92'*
Growth	BVAR SV	1.01	0.96	0.95	0.92	**0.94**	**0.93'**	1.12	0.96	0.95	1.02	0.98	0.98
	BVAR SV−t	1.08	**0.95**	0.95	0.97	0.94	0.93	1.22	**0.96**	0.96	1.12	0.99	0.98
	TVP SV	**0.97**	0.97	0.95	**0.92'**	0.95	0.94	1.04	0.98	0.96	0.98	*0.98*	0.98
	BART hom	1.12	1.00	0.96'	1.08	0.99	0.95'	1.14	1.01	0.96'	1.11	1.07	1.00
	BART SV	1.00	1.01	0.97'	0.99	1.01	0.96'	1.01	1.00	0.96'	1.02	1.06	0.98
	GP hom	1.03	1.01	0.95'	0.97	0.99	0.94'	1.08	1.03	**0.94**	*0.98*	1.05	0.99
	GP SV	1.00	1.01	**0.95'**	0.98	1.00	0.94'	*1.01*	1.01	0.94	0.99	1.05	**0.98**
Inflation	BVAR SV	0.98	0.96'	*0.98*	**0.98**	0.96	0.99	0.98	0.96°	0.97°	0.99	0.96'	0.97
	BVAR SV−t	0.98	1.00	1.00	0.99	1.03	1.02	**0.97**	0.99	0.98'	**0.98**	0.98	0.98
	TVP SV	**0.97**	**0.94°**	0.98	0.98	0.93	0.99	0.97	**0.94°**	0.97	0.99	**0.93°**	0.99
	BART hom	1.09	1.00	1.02	1.07	0.95	1.00	1.09	1.02	1.02	1.09	0.97	0.97'
	BART SV	1.13	0.97	1.01	1.10	**0.92'**	*0.98*	1.15	1.01	1.02	1.14	0.95'	*0.97°*
	GP hom	1.29	1.02	1.02	1.21	0.97	1.00	1.34	1.03	1.03	1.31	0.99	0.99
	GP SV	1.28	1.01	1.03	1.26	0.98	1.01	1.29	1.03	1.03	1.34	1.00	0.99

