IHS Working Paper 56
September 2024

Inflation Forecasting in Turbulent Times

Martin Ertl
Ines Fortin
Jaroslava Hlouskova
Sebastian P. Koch
Robert M. Kunst
Leopold Sögner

INSTITUTE FOR
ADVANCED STUDIES
VIENNA

Author(s)

Martin Ertl, Ines Fortin, Jaroslava Hlouskova, Sebastian P. Koch, Robert M. Kunst, Leopold Sögner

Editor(s)

Robert M. Kunst

Title

Inflation Forecasting in Turbulent Times

Institut für Höhere Studien - Institute for Advanced Studies (IHS)
Josefstädter Straße 39, A-1080 Wien
T +43 1 59991-0
www.ihs.ac.at
ZVR: 066207973

Inflation Forecasting in Turbulent Times

Martin Ertl Ines Fortin Jaroslava Hlouskova Sebastian P. Koch
Robert M. Kunst Leopold Sögner*

Abstract

Recently, many countries were hit by a series of macroeconomic shocks, most notably as a consequence of the COVID-19 pandemic and Russia's invasion in Ukraine, raising inflation rates to multi-decade highs and suspending well-documented macroeconomic relationships. To capture these tail events, we propose a mixed-frequency Bayesian vector autoregressive (BVAR) model with t-distributed innovations or with stochastic volatility. While inflation, industrial production, oil and gas prices are available at monthly frequencies, real gross domestic product (GDP) is observed at a quarterly frequency. Thus, we apply a mixed-frequency framework using the forward-filtering-backward-sampling algorithm to generate monthly real GDP growth rates. We forecast inflation in those euro area countries which extensively import energy from Russia and therefore have been heavily exposed to the recent oil and gas price shocks. To measure the forecast performance of our mixed-frequency BVAR model, we compare these inflation forecasts with those generated by a battery of competing inflation forecasting models. The proposed BVAR models dominate the competition for all countries in terms of the log predictive density score.

Keywords: Bayesian VAR, mixed-frequency, forward-filtering-backward-sampling, inflation forecasting

JEL classification: C5, E3

*All authors: Institute for Advanced Studies (IHS), Josefstädter Straße 39, 1080 Vienna, Austria. Leopold Sögner has a further affiliation with the Vienna Graduate School of Finance (VGSF). The authors would like to thank participants of the 2023 Meeting of the Austrian Economic Association (NOeG), Joshua C.C. Chan, Oscar Fernandez, Sylvia Frühwirth-Schnatter, Helmut Hofer, and two anonymous referees for helpful comments that have contributed to improving the paper. The authors gratefully acknowledge support from Peter Griessl, Christine Lietz, Johannes Nemeth, and Yannic Prohaska. Leopold Sögner acknowledges support by the Cost Action HiTEc – CA21163.

1 Introduction

The COVID-19 pandemic, resulting supply side disruptions, the quick economic recovery, and the energy price shock following Russia's invasion of Ukraine had unforeseen consequences on inflation dynamics and have posed major challenges for inflation forecasting. Evidence has emerged that parameter estimation in time series models widely used for macroeconomic forecasting has become more difficult due to the COVID-19 shock and its aftermath.

For euro area inflation, Bobeica and Hartwig (2023) document that parameter estimates of Bayesian vector autoregressions (BVAR) were strongly affected. They propose to use a fat-tailed distribution for the error terms. To improve the accuracy of euro area inflation forecasts, they also recommend estimating larger models with a tighter prior (compared to standard BVAR specifications) and including off-model information for forecasts, such as information from the ECB Survey of Professional Forecasters (see also Krüger et al., 2017; Banbura et al., 2021).

Other work addressing the recent tail events focuses on the US economy, such as Lenza and Primiceri (2022), Carriero et al. (2022) and Schorfheide and Song (2021). More specifically, Lenza and Primiceri (2022) modify the innovation variance for the pandemic period. They exploit the fact that we know the exact timing of the increase in the innovations' variance during the COVID-19 period (March and subsequent months in 2020). Whereas this might be true for the pandemic, it is harder to disentangle the exact timing of the heterogeneous effects of rising energy prices on inflation rates in different countries. Countries have faced differing dependencies on energy supply from Russia, and governments have been implementing different policies to mitigate rapidly rising prices. For the period after May 2020, the authors simply assume that the residual variance will decay at a monthly rate of 20%.

Carriero et al. (2022) suggest allowing for Student-t distributed innovations and outliers in a vector autorregressive (VAR) model with stochastic volatility. Extreme observations are viewed as outliers that are characterized by transitory increases in volatility, in which case it may be desirable to reduce their influence on model estimates. Their model augments the standard stochastic volatility specification with an outlier state. For the treatment of fat-tailed errors in stochastic volatility, they use t-distributed innovations. Antolin-Diaz et al. (2021) also allow for short-lived outliers that do not lead to a persistent rise in the stochastic volatility process in a dynamic factor model for nowcasting US GDP.

Alternatively, Schorfheide and Song (2021) reconsider a mixed-frequency VAR to generate macroeconomic forecasts for the US during COVID-19. The recommendation is to exclude extreme observations during a few months of the pandemic to improve the forecasting performance. However, this assumes that the timing of outliers is known ex-ante and does not address the subsequent period of uncertainty properly. Furthermore, Clark et al. (2023) apply Bayesian machine learning techniques to account for possible non-linearity. The authors demonstrate that Bayesian regression trees have strong forecasting properties in both the overall level and in the

tails, respectively.

In this paper, we consider a small-scale Bayesian VAR framework with different error variance specifications to forecast inflation in turbulent times for selected European countries. Working with the most recent inflation data, our experience is similar to the evidence described in the above-mentioned literature. Forecasting inflation using VAR models based on longer historical time series until the year 2019 (pre-COVID-19) results in a rapid decay of inflation rates down to rates observed before the recent sharp increase. By contrast, estimating VARs (with Gaussian errors) with the full 2004 to 2023 data set can result in non-stable (explosive) inflation forecasts. Both results are implausible and unsatisfactory.

Therefore, we explore two different specifications of volatility. First, we choose an approach of forecasting post-pandemic inflation that is closely related to that of Bobeica and Hartwig (2023) using t-distributed disturbances. However, we consider a monthly instead of the quarterly frequency in order to use timelier and finer information on inflation dynamics. Thus, employing (quarterly) gross domestic product (GDP) results in a mixed-frequency problem. Further, we use the gas price as an additional energy variable. Second, we consider models with stochastic volatility to capture extreme events. It is a well-known fact that Bayesian VARs with time-varying volatility often provide better point and density forecasts of macroeconomic variables than models with homoscedastic errors terms; see, e.g., Clark (2011) and Clark and Ravazzolo (2015). Specifically, we consider the error terms to be generated by a factor stochastic volatility model as proposed in Kastner (2019).

We include data on inflation, industrial production, and GDP from six euro area countries, namely Austria, Belgium, Finland, Germany, Italy, and Slovakia, which depend strongly on natural gas imports from Russia. To capture exogenous shocks affecting inflation we do not only include the oil price but also the gas price. We consider monthly observations from February 2004 to February 2024, with GDP only observed at a quarterly frequency, while the other data are available at monthly frequencies. Our underlying econometric model is a vector autoregressive model relying on monthly variables. That is, also for GDP the underlying model applies monthly growth rates (see, for instance Proietti and Giovannelli, 2021, for frequentist monthly GDP estimates). To perform parameter estimation this article follows a Bayesian approach. In working with Student-t distributed noise terms we mainly follow Bobeica and Hartwig (2023), but – in contrast to them – we apply a Minnesota type prior to the autoregressive parameter matrices. The same prior for the autoregressive matrices is also applied in the case of stochastic volatility. In addition, we have to account for mixed sampling frequencies. We adapt forward-filtering-backward-sampling, as proposed in Frühwirth-Schnatter (1994), to obtain samples from the posterior distribution of the unobserved monthly GDP growth rates.

We conduct a comprehensive empirical analysis including the period of sharp inflation increases and decreases between mid-2021 and early-2024 in the six euro area countries, which we consider the "turbulent times" in this paper. We find that the proposed variant of a mixed-frequency

Bayesian VAR with fat-tailed errors and the alternative variant with stochastic volatility provide better out-of-sample point and density forecast accuracies than a battery of popular competing inflation forecasting models. The competing models include a univariate version of the proposed model, a version of the Bayesian VAR model with only monthly variables, disregarding GDP, a univariate autoregressive model, and homoscedastic and heteroscedastic versions of an unobserved components model. To evaluate the inflation forecasts we employ traditional measures, such as the mean absolute error and the root mean squared error, as well as log predictive density scores.

This article is organized as follows: Section 2 briefly describes the VAR model. Section 3 introduces the mixed-frequency problem. Then, Section 4 describes our Bayesian approach, in particular, the priors. Details are provided in a separate appendix. Section 5 discusses the performance of different models and then presents forecasts and an impulse response analysis for six European countries. The last section concludes.

2 The Model

In this article we jointly model industrial production, IP_t, inflation, $Infl_t$, the real gross domestic product, GDP_t, the gas price, $p_{gas,t}$, and the oil price, $p_{oil,t}$, by using a vector autoregressive (VAR) model of order p. We consider data at a monthly frequency, and index t denotes the time index. For each country, we stack the variables into the five-dimensional column vector $\mathbf{y}_t = (\Delta \ln IP_t, Infl_t, \Delta \ln GDP_t, \Delta \ln p_{gas,t}, \Delta \ln p_{oil,t})^\top \in \mathbb{R}^{\tilde{k}}$ where $\tilde{k} = 5$. Then we get[1]

$$\mathbf{y}_t \;=\; \mathbf{a} + \sum_{j=1}^{p} \mathcal{A}_j \mathbf{y}_{t-j} + \boldsymbol{\varepsilon}_t \;. \tag{1}$$

In the following we assume that the growth rates of the oil and the gas prices are not affected by $\Delta \ln IP_t$, $Infl_t$, $\Delta \ln GDP_t$ and therefore set the corresponding elements of $\mathcal{A}_j$ to zero (see also Equation (11) in Appendix A). In this article the noise term $\boldsymbol{\varepsilon}_t$ either follows a Student-t distribution (*Case 1*), or is generated by a stochastic volatility model (*Case 2*).

Case 1: Following Bobeica and Hartwig (2023), the noise term $\boldsymbol{\varepsilon}_t$ follows an *iid* multivariate Student-t distribution with mean zero, covariance matrix $\boldsymbol{\Sigma}$, where $0 < \boldsymbol{\Sigma} < \infty$, and ν degrees of freedom. From Bayesian literature (see, e.g., Geweke, 1993; Bobeica and Hartwig, 2023) a Student-t distributed noise term $\boldsymbol{\varepsilon}_t$ with ν degrees of freedom can be obtained by drawing $\boldsymbol{\varepsilon}_t$ from

[1] In this article we apply the following notation: Δx_t denotes $x_t - x_{t-1}$ and $\Delta \ln x_t$ abbreviates $\ln x_t - \ln x_{t-1}$ (that is, growth rates are calculated as logarithmic growth rates). For vectors and matrices we use boldface. If not otherwise stated, the vectors considered are column vectors. $\mathbf{0}_{a\times b}$ and $\mathbf{1}_{a\times b}$ stands for $a \times b$ matrix of zeros and ones and $\mathbf{0}_a$ is used to abbreviate $\mathbf{0}_{a\times 1}$. $\otimes$ denotes the Kronecker product and $\mathbf{I}_n$ the identity matrix of dimension $n \times n$. $vec(\mathbf{M})$ vectorizes the matrix $\mathbf{M}$, while $vech(\mathbf{M})$ vectorizes the lower triangular part of a symmetric matrix $\mathbf{M}$. $\mathcal{N}(\cdot,\cdot)$, $\mathcal{IG}(\cdot,\cdot)$, and $\mathcal{W}(\cdot,\cdot)$ denotes the multivariate normal, the inverse Gamma, and the Wishart distribution, respectively. $\mathcal{U}(\underline{\nu},\bar{\nu})$ abbreviates a uniform distribution on the interval $[\underline{\nu},\bar{\nu}]$. $\propto$ stands for proportional to.

a multivariate normal with mean zero and covariance matrix $\boldsymbol{\Sigma}_t := \lambda_t \boldsymbol{\Sigma}$, where λ_t is sampled from an inverse Gamma distribution $\mathcal{IG}\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$.

Case 2: Alternatively, we consider the noise terms $\boldsymbol{\varepsilon}_t$ to be generated by a factor stochastic volatility model as proposed in Kastner (2019). That is, for a $\tilde{k} \times \tilde{k}$-dimensional matrix $\boldsymbol{\Sigma}_t$ we assume

$$\boldsymbol{\Sigma}_t = \boldsymbol{\Lambda} \mathbf{V}_t \boldsymbol{\Lambda}^\top + \boldsymbol{\Sigma}_{Ut} \ , \tag{2}$$

where $\boldsymbol{\Lambda}$ is a $\tilde{k} \times r$-loading matrix, r is the number of volatility factors $\mathbf{V}_t = diag\left(\exp(h_{1t}), \ldots, \exp(h_{rt})\right) \in \mathbb{R}^{r \times r}$, $\boldsymbol{\Sigma}_{Ut} = diag\left(\exp(h_{r+1t}), \ \ldots \exp(h_{r+\tilde{k}t})\right) \in \mathbb{R}^{\tilde{k} \times \tilde{k}}$, and each h_{jt}, $j = 1, \ldots, \tilde{k}+r$, follows a stable first order autoregressive process with normally distributed noise terms. Then, $\boldsymbol{\varepsilon}_t = \boldsymbol{\Sigma}_t^{1/2} \boldsymbol{\eta}_t$, where $\boldsymbol{\eta}_t$ follows a $\tilde{k}$-dimensional standard normal distribution.

The parameter vector $\boldsymbol{\theta}$ collects all the parameters of the VAR considered in (1), that is $\mathbf{a}$, and the vectorized parameter matrices $\mathcal{A}_j$, $j = 1, \ldots, p$. For t-distributed innovations it also contains $vech\left(\boldsymbol{\Sigma}\right)$, $\lambda_1, \ldots, \lambda_T$, as well as ν, while for the stochastic volatility model it contains all the parameters of the factor stochastic volatility models defined in Kastner (2019). We choose p such that the autocorrelations of the residuals are insignificant, i.e., $p = 4$.

The VAR system defined in (1) results in the matrix polynomial $\mathbf{a}(z) = \mathbf{I}_k - \mathcal{A}_1 z - \cdots - \mathcal{A}_p z^p$, $z \in \mathbb{C}$. We assume that the stability condition (the determinant of $\mathbf{a}(z) \neq 0$, for all $|z| \leq 1$) is met. Let L denote the lag operator. Then, $\mathbf{y}_t = \mathbf{a}(L)^{-1} \boldsymbol{\varepsilon}_t$, $t \in \mathbb{Z}$, provides us with the unique (weakly) stationary (and causal) solution of (1) (see, e.g., Deistler and Scherrer, 2018, Theorem 4.4).

In addition, we conducted a panel VAR analysis. However, the forecasting performance turned out to be better in the country-by-country setup. That is why we focus in the main text on the country-specific VARs.

3 Data and mixed-frequency

We use industrial production, inflation and real gross domestic product for the six countries Austria, Belgium, Finland, Germany, Italy, and Slovakia. The selected countries are all part of the European Economic and Monetary Union (EMU) and also depend strongly on gas imports (from Russia).[2] Thus, they are particularly vulnerable to gas price shocks and natural candidates for analyzing oil and gas price shocks as potential drivers of (energy) inflation. We do not consider, for instance, countries like Portugal or Spain that import little or zero natural gas from Russia. The inflation rates of the Baltic countries may have been affected much more by the Russian invasion in Ukraine due to a generally broader economic interaction with Russia, Belarus and Ukraine and are therefore also not considered. France and the Netherlands are not in the country

[2]For estimates of the number and diversity of gas supply sources, see, for instance, the European Union Agency for the Cooperation of Energy Regulators, https://aegis.acer.europa.eu/chest/dataitems/214/view, last accessed 29.11.2023.

list, as the former extensively implemented anti-inflationary measures while the latter changed its method of calculating inflation[3] during the course of the year 2023 (June 2023).[4]

While industrial production, inflation, and GDP are obviously country-specific, we use international price quotations for Brent oil as well as for TTF gas, thereby implicitly neglecting minor differences in country-specific wholesale prices. Industrial production and GDP are published seasonally adjusted while the harmonized index of consumer prices (HICP) is not. Instead of seasonally adjusting the HICP and using month-over-month percentage changes, we opt to work with price changes on a year-over-year basis (annual inflation), which effectively acts as some sort of seasonal adjustment. All variables measured in prices are denominated in Euro with the exception of the oil price, which is originally measured in US Dollar and then converted to Euro using the US Dollar/Euro exchange rate. The inflation rate is measured in percent.

Further, we apply the following data transformations: we calculate month-over-month logarithmic growth rates for industrial production, oil and gas prices, and quarter-over-quarter growth rates for GDP. That is, we get the transformed variables $\Delta \ln IP_t$, $\Delta \ln p_{gas,t}$, $\Delta \ln p_{oil,t}$ observed on a monthly basis, and $\ln GDP_q - \ln GDP_{q-1}$, observed on a quarterly basis, where $q, q+1, \ldots$ denotes a quarterly time scale. For final estimation we consider the period February 2004 to February 2024. The starting date of our sample is determined by the availability of gas prices.[5] The data, its sources and transformations are summarized in Table 1.

Variable	*Abbreviation*	Transformation	Source	Dataset or Code
Industrial production	IP_t	$\Delta \ln IP_t$	Eurostat	`sts_inpr_m`
HICP inflation rate	$Infl_t$		Eurostat	`prc_hicp_manr`
Real gross domestic product	GDP_q	$\Delta \ln GDP_q$	Eurostat	`namq_10_gdp`
TTF NL natural gas future	$p_{gas,t}$	$\Delta \ln p_{gas,t}$	Refinitiv Eikon	`TRNLTTD`
Brent oil price in Euro	$p_{oil,t}$	$\Delta \ln p_{oil,t}$	Refinitiv Eikon	`OILBREN/USEURSP`
US Dollar/Euro exchange rate			Refinitiv Eikon	`USEURSP`

Table 1: Included variables.
t represents monthly frequency, q represents quarterly frequency. Note that inflation is calculated as the year-over-year growth rate of the Harmonized Index of Consumer Prices (HICP). For oil and gas prices as well as the exchange rate we use monthly averages of daily quotes.

Observational Scheme: Equation (1) describes the data generating process for $\mathbf{y}_t$ on a monthly basis. The variables observed at a monthly frequency are called fast variables, $\mathbf{y}_t^f$, while the variable GDP growth observed at a quarterly frequency is called a slow variable, y_t^s, with the sampling rate of the slow variable being three. In the data described above the growth rate of industrial production, $\Delta \ln IP_t$, inflation, $Infl_t$, the change of the gas price, $\Delta \ln p_{gas,t}$, and the change of the oil price, $\Delta \ln p_{oil,t}$, are observed at a monthly basis and are therefore fast variables. By contrast, GDP is observed at a quarterly rate and is a slow variable. Monthly real

[3]Switching from including only new energy contracts to a method that reflects all (new and existing) contracts.
[4]See, for instance, Armendariz et al. (2023).
[5]Our transformations ensure stationarity, which is confirmed by unit root tests.

GDP growth rates $\Delta \ln GDP_t$ are not observed. Note that GDP is a flow variable, such that the quarterly growth rate is $\sum_{j=0}^{2} y_{3,t-j} = \sum_{j=0}^{2} \Delta \ln GDP_{t-j} = \ln GDP_t - \ln GDP_{t-3}$. This variable is observed for some $t \in 3\mathbb{Z} + t_j$, where $t_j \in \{0, 1, 2\}$, depending on the starting month of the monthly series. To simplify the notation we consider the case where $t_j = 0$ in the following.

Let $\mathscr{Y}_T$ collect all high-frequency data, that is $\mathbf{y}_t$, $t = 1, \ldots, T$, and the initial values $\mathbf{y}_0, \ldots, \mathbf{y}_{0-p+1}$. $\mathscr{Y}_T^{obs}$ denotes the data observed, that is $\mathbf{y}_t^f$, for $t \in \mathbb{Z}$, and $\ln GDP_t - \ln GDP_{t-3}$, for $t \in 3\mathbb{Z}$. Finally, $\mathscr{Y}_T^{miss}$ collects non-observed elements of $\mathscr{Y}_T$, and it will be estimated by means of Bayesian methods given the data observed $\mathscr{Y}_T^{obs}$. See Appendix B.1 for more details.

4 Bayesian Analysis

By the Bayes theorem

$$\pi\left(\boldsymbol{\theta}, \mathscr{Y}_T^{miss} | \mathscr{Y}_T^{obs}\right) \quad \propto \quad f\left(\mathscr{Y}_T^{obs} | \mathscr{Y}_T^{miss}, \boldsymbol{\theta}\right) \pi\left(\mathscr{Y}_T^{miss} | \boldsymbol{\theta}\right) \pi\left(\boldsymbol{\theta}\right) \ , \tag{3}$$

where $\pi\left(\boldsymbol{\theta}, \mathscr{Y}_T^{miss} | \mathscr{Y}_T^{obs}\right)$ is the joint posterior density of the parameter $\boldsymbol{\theta}$ and the missing observations $\mathscr{Y}_T^{miss}$. $f\left(\mathscr{Y}_T^{obs} | \mathscr{Y}_T^{miss}, \boldsymbol{\theta}\right)$ is the (conditional) likelihood (see also Equation (14) in the Appendix). With a slight abuse of notation, the non-observed high-frequency observations contained in $\mathbf{y}_t$, $t = 1, \ldots, T$ (in our application the monthly growth rates of GDP), are replaced by the corresponding samples $\mathscr{Y}_T^{miss}$. $\pi\left(\mathscr{Y}_T^{miss} | \boldsymbol{\theta}\right)$ and $\pi\left(\boldsymbol{\theta}\right)$ denote the priors of the missing observations and the model parameters, respectively.

Priors

Prior for the covariance matrix $\boldsymbol{\Sigma}$, *Case 1 – t*-distributed innovations: We follow the literature and commence from an inverse Wishart prior with positive definite scale matrix $\mathbf{S}_{0\Sigma}$ and degrees of freedom parameter $n_{0\Sigma}$.

Prior for the covariance matrix $\boldsymbol{\Sigma}$, *Case 2* – stochastic volatility: Also for the parameters of the stochastic volatility model priors have to be specified (see Kastner, 2019, Section 2.2). Here we use the default values suggested by the `factorstochvol` package (see Hosszejni and Kastner, 2021).