L

Variable	Model	CRPS 1	CRPS 4	CRPS 8	CRPS−L 1	CRPS−L 4	CRPS−L 8	CRPS−R 1	CRPS−R 4	CRPS−R 8	MAE 1	MAE 4	MAE 8
Government Debt Ratio	BVAR SV	**0.86'**	**0.97**	*1.06*	**0.88'**	**0.92'**	*1.05*	**0.86°**	1.01	1.07	**0.82'**	**0.95**	*1.10*
	BVAR SV−t	1.12	1.18	1.22	1.20	1.17	1.21	1.06	1.18	1.20	1.04	1.24	1.30
	TVP SV	0.99	1.10	1.22	1.04	1.08	1.25	0.97	1.11	1.20	0.94	1.09	1.21
	BART hom	1.04	1.28	1.40	1.08	1.38	1.45	1.02	1.20	1.34	0.97	1.31	1.48
	BART SV	0.99	1.23	1.37	1.01	1.25	1.38	0.99	1.21	1.33	0.95	1.30	1.49
	GP hom	1.07	1.32	1.41	1.08	1.40	1.46	1.08	1.25	1.35	1.01	1.38	1.53
	GP SV	1.05	1.28	1.38	1.06	1.28	1.39	1.06	1.26	1.35	1.00	1.36	1.52
Deficit Ratio	BVAR SV	0.93	1.04	1.07	0.92	*1.01*	1.05	0.94	1.09	1.12	0.89	*1.01*	1.10
	BVAR SV−t	1.24	1.15	1.22	1.23	1.08	1.15	1.26	1.25	1.32	1.24	1.15	1.27
	TVP SV	**0.88'**	1.12	1.21	**0.90'**	1.08	1.14	**0.89**	1.19	1.34	**0.80'**	1.09	1.25
	BART hom	1.09	1.09	0.95°	1.22	1.12	0.97	0.97	1.05	0.90*	1.03	1.07	0.94'
	BART SV	1.02	1.06	0.96'	1.11	1.11	0.99	0.93	0.99	0.90*	1.01	1.09	0.97
	GP hom	1.09	1.04	*0.93°*	1.24	1.09	*0.96*	0.95	0.97	*0.86**	1.02	1.02	0.91'
	GP SV	1.06	*1.02*	0.93°	1.15	1.07	0.97	0.96	**0.94**	0.87*	1.03	1.02	*0.90'*
Growth	BVAR SV	*1.00*	1.00	1.02	*0.99*	1.01	1.02	1.01	1.00	1.02	1.00	0.99	1.03
	BVAR SV−t	1.07	*0.99*	1.03	1.05	0.98	1.02	1.10	*0.99*	1.04	1.06	0.98	1.07
	TVP SV	1.13	1.11	1.12	1.12	1.10	1.14	1.15	1.14	1.12	1.15	1.01	1.07
	BART hom	1.04	1.00	*0.95*	1.01	0.99	*0.95*	1.05	1.02	*0.95*	1.00	1.06	*0.99*
	BART SV	1.01	1.05	0.99	1.03	1.06	0.99	0.99	1.04	1.00	*0.98*	1.07	0.99
	GP hom	1.07	1.03	0.96	1.03	1.01	0.96	1.08	1.04	0.96	1.00	1.08	1.00
	GP SV	1.02	1.02	0.97	1.04	1.01	0.97	*0.99*	1.02	0.97	1.00	1.07	1.00
Inflation	BVAR SV	1.04	1.03	0.98	1.05	1.04	0.96	1.03	1.03	1.00	1.09	1.02	1.00
	BVAR SV−t	1.22	1.09	0.97	1.29	1.10	0.97	1.16	1.08	0.98	1.25	1.04	0.99
	TVP SV	1.06	1.03	1.00	1.09	1.02	1.05	1.04	1.03	*0.97*	1.12	1.04	1.08
	BART hom	1.42	1.04	1.01	1.37	0.98	0.97'	1.45	1.06	1.03	1.46	1.00	0.96°
	BART SV	1.45	*0.99*	**0.97**	1.46	*0.94*	**0.90***	1.44	1.03	1.02	1.61	1.00	**0.96***
	GP hom	1.46	1.01	0.98	1.34	0.96	0.93°	1.54	1.04	1.01	1.49	1.00	0.96'
	GP SV	1.39	1.00	0.98	1.37	0.96	0.93°	1.40	*1.02*	1.00	1.49	*0.99*	0.96'

Notes: The table shows the continuous ranked probability score (CRPS) and quantile-weighted version for the left (qwCRPS-L) and right (qwCRPS-R) tail, together with the mean absolute error (MAE) scores for key variables and different models for horizon $h = \{1, 4, 8\}$ as ratio to the BVAR-SV model based on the *medium* (**M**) information set. Results in the left panel are based on the *small* (**S**) information set; on the right panel on the *large* (**L**) information set.

Table B2: Evaluation Based on Quantile Scores: Small and Large Information Sets.