Priors for the parameters $\mathbf{a}$, $\mathcal{A}_l$, $l = 1, \ldots, p$: We stack the non-restricted elements of $\mathbf{a}$ and $\mathcal{A}_l$, $l = 1, \ldots, p$, into the column vector $\boldsymbol{\alpha}$. For $\boldsymbol{\alpha}$ we consider a *Minnesota type prior* (based on works of Litterman, see, e.g., Kilian and Lütkepohl, 2017, p. 155): For the intercept terms $\mathbf{a}$ we apply a normal prior with mean parameter $\mathbf{b}_{0\mathbf{a}}$ and covariance matrix $\mathbf{B}_{0\mathbf{a}}$. Next, we consider the non-zero elements of the matrices $\mathcal{A}_l$, $l = 1, \ldots, p$. Elements of these matrices are abbreviated by $\mathcal{A}_{\iota j,l}$, while their prior means are $b_{\iota j,l}$. By collecting terms we get the vector of mean parameters $\mathbf{b}_{0\boldsymbol{\alpha}}$. If not otherwise stated for $\mathcal{A}_{\iota j,l}$ we set the prior means $b_{\iota j,l}$ equal to zero. The prior variances

of $\mathcal{A}_{\iota j,l}$ are

$$B_{\iota j,l} = \begin{cases} (\gamma_0/l)^2 & \text{if } \iota = j \ , \\ (\gamma_0 \psi_0 \sigma_{0,y,\iota}/(l\sigma_{0,y,j}))^2 & \text{if } \iota \neq j \ , \end{cases} \tag{4}$$

where γ_0 is the prior standard deviation of $\mathcal{A}_{\iota\iota,1}$. The parameter ψ_0, $0 < \psi_0 < 1$, results in a shrinkage to zero prior of the off-diagonal elements of the autoregressive matrices. Finally, the hyper-parameters $\sigma_{0,y,j} > 0$, $j = 1, \ldots, \tilde{k}$, are introduced. By collecting terms we get the diagonal matrix of prior variances $\mathbf{B}_{0\boldsymbol{\alpha}}$.

Different values for $b_{\iota\iota,1}$ have been proposed in literature (see, e.g., Kilian and Lütkepohl, 2017; Geweke et al., 2011; Koop and Korobilis, 2021). For example, $b_{\iota\iota,1} = 1$, $\iota = 1, \ldots, \tilde{k}$, which implies that we a-priori assume the process follows a random walk, or $b_{\iota\iota,1} = 0$ for stationary time series, etc. In the following, for $\iota = 1, 2, 4, 5$ we set $b_{\iota\iota,1}$ (approximately) equal to the first order sample autocorrelation of the corresponding coordinate of $\mathbf{y}_t$, while for the slow variable (i.e., the growth rate of GDP) we set the corresponding $b_{33,1} = 0$. Here we used the 2004 to 2019 subsample. Since $\sigma_{0,y,j}$ does not depend on $\boldsymbol{\Sigma}$ we consider independent priors on $\boldsymbol{\alpha}$ and $\boldsymbol{\Sigma}$. That is, $\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\alpha}, \boldsymbol{\Sigma}) = \pi(\boldsymbol{\alpha})\,\pi(\boldsymbol{\Sigma})$.

Prior on λ and ν (in the case of t-distributed noise terms): We follow Bobeica and Hartwig (2023) and impose $\lambda_t \sim \mathcal{IG}\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$ and $\nu \sim \mathcal{U}(\underline{\nu}, \bar{\nu})$.

In our empirical analysis we apply the following priors: $n_{0\Sigma} = 25$. $\mathbf{S}_{0\Sigma} = n_{0\Sigma}\widehat{\boldsymbol{\Sigma}}^{OLS}$, where $\widehat{\boldsymbol{\Sigma}}^{OLS}$ denotes the estimate of covariance $\boldsymbol{\Sigma}$ following from OLS residuals for *Case 1* where t-distributed innovations are used. For the model with stochastic volatility (Case 2) the default values suggested by the `factorstochvol` package (see Hosszejni and Kastner, 2021) are applied, where the number of common volatility factors is set to $r = 1$. $\mathbf{b}_{0\mathbf{a}} = \mathbf{0}_5$, covariance matrix $\mathbf{B}_{0\mathbf{a}} = 1000\,\mathbf{I}_5$, $\gamma_0 = 5$ and $\psi_0 = 0.7$. For $\sigma_{0,y,j}$ we use the sample standard deviation of the data observed $y_{j,t}$ (for the slow variables this estimate is based on approximately T/N observations). Finally, $\underline{\nu} = 2$, and $\bar{\nu} = 50$.

In our Bayesian sampler we consider $M = M_0 + M_1$ sampling steps, where M_0 is the number of burn-in steps. To abbreviate individual samples we use m. In our analysis we apply $M_0 = 2,000$ and $M_1 = 8,000$. For convergence and mixing of our Markov Chain Monte Carlo (MCMC) sampler see Appendix B.2.

5 Results

5.1 Forecast evaluation

In this section we compare the inflation forecasting performance of the models described in Table 2. As argued above we are confident that the Bayesian VAR model with either stochastic volatility or with t-distributed disturbances predicts inflation well in turbulent times. The models including GDP, $M.5.t$ and $M.5.sv$, are mixed-frequency models applying the estimation procedure described in Appendix B.1. All models except for the univariate autoregressive model, $M.1.ar$, and the unobserved component model, $M.1.uc$ (see Stock and Watson, 2007), are estimated by means of Bayesian methods. We consider these two simple models to include standard classical (frequentist) benchmarks. For the univariate autoregressive model $M.1.ar$ the parameters are estimated by means of ordinary least squares and the lag order (from 1 to 12) is determined by the best forecast performance. In addition we consider the unobserved component model with stochastic volatility, $M.1.ucsv$, which is a popular model for inflation forecasting (Chan, 2013; Kroese et al., 2014, see Appendix C for a short description of the unobserved component models).

Model	Variables ($\mathbf{y_t}$)	Volatility Model	Estimation
$M.5.sv$	IP_t, $Infl_t$, GDP_t, $p_{gas,t}$, $p_{oil,t}$	stochastic volatility	Bayesian
$M.5.t$	IP_t, $Infl_t$, GDP_t, $p_{gas,t}$, $p_{oil,t}$	t-distributed noise terms	Bayesian
$M.4.sv$	IP_t, $Infl_t$, $p_{gas,t}$, $p_{oil,t}$	stochastic volatility	Bayesian
$M.4.t$	IP_t, $Infl_t$, $p_{gas,t}$, $p_{oil,t}$	t-distributed noise terms	Bayesian
$M.1.sv$	$Infl_t$	stochastic volatility	Bayesian
$M.1.t$	$Infl_t$	t-distributed noise terms	Bayesian
$M.1.ucsv$	$Infl_t$	stochastic volatility	Bayesian
$M.1.ar$	$Infl_t$	white noise	Ordinary least squares
$M.1.uc$	$Infl_t$	white noise	Maximum likelihood

Table 2: Models considered to forecast inflation ($Infl_t$).

Our dataset on monthly observations of inflation spans the period from February 2004 to February 2024. The beginning of the out-of-sample (evaluation) forecasting period is July 2021, and the end of the data sample is February 2024, i.e., $L = 32$ observations. First we estimate the models with the data ranging from February 2004 to June 2021 (T_0). Based on these estimates we compute one- to six-steps ahead forecasts, i.e., forecasts for July 2021 to December 2021. Then we expand the estimation sample by one observation (i.e., we use data from February 2004 to July 2021) and, again, generate one- to six-steps ahead forecasts, i.e., from August 2021 to January 2022. The estimation-forecast procedure is repeated until the end of the total sample, February 2024 ($T = T_0 + L$). Finally, we evaluate the forecasts using different performance criteria. Note that towards the end of the sample the forecasting horizon decreases from six to one when evaluating the forecasts. Let $\widehat{Infl}_{t+h|t}$ be the $h-$step ahead (inflation) point forecast for time $t + h$, conditional on the information available at time t (note that $h = 1, \ldots, 6$). To

obtain Bayesian point forecasts, the sample median is applied.

To evaluate and compare inflation forecasts we employ traditional loss measures, such as the mean absolute error, MAE, and the root mean squared error, $RMSE$, as well as the log predictive density score, $LPDS$, that takes into account the whole predictive distribution. We consider MAE and $RMSE$ for each forecast horizon separately, $h = 1, \ldots, 6$

$$\begin{aligned} MAE_{T_0,h} &= \frac{1}{L-h+1} \sum_{\ell=0}^{L-h} \left| \widehat{Infl}_{T_0+\ell+h|T_0+\ell} - Infl_{T_0+\ell+h} \right| \\ RMSE_{T_0,h} &= \sqrt{\frac{1}{L-h+1} \sum_{\ell=0}^{L-h} \left(\widehat{Infl}_{T_0+\ell+h|T_0+\ell} - Infl_{T_0+\ell+h} \right)^2} \end{aligned} \tag{5}$$

as well as their aggregated level

$$\begin{aligned} MAE_{T_0} &= \frac{1}{N_f} \sum_{\ell=0}^{L-1} \sum_{i=1}^{h_\ell} \left| \widehat{Infl}_{T_0+\ell+i|T_0+\ell} - Infl_{T_0+\ell+i} \right| \\ RMSE_{T_0} &= \sqrt{\frac{1}{N_f} \sum_{\ell=0}^{L-1} \sum_{i=1}^{h_\ell} \left(\widehat{Infl}_{T_0+\ell+i|T_0+\ell} - Infl_{T_0+\ell+i} \right)^2} \end{aligned} \tag{6}$$

where

$$h_\ell = \begin{cases} h, & \ell < L-h \\ L-\ell, & \ell \geq L-h \end{cases} \tag{7}$$

and $L = 32$, $h = 6$, $N_f = hL - \frac{h(h-1)}{2}$.

We obtain the log predictive density score (see, e.g., Gneiting and Raftery, 2007; Martin et al., 2024) to compare forecasts of the models estimated by Bayesian methods, i.e., $M.5.sv$, $M.5.t$, $M.4.sv$, $M.4.t$, $M.1.sv$, $M.1.t$, and $M.1.ucsv$. The log predictive density score for the variable inflation and forecast horizons $h = 1, \ldots, 6$, is defined as follows

$$\begin{aligned} LPDS_{T_0,h} &:= \sum_{\ell=0}^{L-h} \log \pi \left(Infl_{T_0+\ell+h} | \mathbf{y}_1, \ldots, \mathbf{y}_{T_0+\ell} \right) \text{ , where} \\ \log \pi \left(Infl_{T_0+\ell+h} | \mathbf{y}_1, \ldots, \mathbf{y}_{T_0+\ell} \right) &= \log \int \pi \left(Infl_{T_0+\ell+h} | \mathbf{y}_1, \ldots, \mathbf{y}_{T_0+\ell}, \boldsymbol{\theta} \right) \pi \left(\boldsymbol{\theta} | \mathbf{y}_1, \ldots, \mathbf{y}_{T_0+\ell} \right) d\boldsymbol{\theta} \text{ ,} \end{aligned} \tag{8}$$

where $\pi \left(Infl_{T_0+\ell+h} | \mathbf{y}_1, \ldots, \mathbf{y}_{T_0+\ell}, \boldsymbol{\theta} \right)$ is the conditional predictive density, and $\pi \left(\boldsymbol{\theta} | \mathbf{y}_1, \ldots, \mathbf{y}_{T_0+\ell} \right)$ is the posterior density.

By the state space structure of our mixed-frequency Bayesian VAR, the *conditionally optimal Kalman mixture approximation* proposed in Bitto and Frühwirth-Schnatter (2019) can be applied to approximate the $h-$step ahead predictive density of the variable inflation, that is, of $\pi \left(Infl_{T_0+\ell+h} | \mathbf{y}_1, \ldots, \mathbf{y}_{T_0+\ell} \right)$. From the Bayesian sampler we obtain the posterior samples

$\boldsymbol{\theta}^{(m)}$, $m = 1, \ldots, M$. Conditional on $\boldsymbol{\theta}^{(m)}$ the conditional $h-$step ahead predictive density for the variable $Infl_{T_0+\ell+h}$, that is, $\pi\left(Infl_{T_0+\ell+h}|\mathbf{y}_1, \ldots, \mathbf{y}_{T_0+\ell}, \boldsymbol{\theta}\right)$, is a normal density with mean $\mu\left(Infl_{T_0+\ell+h}, (m)\right)$ and variance $\sigma^2\left(Infl_{T_0+\ell+h}, (m)\right)$, where both terms are obtained by running the Kalman filter.[6] This allows to approximate the LPDS by means of

$$\begin{aligned} LPDS_{T_0,h} &\approx \sum_{\ell=0}^{L-h}\left[\log\frac{1}{M}\sum_{m=1}^{M}\pi\left(Infl_{T_0+\ell+h}|\mathbf{y}_1, \ldots, \mathbf{y}_{T_0+\ell}, \boldsymbol{\theta}^{(m)}\right)\right], \\ &= \sum_{\ell=0}^{L-h}\left[\log\frac{1}{M}\sum_{m=1}^{M} f_{\mathcal{N}}\left(Infl_{T_0+\ell+h}|\mu\left(Infl_{T_0+\ell+h}, (m)\right), \sigma^2\left(Infl_{T_0+\ell+h}, (m)\right)\right)\right], \end{aligned} \tag{9}$$

where $f_{\mathcal{N}}\left(x|\mu, \sigma^2\right)$ denotes the normal density with mean μ and variance σ^2. We estimate log predictive density scores $LPDS_{T_0,h}$ for the forecasting horizons $h = 1, \ldots, 6$ separately. In addition we compute an aggregate measure by summing the separate density scores, i.e., $LPDS_{T_0,agg} = \sum_{h=1}^{6} LPDS_{T_0,h}$.

Figure 1 presents results on the forecast performance of inflation with respect to the log predictive density score (LPDS), MAE, and RMSE for Austria, Belgium, Germany, Finland, Italy, and Slovakia for the models presented in Table 2. We observe two main findings. First, the LPDS is largest (i.e., forecast accuracy is best) for five-variable models (M.5.sv and M.5.t), it is smaller for four-variable models (M.4.sv and M.4.t) and it is smallest for (univariate) one-variable models (M.1.sv, M.1.t, and M.1.ucsv). Second, when the forecast performance is measured by MAE and RMSE then the best forecast accuracy, except for Germany and Finland, occurs for five-variable and one-variable models. In more detail, for Germany only one-variable models perform best while for Finland the best performing models are one-variable and four-variable models.[7] When evaluating the performance of the autoregressive model M.1.ar recall that the lag length of the AR model itself was already determined by their forecast performance. Note in addition that the worst performing model with respect to LPDS is the M.1.t model, while the worst performing model with respect to MAE and RMSE is the four-variable model (M.4.sv or M.4.t) except for Italy[8]. Finally note that – although the four-variable models perform relatively well with respect to LPDS – the MAE and RMSE of these models are quite large, in particular for Austria, Belgium, Finland, and Slovakia. We claim that the main source of these large forecasting errors is the relation inferred between inflation and industrial production. The strong variation of the growth rate of industrial production results in volatile inflation forecasts and in high forecasting errors. When including GDP this effect is not observed anymore, the forecasts as well as the impulse responses become less volatile.

[6] In particular, by means of first two equations in (22) (see Appendix B.1) and a selector matrix $\mathbf{S}_{Infl_t}$, which picks inflation out of the vector $\mathbf{x}_{T_0+\ell+h|T_0+\ell}$ we get $\mu\left(Infl_{T_0+\ell+h}, (m)\right) = \mathbf{S}_{Infl_t}\mathbf{x}^{(m)}_{T_0+\ell+h|T_0+\ell}$ and $\sigma^2\left(Infl_{T_0+\ell+h}, (m)\right) = \mathbf{S}_{Infl_t}\boldsymbol{\Pi}^{(m)}_{T_0+\ell+h|T_0+\ell}\mathbf{S}^{\top}_{Infl_t}$. For both, the model with Student-t distributed innovations as well as for the stochastic volatility model we get a normal distribution.

[7] In the case of Italy the four-variable model performs best for the 3-month forecast horizon.

[8] In the case of Italy, the largest MAE and RMSE are observed for the M.1.ucsv model.

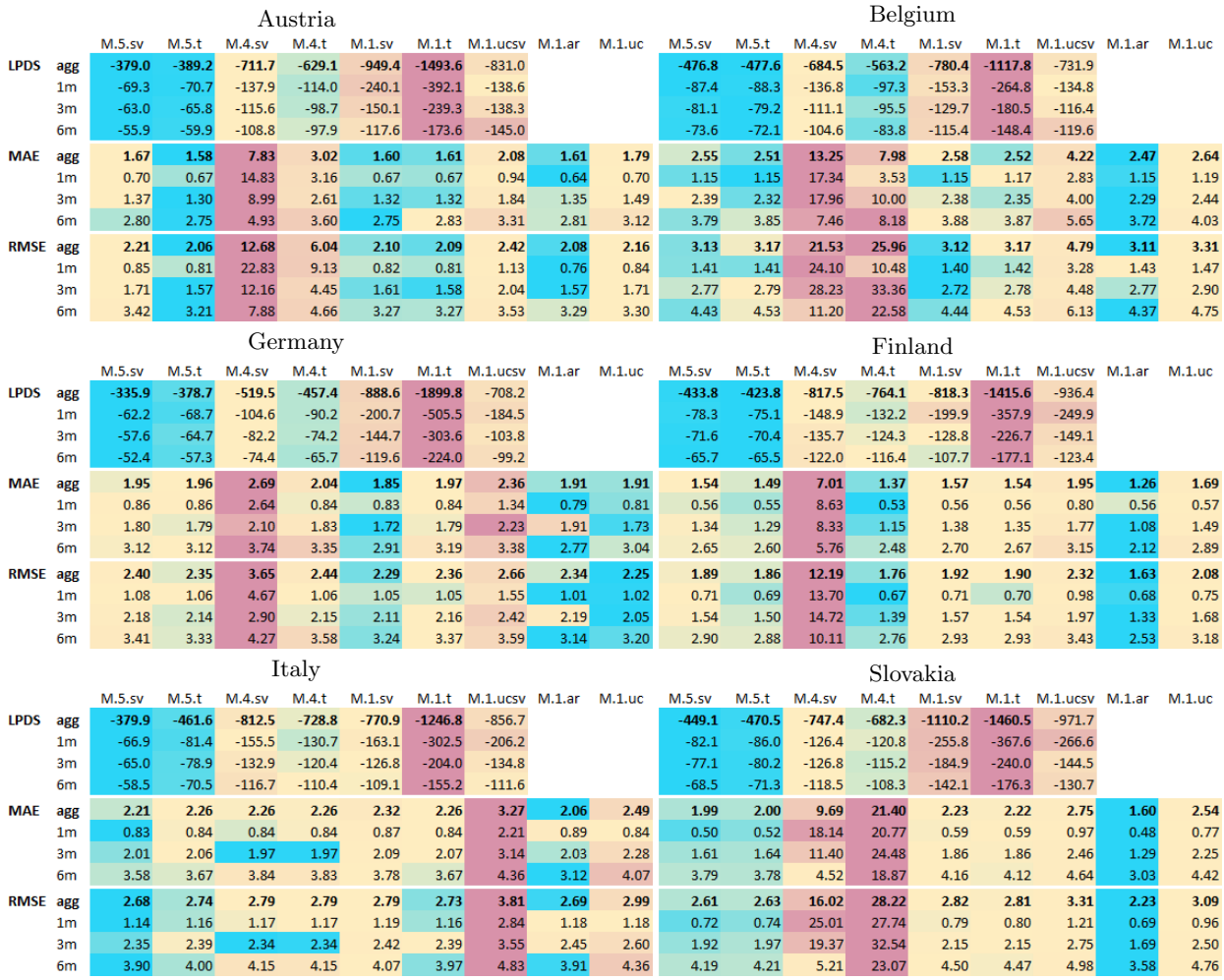

Figure 1: Forecast accuracy.
The figure shows inflation forecast accuracies of different models for Austria, Belgium, Germany, Finland, Italy, and Slovakia based on out-of-sample period from July 2021 to February 2024. We present performance measures for 1, 3, and 6 months ahead as well as aggregate performance measures. The first block shows the LPDS (log predictive density score), the second block the MAE and the third block the RMSE. The aggregate LPDS is the sum of the LPDS for 1, 2, ..., 6 months. The estimation sample starts in February 2004. The Bayesian forecasts are based on 8,000 samples and 2,000 burn-in steps. The colour coding is to be read per row, per country. The best forecast is blue, the worst forecast is red, the median forecast is yellow.

The violin plots in Figure 2 summarize the distribution of the forecast errors, $\widehat{Infl}_t - Infl_t$, of the various models for Austria. Results presented in this figure are complementary to the results presented in Figure 1 (mainly, aggregate MAE) as they provide insight about the distribution of forecast errors in contrast to a single summary statistic, and in addition give an idea about under- or over-estimation of inflation, which is not captured by the MAE. With the exception of the four-variable models the distribution of the forecast errors is rather similar across models.

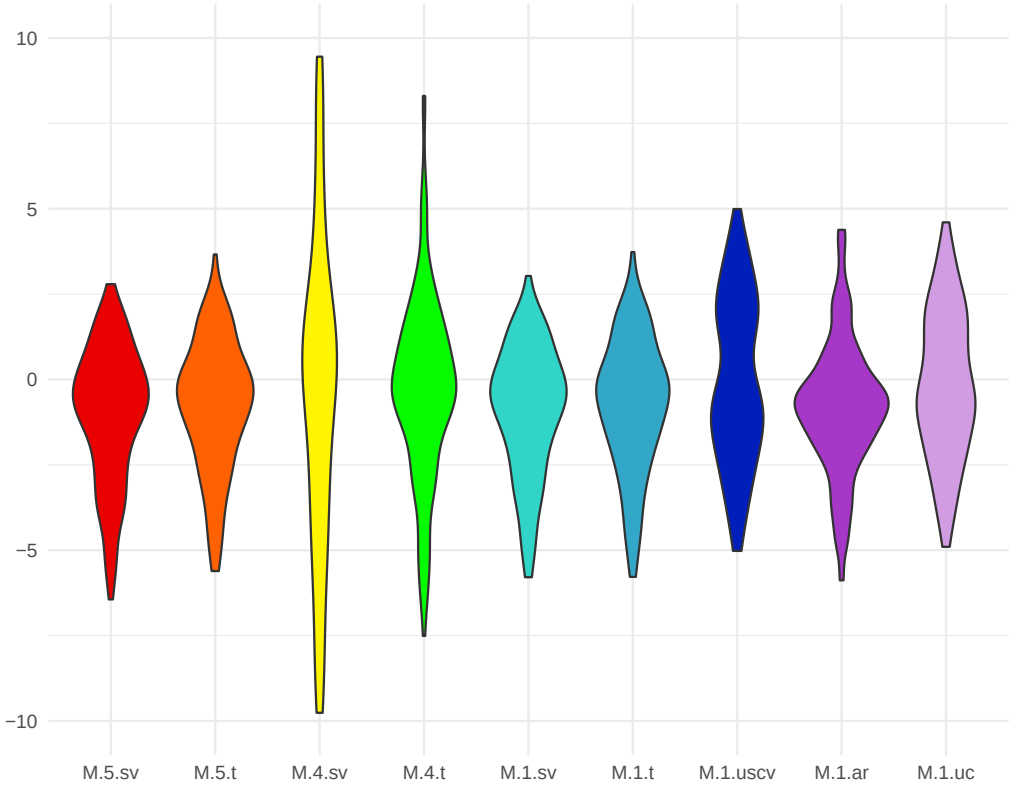

Figure 2: Inflation forecasts errors for Austria.
The violin plot (two-sided kernel density plots) summarizes the distribution of the forecast errors (i.e., forecasted inflation minus observed inflation) of all included models over all horizons ($h = 1, \ldots, 6$). The scale of the vertical axis was limited to ± 10 to exclude extreme forecast errors (in the case of M.4.sv and M.4.t models).