S

		QS10			QS25			QS75			QS90		
	Horizon	1	4	8	1	4	8	1	4	8	1	4	8
Government Debt Ratio	BVAR SV	1.24	1.04	1.15	1.05	1.13	1.23	0.97	**0.94**	1.16	0.93	**0.94**	1.26
	BVAR SV−t	1.65	1.54	1.47	1.15	1.42	1.50	1.21	1.13	1.27	1.54	1.05	1.25
	TVP SV	1.41	*1.03*	*1.11*	1.09	*1.01*	*1.22*	1.00	0.97	*1.13*	**0.88**	0.97	*1.24*
	BART hom	1.25	1.64	1.36	1.01	1.40	1.41	0.89	1.11	1.26	1.18	1.07	1.26
	BART SV	*1.09*	1.42	1.25	*1.00*	1.35	1.39	*0.88*	1.14	1.30	1.11	1.11	1.29
	GP hom	1.24	1.60	1.30	1.05	1.41	1.39	0.91	1.16	1.35	1.18	1.12	1.44
	GP SV	1.12	1.29	1.26	1.02	1.34	1.40	0.93	1.18	1.36	1.19	1.16	1.47
Deficit Ratio	BVAR SV	1.06	0.97	0.97'	1.28	1.00	0.96'	1.04	1.05	1.02	1.05	1.05	0.95'
	BVAR SV−t	1.55	**0.95'**	0.97°	1.44	1.04	0.97'	1.04	1.05	1.03	1.05	1.01	0.97
	TVP SV	*1.01*	0.98	0.99	*0.96*	*1.00*	0.98	1.01	1.09	1.01	1.04	1.08	0.96
	BART hom	1.33	1.09	0.96	1.23	1.10	0.98	0.90	1.00	0.89°	0.93	0.86	0.80°
	BART SV	1.08	1.06	0.98	1.14	1.09	1.00	*0.87'*	0.99	0.90°	**0.83°**	0.81°	0.81'
	GP hom	1.42	1.10	**0.95°**	1.33	1.11	**0.94'**	0.91	0.97	0.85°	0.88	0.86	**0.79'**
	GP SV	1.30	1.11	0.96'	1.29	1.12	0.95	0.89'	*0.92*	*0.83°*	0.98	**0.79'**	0.82'
Growth	BVAR SV	**0.84°**	*0.92*	0.91	**0.90'**	**0.93**	**0.92'**	1.14	1.01	0.97	1.26	0.90	0.92
	BVAR SV−t	0.86°	0.92	0.90	0.92	0.93	0.92	1.23	*1.00*	0.97	1.38	**0.87'**	0.92
	TVP SV	0.86°	0.93	0.92	0.91°	0.95	0.93	1.04	1.01	0.97	1.13	0.92	0.93
	BART hom	1.04	0.92	0.91'	1.09	0.99	0.95	1.05	1.09	0.96	1.37	0.88	0.91
	BART SV	0.96	0.96	0.92	0.99	1.01	0.96	*1.00*	1.03	0.98	1.07	0.91	0.92
	GP hom	0.94	0.92	**0.90'**	0.97	0.99	0.95	1.04	1.12	*0.95*	1.30	0.91	**0.90**
	GP SV	0.95°	0.94	0.91'	0.99	1.00	0.94	1.02	1.08	0.96	*1.04*	0.90	0.91
Inflation	BVAR SV	0.97	0.92	1.06	*0.97*	0.96	1.00	0.98	0.97°	0.96°	0.98	0.96°	0.99
	BVAR SV−t	1.01	1.10	1.13	0.99	1.05	1.03	*0.97*	0.98	0.97°	0.94	0.99	0.99
	TVP SV	0.96	0.90	*1.00*	0.98	0.94	1.00	0.97	**0.95°**	0.98	**0.94°**	**0.92**	*0.95*
	BART hom	1.10	0.93	1.16	1.05	0.94	0.98	1.10	1.02	1.00	1.08	1.09	1.11
	BART SV	1.08	*0.83'*	1.04	1.08	**0.91'**	*0.96*	1.17	1.00	0.99	1.13	1.07	1.11
	GP hom	1.09	0.90	1.08	1.16	0.97	0.99	1.35	1.04	1.00	1.38	1.08	1.10
	GP SV	1.20	0.91	1.11	1.22	0.98	1.01	1.32	1.03	1.01	1.23	1.06	1.11