Densities are mostly highest for forecast errors just below zero. Negative forecast errors, that is, an underestimation of inflation, are observed more often than positive ones in our evaluation sample. The two unobserved components models show a tendency towards a bimodal forecast error distribution.

5.2 Forecasts

This section presents the inflation forecasts for the six countries considered. Figure 3 presents inflation forecasts of the (country-by-country) BVAR models $M.5.sv$ for Austria, Belgium, Germany, Finland, Italy, and Slovakia for the period June 2023 to January 2026, i.e., for 32 months ahead. The solid black lines are posterior median estimates based on 8,000 MCMC samples, and the four types of blue areas represent the 90%, 60%, 50%, and 30% forecasting intervals, respectively. As the realized values of inflation span until February 2024, we can thus observe

how well (or not) inflation was forecasted. Note that for Austria and Slovakia the nine-months ahead forecasts are inside the 90% forecasting intervals, while for Finland and Italy none of the inflation forecasts are within the 90% forecasting intervals.

Austria, Germany, and Italy behave very similarly regarding past and forecasted inflation rates. With respect to the realized values we observe in the case of Italy a sharpe decline of inflation, namely from approximately 8% in June 2023 to approximately 0.5% in February 2024. The highest inflation rates are observed in the fourth quarter of 2022 reaching rates between 11.6% (in Germany and Austria) and 12.6% (in Italy). Also the forecasts with regard to the level of inflation at the end of the forecasting horizon as well as the forecasting intervals are very much alike.

Belgium, Finland and Slovakia are different with respect to past as well as forecasted inflation rates. The strong increase as well as the sharp decrease of the Belgium inflation rate might be affected by Belgium HICP measurement.[9] The forecast for Belgium first declines below the 2% inflation target of the European Central Bank (ECB) and then approaches this target from below. Finland stands out with comparably low inflation rates. This does not come as a surprise taking into account that gas in Finland is used almost entirely by the industrial sector[10] (e.g., pulp production) and only very marginally by households.[11] The inflation development in Slovakia is different, because its peak is the largest and occurs later than in other countries. Also the forecast stands out, as it has much broader forecasting intervals with generally higher inflation rates.

Figure 4 presents six snapshots of two-years ahead inflation forecasts for Austria with forecasts starting at six different time points, namely at July 2021, January and July 2022, January and July 2023, as well as January 2024. In all six cases inflation is forecasted to decline and the inflation forecasts decrease faster when the starting points of inflation forecasts are part of the more turbulent time period when inflation in Austria was highest (July 2022 and January 2023). We also observe that the forecasting intervals are largest during more turbulent times suggesting larger forecast uncertainty.

5.3 Impulse responses

Figures 5 and 6 present impulse response functions of inflation with respect to a (positive) one-standard-deviation shock in the oil price change (i.e., about 10%) and with respect to a (positive) one-standard-deviation shock in the gas price change (i.e., about 16%) over 24 months for model $M.5.sv$. Estimates are obtained using the generalized impulse response analysis for vector autoregressive models as presented in Pesaran and Shin (1998).[12] Note that the calculation of the

[9] Note, that in Belgium only new energy contracts are included in HICP measurement and not all contracts (existing and new) (see, e.g., Jonckheere, 2022).

[10] See, for instance, Vaden et al. (2022).

[11] According to Eurostat the HICP weight of gas consumption (i.e. by Finish households) is zero.

[12] This approach does not require orthogonalization of shocks and is invariant to the ordering of the variables in the VAR.

generalized impulse response function requires estimates of the covariance matrix $\boldsymbol{\Sigma}_t$, which is time dependent for a model with stochastic volatility. When applying the stochastic volatility model we use the samples $\boldsymbol{\alpha}^{(m)}$ and the samples $\boldsymbol{\Sigma}_t^{(m)}$ to obtain the generalized impulse response function. The time point used is June 2023.[13] The solid black lines are again median estimates and the four types of blue areas represent 90%, 60%, 50% and 30% credible intervals. We observe a positive though small impact of an increase in the oil price on inflation for all six countries, with inflation first increasing and then gradually decreasing. The inflation impulse responses peak in all countries in the first year (after the shock), and the earliest inflation peaks occur for Germany and Finland, while the latest one occurs for Slovakia. Finally, the largest uncertainty (in terms of the width of the credible intervals) can be observed for Slovakia. In principle, the effects implied by a shock in the gas price are similar to the ones implied by a shock in the oil price, only smaller. We observe a positive impact of an increase in the gas price on inflation. The effect is largest for Slovakia although surrounded also by the highest uncertainty. Note that the recently observed increases in inflation are much larger than the shocks of one standard deviation applied in the impulse response functions shown in Figures 5 and 6. The prices of oil and gas rose by over 40% during the most turbulent times, while the shocks assumed in the impulse response functions are around 10% and 16%, respectively.

[13] For the stochastic volatility model the impulse responses depend on time. We show them for June 2023. The ones for February 2024 (last month) are rather similar.

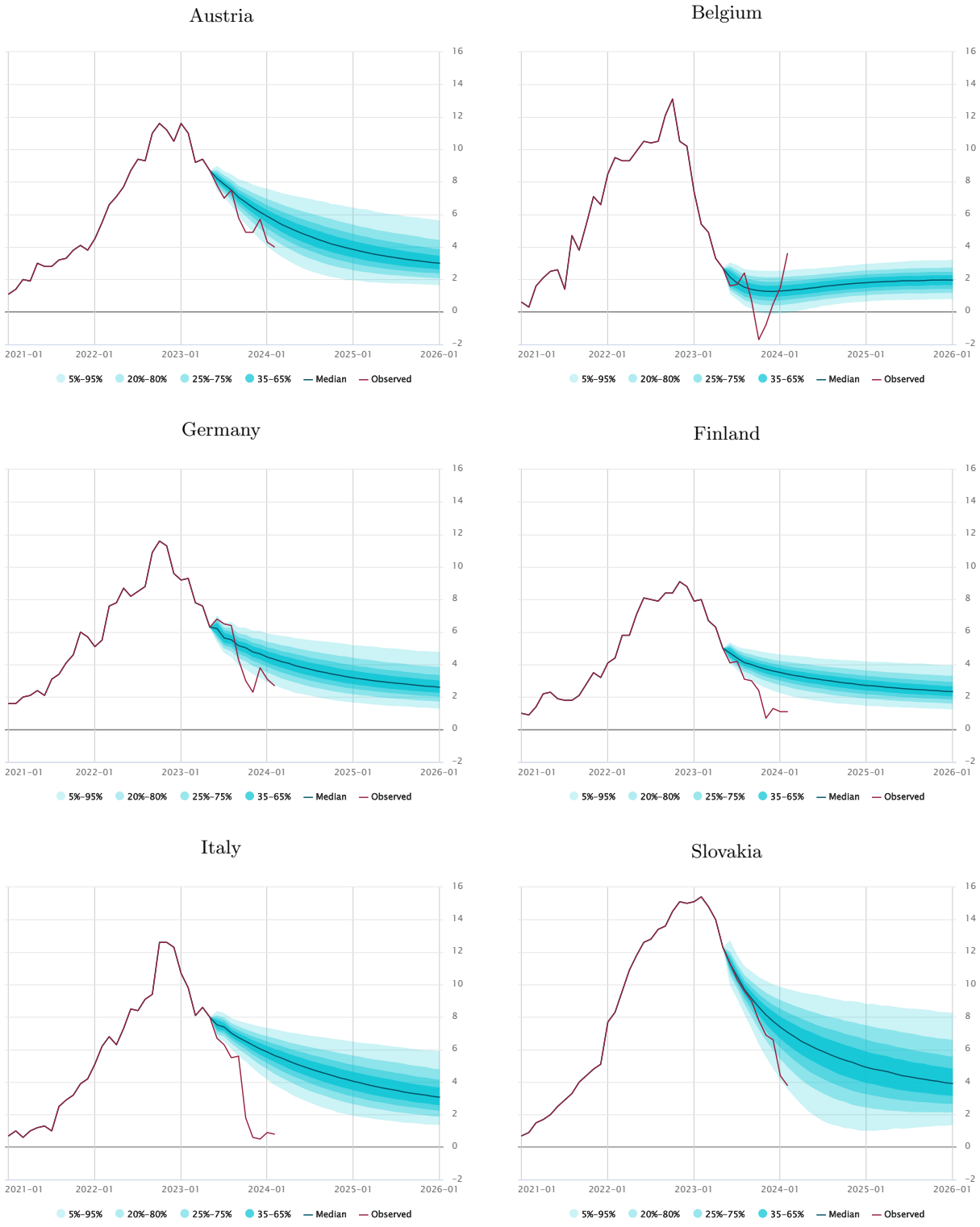

Figure 3: Inflation forecasts and forecasting intervals.
The figure shows inflation forecasts and forecasting intervals for 32 months ahead for Austria, Belgium, Germany, Finland, Italy, and Slovakia from June 2023 to January 2026. The estimation sample is February 2004 to May 2023. The forecasts are based on 8,000 samples, 2,000 burn-in steps.

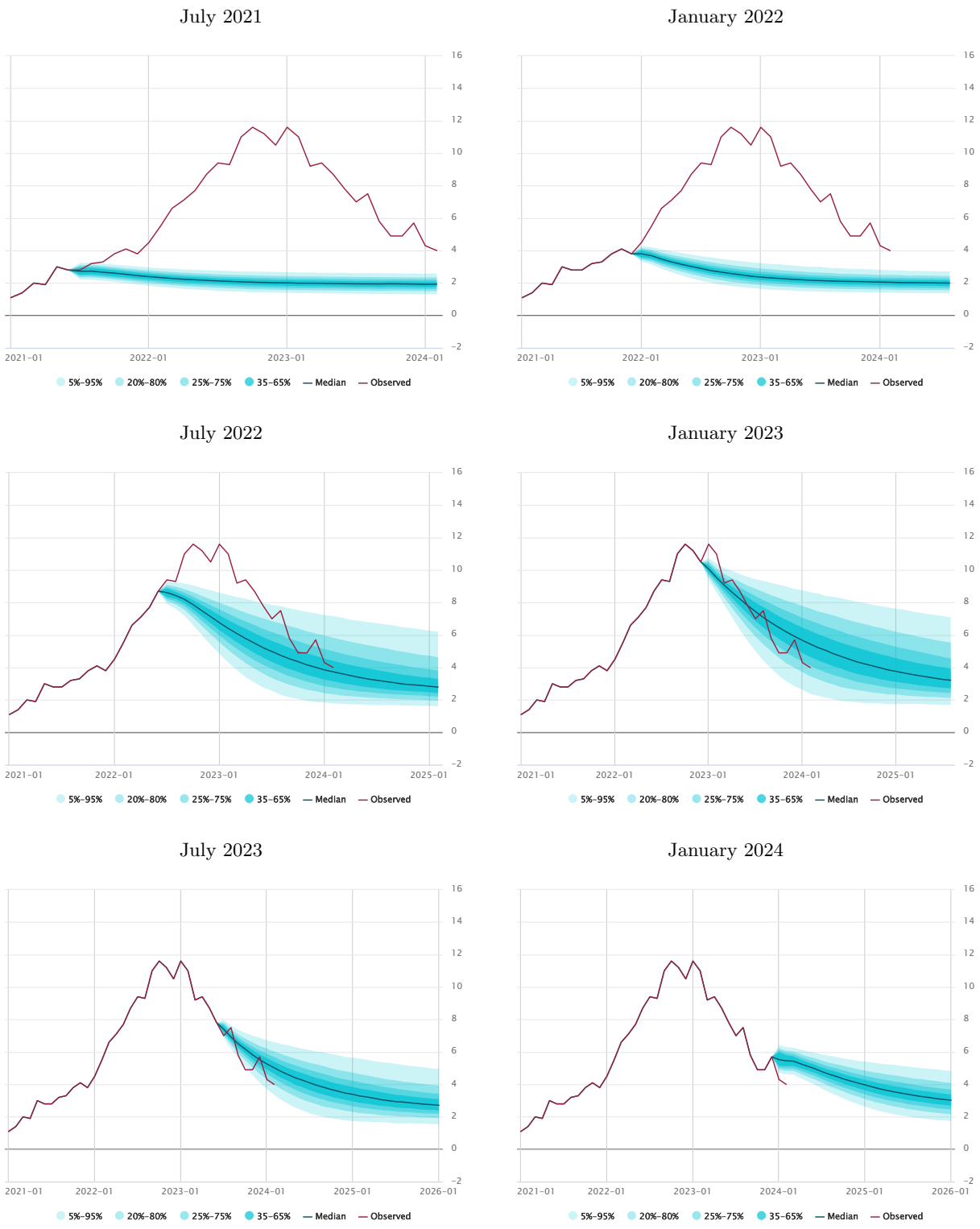

Figure 4: Inflation forecasts and forecasting intervals for Austria.
The figure shows inflation forecasts and forecasting intervals for Austria for six different starting points (July 2021, January and July 2022, January and July 2023, as well as January 2024). The estimation sample ranges from February 2004 to the month previous to the indicated starting points. The forecasts are based on 8,000 samples, 2,000 burn-in steps.

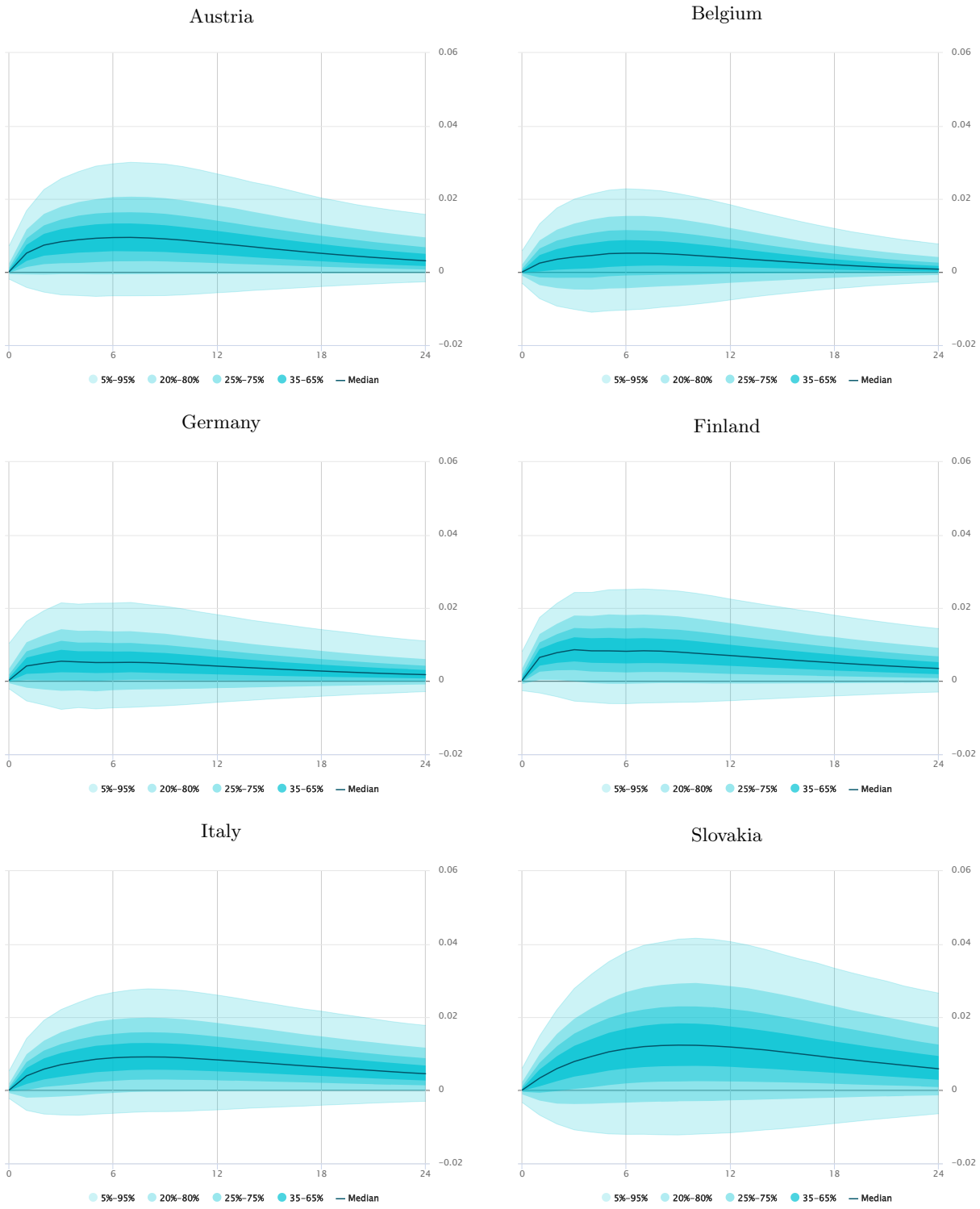

Figure 5: Generalized impulse responses of inflation with respect to the oil price.
The figure shows generalized impulse response functions of inflation (in %) with respect to a one-standard-deviation shock in the oil price (i.e., $\approx 10\%$), for Austria, Belgium, Germany, Finland, Italy, and Slovakia for 24 months. We apply samples of the covariance matrix $\mathbf{\Sigma}_t^{(m)}$ for the time point June 2023.

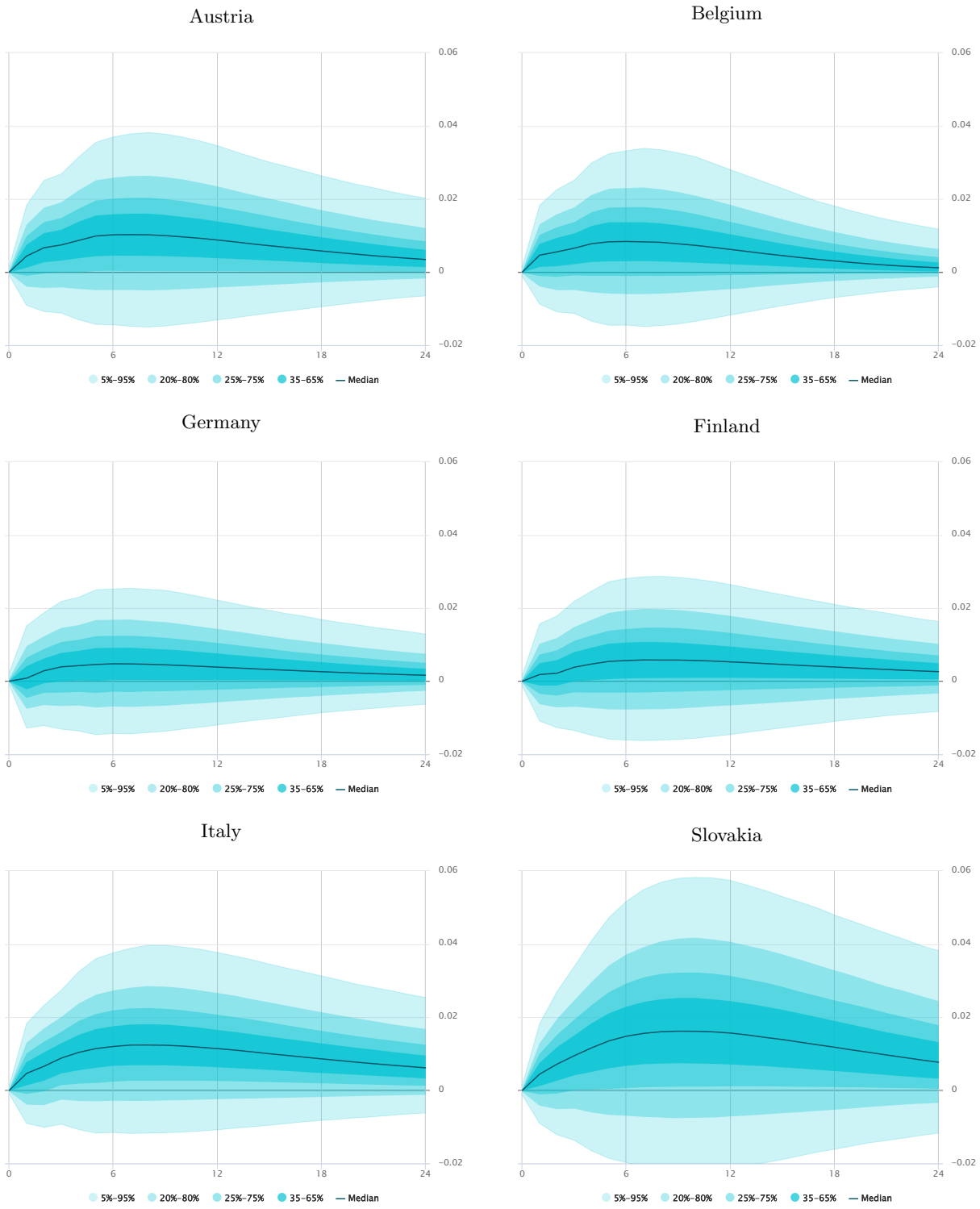

Figure 6: Generalized impulse responses of inflation with respect to the gas price.
The figure shows generalized impulse response functions of inflation (in %) with respect to a one-standard-deviation shock in the gas price (i.e., $\approx 16\%$), for Austria, Belgium, Germany, Finland, Italy, and Slovakia for 24 months. We apply samples of the covariance matrix $\mathbf{\Sigma}_t^{(m)}$ for the time point June 2023.

6 Conclusions

A series of macroeconomic shocks hit many countries between the years 2020 and 2022, primarily triggered by the COVID-19 pandemic and Russia's invasion in Ukraine, which raised inflation rates across Europe to multi-decade highs and put well-documented relationships among macroeconomic variables under scrutiny. In particular, inflation forecasting became much more difficult. We propose a mixed-frequency Bayesian vector autoregressive model and, accounting for the recent tail events, we assume Student-t distributed innovations or, alternatively, stochastic volatility. We include the variables inflation, industrial production, gross domestic product, oil and gas prices. We forecast inflation in selected euro area countries, which have been heavily exposed to energy supply from Russia and, thus, to the recent oil and gas price shocks.

We compare the forecast performance of our model with the forecast performance of several competing models of inflation in the out-of-sample period from July 2021 to February 2024. In the out-of-sample forecast evaluation it turns out that with respect to log predictive density scores the mixed-frequency BVAR models dominate the competing models. When the forecast performance is measured by MAE and RMSE, then the best forecast accuracy, except for Germany and Finland, occurs again for mixed-frequency BVAR models, though univariate models are strong competitors. Against pre-COVID-19 evidence (see, e.g., Koop and Korobilis, 2019), BVAR forecasts in a panel set-up are strongly dominated by our country-specific BVAR models, which might be due to the described country heterogeneity. Our results rather support the recent emphasis on fat-tailed noise terms for inflation modeling in the post-pandemic world as well as the vast evidence that stochastic volatility is pivotal for inflation forecasting.