L

		QS10			QS25			QS75			QS90		
	Horizon	1	4	8	1	4	8	1	4	8	1	4	8
Government Debt Ratio	BVAR SV	1.05	**0.95**	0.97	**0.89**	**0.90**	*1.04*	**0.84'**	1.03	1.05	0.99	1.04	1.06
	BVAR SV−t	1.57	1.02	**0.90**	1.21	1.16	1.27	1.01	1.16	1.16	1.15	1.16	1.14
	TVP SV	1.14	1.18	1.41	1.09	1.07	1.20	0.94	1.14	1.21	1.07	1.10	1.18
	BART hom	1.33	1.61	1.28	1.10	1.41	1.49	0.95	1.18	1.29	1.22	1.11	1.28
	BART SV	1.13	1.13	1.02	1.02	1.26	1.42	0.94	1.20	1.28	1.15	1.14	1.24
	GP hom	1.25	1.53	1.27	1.09	1.42	1.48	1.02	1.24	1.31	1.27	1.12	1.23
	GP SV	1.19	1.09	1.04	1.06	1.32	1.42	1.02	1.26	1.31	1.22	1.17	1.24
Deficit Ratio	BVAR SV	**0.98**	1.03	1.01	0.94	*1.00*	1.03	0.96	1.17	1.15	1.02	1.11	1.12
	BVAR SV−t	1.24	1.04	1.07	1.24	1.04	1.13	1.25	1.34	1.36	1.32	1.36	1.42
	TVP SV	1.06	1.07	1.02	*0.91*	1.08	1.16	0.91	1.28	1.40	1.03	1.26	1.51
	BART hom	1.43	1.12	0.96	1.30	1.15	0.99	0.94	1.04	0.87*	0.88'	0.98	*0.79°*
	BART SV	1.17	1.10	0.97	1.20	1.15	1.01	*0.85°*	0.94	0.84*	0.91	0.86'	0.81°
	GP hom	1.46	1.12	*0.96*	1.33	1.13	*0.98*	0.90	0.94	0.81*	*0.83'*	0.88	0.80'
	GP SV	1.20	1.09	0.98	1.27	1.11	0.98	0.87'	**0.91**	*0.81°*	0.97	*0.81°*	0.85'
Growth	BVAR SV	*0.95*	1.00	1.01	*1.01*	1.01	1.02	0.99	*1.03*	1.03	*1.01*	0.97	0.99
	BVAR SV−t	0.99	0.97	0.98	1.06	*0.97*	1.02	1.09	1.07	1.07	1.13	0.93'	0.99
	TVP SV	1.07	1.12	1.16	1.15	1.11	1.17	1.08	1.21	1.17	1.29	1.15	1.11
	BART hom	0.98	**0.90**	*0.91*	1.03	0.99	*0.95*	0.97	1.08	**0.95**	1.22	*0.88*	*0.90*
	BART SV	1.01	1.02	0.99	1.07	1.07	0.99	*0.92'*	1.04	1.01	1.07	1.01	1.00
	GP hom	1.00	0.92	0.93	1.06	1.02	0.97	1.02	1.12	0.98	1.32	0.93	0.91
	GP SV	1.02	0.95	0.95	1.08	1.02	0.97	0.96	1.04	0.97	1.01	0.94	0.94
Inflation	BVAR SV	1.07	1.11	0.94	1.03	1.02	0.94	1.02	1.04	1.01	0.97	1.04	0.98
	BVAR SV−t	1.50	1.32	0.96	1.27	1.08	0.95	1.16	1.09	0.99	1.05	1.12	0.95
	TVP SV	1.14	1.02	1.07	1.06	0.99	1.01	1.02	1.06	*0.99*	0.97	*0.99*	**0.81**
	BART hom	1.24	0.94	1.03	1.33	0.97	0.95°	1.46	1.06	1.01	1.44	1.14	1.14
	BART SV	1.25	**0.81'**	**0.77'**	1.42	*0.92*	**0.88***	1.49	1.05	1.01	1.19	1.07	1.11
	GP hom	1.15	0.87'	0.93	1.27	0.95	0.91°	1.55	1.04	0.99	1.60	1.10	1.09
	GP SV	1.22	0.89	0.91	1.32	0.95	0.91°	1.45	*1.03*	0.99	1.27	1.06	1.08

Notes: The table shows the quantile scores for quantiles $\tau = \{0.10, 0.25, 0.75, 0.90\}$ for key variables and different models for horizon $h = \{1, 4, 8\}$ as ratio to the BVAR-SV model based on the *medium* (**M**) information set. Results in the left panel are based on the *small* (**S**) information set; on the right panel on the *large* (**L**) information set.