In our forecasting exercise, we present inflation forecasts starting in June 2023, a time of still high inflation in most countries, until January 2026. For Austria, Germany, Finland, and Italy inflation forecasts behave similarly, they slowly decrease to rates between approximately 2.5% and 3% in January 2026. The inflation trajectory for Belgium is different, since it falls below 2% and, afterwards, convergences smoothly towards levels close to the European Central Bank's 2% inflation target. Finally, the inflation forecast for Slovakia exhibits the highest uncertainty. When forecasting inflation for Austria in different points in time, we demonstrate that the highest forecasting uncertainty occurs in the most turbulent time periods.

The methodology developed in this article can be extended in several ways: First, one can split up HICP inflation into its components. Separate modeling of these components allows to infer shocks of oil and gas prices on these components of inflation. For example, the transmission of oil and gas prices on the energy component is of particular interest. Second, instead of a reduced form VAR one may consider a structural VAR, with the goal to identify structural shocks and to model the instantaneous effects, e.g., of energy prices on inflation in more detail. Third, the structural stability of the relationship between the variables considered can be further investigated, for example, whether there are significant changes in the relationship between energy prices and

inflation during turbulent times.

A The Panel Model

This Appendix considers a panel VAR. Due to the forecasting performance of the panel model, country specific models are considered in the main text. To get the country specific analogs of the panel model (10) simply set $n = 1$.

We consider a panel *vector autoregressive* (VAR) model of order p as a starting point (in this section we mainly follow Lütkepohl, 2006; Kilian and Lütkepohl, 2017):

$$\mathbf{y}_{it} \quad = \quad \mathbf{a}_i + \sum_{j=1}^{p} \mathbf{A}_{y,j}\mathbf{y}_{it-j} + \boldsymbol{\varepsilon}_{it} \ . \tag{10}$$

The time series dimension is $t = 1, \ldots, T$, while $i = 1, \ldots, n$ denotes the cross-sectional dimension. The variables $\mathbf{y}_{it} \in \mathbb{R}^k$, the intercept terms are allowed to be country dependent, $\mathbf{a}_i$, $i = 1, \ldots, n$, while the autoregressive matrices $\mathbf{A}_{y,j}$ are the same for all countries $i = 1, \ldots, n$. The noise terms are $\boldsymbol{\varepsilon}_{it} \in \mathbb{R}^k$, $i = 1, \ldots, n$. In addition, we include common variables $\mathbf{y}_{ct} \in \mathbb{R}^{k_c}$. In the empirical application discussed in Section 5, $k = 3$ and $k_c = 2$. The vector $\mathbf{y}_{it}$ contains growth rates in industrial production, inflation and the monthly GDP growth rate. The common variables are the growth rates in the oil and gas prices.

Let $\mathbf{y}_t := \left(\mathbf{y}_{1t}^\top, \ldots, \mathbf{y}_{nt}^\top, \mathbf{y}_{ct}^\top\right)^\top \in \mathbb{R}^{nk+k_c}$, $\mathbf{a} := \left(\mathbf{a}_1^\top, \ldots, \mathbf{a}_n^\top, \mathbf{a}_c^\top\right)^\top \in \mathbb{R}^{nk+k_c}$, and $\boldsymbol{\varepsilon}_t := \left(\boldsymbol{\varepsilon}_{1t}^\top, \ldots, \boldsymbol{\varepsilon}_{nt}^\top, \boldsymbol{\varepsilon}_{ct}^\top\right)^\top \in \mathbb{R}^{nk+k_c}$.[14] Then we describe the country models including the common variables by one joint VAR system, that is

$$\mathbf{y}_t \quad = \quad \mathbf{a} + \sum_{j=1}^{p} \underbrace{\begin{pmatrix} (\mathbf{I}_n \otimes \mathbf{A}_{y,j}) & \mathbf{A}_{yc,j} \\ \mathbf{0} & \mathbf{A}_{c,j} \end{pmatrix}}_{\mathcal{A}_j \in \mathbb{R}^{(nk+k_c)\times(nk+k_c)}} \mathbf{y}_{t-j} + \boldsymbol{\varepsilon}_t \ , \tag{11}$$

where $\mathbf{A}_{yc,j}$ is a $kn \times k_c$ matrix. In (11) we imposed the simplifying assumption that $\mathbf{y}_{it}$, $i = 1, \ldots, n$, do not Granger cause $\mathbf{y}_{ct}$. The noise term $\boldsymbol{\varepsilon}_t$ follows an *iid* multivariate Student-t distribution with mean zero and covariance matrix $\boldsymbol{\Sigma}$ and ν degrees of freedom. By following, e.g., Geweke (1993), by means of sampling from a normal with mean zero and covariance matrix $\boldsymbol{\Sigma}_t = \lambda_t \boldsymbol{\Sigma}$, where $0 < \boldsymbol{\Sigma} < \infty$ and $\lambda_t \sim \mathcal{IG}\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$, we obtain samples from a multivariate t-distribution. In addition, we also consider stochastic volatility (Kastner, 2019).

The VAR system defined in (11) results in the polynomial $\mathbf{a}(z) = \mathbf{I}_k - \mathcal{A}_1 z - \cdots - \mathcal{A}_p z^p$, $z \in \mathbb{C}$. We assume that the *stability condition* (the determinant of $\mathbf{a}(z) \neq 0$, for all $|z| \leq 1$) is met. Let L denote the lag operator. Then, $\mathbf{y}_t = \mathbf{a}(L)^{-1}\boldsymbol{\varepsilon}_t$, $t \in \mathbb{Z}$, provides us with the unique stationary (and causal) solution of (1) (see, e.g., Deistler and Scherrer, 2018, Theorem 4.4).

Let $\boldsymbol{\alpha}$ be obtained by stacking $\mathbf{a}_1, \ldots, \mathbf{a}_n, \mathbf{a}_c$ and $\mathbf{A}_{y,1}, \mathbf{A}_{yc,1}, \ldots, \mathbf{A}_{y,p}, \mathbf{A}_{yc,p}, \mathbf{A}_{c,1}, \ldots, \mathbf{A}_{c,p}$ column wise. By means of $\mathbf{Z}_{it} := (\mathbf{y}_{it-1}^\top, \mathbf{y}_{ct-1}^\top, \ldots, \mathbf{y}_{it-p}^\top, \mathbf{y}_{ct-p}^\top)^\top \in \mathbb{R}^{(k+k_c)p}$, $\mathbf{Z}_{ct} := (\mathbf{y}_{ct-1}^\top, \ldots, \mathbf{y}_{ct-p}^\top)^\top \in$

[14] In the main text, $n = 1$, $k = 3$, $k_c = 2$, and $\tilde{k} = nk + k_c = 5$.

$\mathbb{R}^{k_c p}$, and

$$\mathbf{Z}_t := \left(\begin{array}{ccccc|c|c} \mathbf{I}_k & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{Z}_{1t}^\top \otimes \mathbf{I}_k & \mathbf{0} \\ & \ddots & & & & \vdots & \vdots \\ & & \ddots & & & \vdots & \vdots \\ & & & \mathbf{I}_k & \mathbf{0} & \mathbf{Z}_{nt}^\top \otimes \mathbf{I}_k & \mathbf{0} \\ & & & & \mathbf{I}_{k_c} & \mathbf{0}_{k_c \times k(k+k_c)p} & \underbrace{\mathbf{Z}_{ct}^\top \otimes \mathbf{I}_{k_c}}_{\in \mathbb{R}^{k_c \times k_c^2 p}} \end{array}\right) \in \mathbb{R}^{(kn+k_c)\times(nk+k_c+kp(k_c+k)+k_c^2 p)} , \tag{12}$$

we get

$$\underbrace{\begin{pmatrix} \mathbf{y}_{1t} \\ \vdots \\ \mathbf{y}_{nt} \\ \mathbf{y}_{ct} \end{pmatrix}}_{\mathbf{y}_t \in \mathbb{R}^{kn+k_c}} = \underbrace{\mathbf{Z}_t}_{\in \mathbb{R}^{(kn+k_c)\times(nk+k_c+kp(k_c+k)+k_c^2 p)}} \underbrace{\boldsymbol{\alpha}}_{\in \mathbb{R}^{nk+k_c+kp(k_c+k)+k_c^2 p}} + \underbrace{\begin{pmatrix} \boldsymbol{\varepsilon}_{1t} \\ \vdots \\ \boldsymbol{\varepsilon}_{nt} \\ \boldsymbol{\varepsilon}_{ct} \end{pmatrix}}_{\boldsymbol{\varepsilon}_t \in \mathbb{R}^{kn+k_c}} , \quad t = 1, \dots, T . \tag{13}$$

For Case 1, where the noise terms are t-distirubuted, let $\boldsymbol{\theta} := \left(\boldsymbol{\alpha}^\top, vech(\boldsymbol{\Sigma})^\top, \lambda_1, \dots, \lambda_T, \nu\right)^\top$ denote the vector of model parameters, while $\mathbf{y}^T$ abbreviates $\mathbf{y}_1, \dots, \mathbf{y}_T$. Since $\boldsymbol{\varepsilon}_t|\lambda_t$ is iid normally distributed, the complete likelihood (conditional on $\lambda_1, \dots, \lambda_T$) is

$$\begin{aligned} f\left(\mathbf{y}^T, (\mathbf{y}_0, \dots, \mathbf{y}_{0-p+1})|\boldsymbol{\theta}\right) &= g\left(\mathbf{y}^T|(\mathbf{y}_0, \dots, \mathbf{y}_{0-p+1}), \boldsymbol{\theta}\right) \pi\left((\mathbf{y}_0, \dots, \mathbf{y}_{0-p+1})|\boldsymbol{\theta}\right) \text{ , where} \\ g\left(\mathbf{y}^T|(\mathbf{y}_0, \dots, \mathbf{y}_{0-p+1}), \boldsymbol{\theta}\right) &= \prod_{t=1}^{T} \frac{1}{\sqrt{(2\pi)^{(kn+k_c)}|\boldsymbol{\Sigma}_t|}} \exp\left(\frac{1}{2}\left(\mathbf{y}_t - \mathbf{Z}_t\boldsymbol{\alpha}\right)^\top \boldsymbol{\Sigma}_t^{-1}\left(\mathbf{y}_t - \mathbf{Z}_t\boldsymbol{\alpha}\right)\right) , \\ &\quad \text{where } \boldsymbol{\Sigma}_t = \lambda_t \boldsymbol{\Sigma} . \end{aligned} \tag{14}$$

where $|\boldsymbol{\Sigma}_t|$ denotes the determinant of the matrix $\boldsymbol{\Sigma}_t$ and $\pi((\mathbf{y}_0, \dots, \mathbf{y}_{0-p+1})|\boldsymbol{\theta})$ denotes the density of the (unobserved) initial values.

For Case 2 (stochastic volatility), $\boldsymbol{\theta}$ contains the augmented parameters $\boldsymbol{\alpha}$, $\boldsymbol{\Sigma}_t$, and all the parameters of the stochastic volatility model. In this case the likelihood also follows from the first two rows in (14), but $\boldsymbol{\Sigma}_t$ follows from the Bayesian stochastic volatility sampler.

B Bayesian Sampling

The posterior distribution and the priors were already defined in Section 4. To obtain samples from the joint posterior distribution $\pi\left(\boldsymbol{\theta}, \mathscr{Y}_T^{miss}|\mathscr{Y}_T^{obs}\right)$ we apply the following Algorithm 1 for the Case of Student-t distributed noise terms, which except for Sampling Step 3, is mainly based on Bobeica and Hartwig (2023). For stochastic volatility we apply Algorithm 2:

Algorithm 1 (MCMC estimation). Choose starting values for $\boldsymbol{\alpha}^{(0)}$, $\boldsymbol{\Sigma}^{(0)}$ and $\mathscr{Y}_T^{miss,(0)}$, $\lambda_1^{(0)}, \dots, \lambda_T^{(0)}$, and $\nu^{(0)}$. For $m = 1, \dots, M_0 + M_1$ we draw from the conditional posterior distributions:

(1) Sample $\boldsymbol{\Sigma}^{(m)}$ from $\pi\left(\boldsymbol{\Sigma}|\boldsymbol{\alpha}^{(m-1)}, \lambda_1^{(m-1)}, \dots, \lambda_T^{(m-1)}, \mathscr{Y}_T^{obs}, \mathscr{Y}_T^{miss,(m-1)}\right)$.

(2) Sample $\boldsymbol{\alpha}^{(m)}$ from $\pi\left(\boldsymbol{\alpha}|\boldsymbol{\Sigma}^{(m)}, \lambda_1^{(m-1)}, \ldots, \lambda_T^{(m-1)}, \mathscr{Y}_T^{obs}, \mathscr{Y}_T^{miss,(m-1)}\right)$.

(3) Sample $\mathscr{Y}_T^{miss,(m)}$ from $\pi\left(\mathscr{Y}_T^{miss}|\boldsymbol{\alpha}^{(m)}, \boldsymbol{\Sigma}^{(m)}, \lambda_1^{(m-1)}, \ldots, \lambda_T^{(m-1)}, \mathscr{Y}_T^{obs}\right)$.

(4) Sample $\lambda_1^{(m)}, \ldots, \lambda_T^{(m)}$ from $\pi\left(\lambda_t|\boldsymbol{\alpha}^{(m)}, \boldsymbol{\Sigma}^{(m)}, \mathscr{Y}_T^{obs}, \mathscr{Y}_T^{miss,(m)}\right)$.

(5) Sample $\nu^{(m)}$ from $\pi\left(\nu|\boldsymbol{\alpha}^{(m)}, \boldsymbol{\Sigma}^{(m)}, \lambda_1^{(m)}, \ldots, \lambda_T^{(m)}, \mathscr{Y}_T^{obs}, \mathscr{Y}_T^{miss,(m)}\right)$.

We discard the first M_0 draws (burn-in), which results in M_1 draws from the posterior distribution. Let $M = M_0 + M_1$.

Initialization: To start the sampler we have to choose $\boldsymbol{\alpha}^{(0)}$, $\boldsymbol{\Sigma}^{(0)}$ and $\mathscr{Y}_T^{miss,(0)}$, and $\lambda_1^{(0)}, \ldots, \lambda_T^{(0)}$, and $\nu^{(0)}$. Let N denote the sampling rate, in our case $N = 3$ since GDP is observed on a quarterly basis while the other variables are observed on a monthly frequency. In this article we first obtain the missing monthly GDP growth rates by $\Delta \ln GDP_{i,s}^{(0)}$, for $s = t-2, t-1, t$, $t \in N\mathbb{Z}$, by means of $\Delta \ln GDP_{i,q}/3$ (the observed quarterly growth rates of country i, $i = 1, \ldots, n$). This provides us with $\mathbf{y}_t^{(0)}$, $t = 1, \ldots, T$. For the initial values $\mathbf{y}_s^{(0)}$, $s = p-1, \ldots, 0$, we simply use $\mathbf{y}_s^{(0)} = \mathbf{y}_1^{(0)}$. By stacking $\mathbf{y}_s^{(0)}$, $s = p-1, \ldots, 0$, we get $\mathbf{x}_0$. By that we get $\mathscr{Y}_T^{miss,(0)}$ and $\mathscr{Y}_T^{(0)}$. $\mathscr{Y}_T^{(m)}$ follows from $\mathscr{Y}_T^{miss,(m)}$ and $\mathscr{Y}_T^{obs}$. Hence, we write $\mathscr{Y}_T^{(m)}$ in the following to simplify the notation. We use $\mathbf{y}_t^{(0)}$, $t = 1, \ldots, T$, and apply ordinary least squares estimation to obtain $\boldsymbol{\alpha}^{(0)}$. The corresponding residuals are used to calculate $\boldsymbol{\Sigma}^{(0)}$. Then we use the conditional density described in Sampling Step 4 to sample $\tilde{\lambda}_t^{(0)}$, that is $\tilde{\lambda}_t^{(0)} \sim \pi\left(\lambda_t|\boldsymbol{\alpha}^{(0)}, \boldsymbol{\Sigma}^{(0)}, \mathscr{Y}_T^{(0)}\right)$, and obtain $\lambda_t^{(0)} = \frac{1}{\sum_{s=1}^{T} \tilde{\lambda}_s^{(0)}} \tilde{\lambda}_t^{(0)}$, $t = 1, \ldots, T$. Finally we sample $\nu^{(0)}$ as will be described in Sampling Step 5.

Note that the priors of $\boldsymbol{\alpha}$ and $\boldsymbol{\Sigma}$ are conditional independent given $\boldsymbol{\lambda} := (\lambda_1, \ldots, \lambda_T)^\top$. Hence, $\boldsymbol{\alpha}$ and $\boldsymbol{\Sigma}$ can be sampled as, e.g., demonstrated in Kilian and Lütkepohl (2017)[Section 5.2.5]:

Sampling from $\pi\left(\boldsymbol{\Sigma}|\boldsymbol{\alpha}^{(m-1)}, \boldsymbol{\lambda}^{(m-1)}, \mathscr{Y}_T^{(m-1)}\right)$: The variance covariance matrix $\boldsymbol{\Sigma}$ can be sampled from the conjugate inverted Wishart posterior with shape parameter $n_\pi = n_{0\Sigma} + T$ and scale parameter matrix $\boldsymbol{\Sigma}_{0\Sigma} + T\widehat{\boldsymbol{\Sigma}}$. The matrix $\widehat{\boldsymbol{\Sigma}} := \frac{1}{T}\sum_{t=1}^{T} \frac{1}{\lambda_t^{(m-1)}} \widehat{\boldsymbol{\varepsilon}}_t \widehat{\boldsymbol{\varepsilon}}_t^\top$ and $\widehat{\boldsymbol{\varepsilon}}_t$, $t = 1, \ldots, T$, denote the residuals obtained from the current sample of $\boldsymbol{\alpha}$ (see, e.g., Kilian and Lütkepohl, 2017, Chapter 5.5.2).

Sampling from $\pi\left(\boldsymbol{\alpha}|\boldsymbol{\Sigma}^{(m)}, \boldsymbol{\lambda}^{(m-1)}, \mathscr{Y}_T^{(m-1)}\right)$: Note that conditionally on $\boldsymbol{\Sigma}$, $\boldsymbol{\lambda}$, and the data, $\boldsymbol{\alpha}$ can be obtained by means of a Gibbs sampling step. That is, $\boldsymbol{\alpha}$ is sampled from a multivariate normal distribution (see, e.g., Koop and Korobilis, 2021, Equation (9))

$$
\begin{aligned}
\boldsymbol{\alpha}|\boldsymbol{\Sigma}^{(m)}, \mathscr{Y}_T^{(m-1)} \;\sim\; & \mathcal{N}\left(\mathbf{b}_{T\alpha}, \mathbf{B}_{T\alpha}\right) \;, \quad \text{where} \\
& \mathbf{B}_{T\alpha} = \left(\mathbf{B}_{0\alpha}^{-1} + \sum_{t=1}^{T} \left(\mathbf{Z}_t^{(m-1)}\right)^\top \left(\boldsymbol{\Sigma}_t^{(m)}\right)^{-1} \mathbf{Z}_t^{(m-1)}\right)^{-1} \;, \\
& \mathbf{b}_{T\alpha} = \mathbf{B}_{T\alpha}\left(\mathbf{B}_{0\alpha}^{-1}\mathbf{b}_{0\alpha} + \sum_{t=1}^{T} \left(\mathbf{Z}_t^{(m-1)}\right)^\top \left(\boldsymbol{\Sigma}_t^{(m)}\right)^{-1} \mathbf{y}_t^{(m-1)}\right) \;. \qquad (15)
\end{aligned}
$$

In the current Sampling Step 2, $\boldsymbol{\Sigma}_t^{(m)} = \lambda_t^{(m-1)}\boldsymbol{\Sigma}^{(m)}$, $t = 1, \ldots, T$. Since $\mathbf{Z}_t$ and $\mathbf{y}_t$ contain unobserved monthly GDP growth rates, the corresponding index follows from the last sample of

the missing observations, namely $\mathscr{Y}_T^{miss,(m-1)}$.

Sampling from $\pi\left(\mathscr{Y}_T^{miss}|\boldsymbol{\alpha}^{(m)}, \boldsymbol{\Sigma}^{(m)}, \lambda_1^{(m-1)}, \ldots, \lambda_T^{(m-1)}, \mathscr{Y}_T^{obs}\right)$: To get samples of the monthly gross domestic product and all initial values we augment the forward-filtering-backward-sampling algorithm proposed in Frühwirth-Schnatter (1994). More details are provided in Appendix B.1.

Sampling from $\pi\left(\lambda_t|\boldsymbol{\alpha}^{(m)}, \boldsymbol{\Sigma}^{(m)}, \mathscr{Y}_T^{(m)}\right)$: We follow Bobeica and Hartwig (2023)[Equation (9)] and sample λ_t from an

$$\mathcal{IG}\left(\frac{\nu^{(m-1)} + k_c + nk}{2}, \frac{1}{2}\left(\nu^{(m-1)} + \left(\mathbf{y}_t^{(m)} - \mathbf{Z}_t^{(m)}\boldsymbol{\alpha}^{(m)}\right)^\top \left(\boldsymbol{\Sigma}^{(m)}\right)^{-1}\left(\mathbf{y}_t^{(m)} - \mathbf{Z}_t^{(m)}\boldsymbol{\alpha}^{(m)}\right)\right)\right)$$

distribution. Since new values of the missing observations are sampled in Sampling Step 3, $\mathbf{Z}_t^{(m)}$ and $\mathbf{y}_t^{(m)}$ contain the sampling index (m).

Sampling of $\nu^{(m)}$: from $\pi\left(\nu|\boldsymbol{\alpha}^{(m)}, \boldsymbol{\Sigma}^{(m)}, \lambda_1^{(m)}, \ldots, \lambda_T^{(m)}, \mathscr{Y}_T^{(m)}\right)$. Following Chan and Hsiao (2014) and Bobeica and Hartwig (2023)[Equation (10)], the conditional posterior density of the degrees of freedom parameter

$$\begin{aligned}\pi\left(\nu|\boldsymbol{\alpha}^{(m)}, \boldsymbol{\Sigma}^{(m)}, \boldsymbol{\lambda}^{(m)}, \mathscr{Y}_T^{(m)}\right) &= \pi\left(\nu|\boldsymbol{\lambda}^{(m)}\right) \\ &\propto \frac{\frac{\nu}{2}^{T\nu/2}}{\Gamma\left(\frac{\nu}{2}\right)^T}\left(\prod_{t=1}^{T}\lambda_t^{(m)}\right)^{-\frac{\nu}{2}+1}\exp\left(-\frac{\nu}{2}\sum_{t=1}^{T}\frac{1}{\lambda_t^{(m)}}\right)\mathbf{1}_{\nu\in[\underline{\nu},\bar{\nu}]}\,, \end{aligned} \tag{16}$$

where $\Gamma(\cdot)$ denotes the Euler Gamma function, $\mathbf{1}_{\nu\in[\underline{\nu},\bar{\nu}]}$ an indicator function. By using (16) and a random walk on ν, that is $\ln\nu^{new} = \ln\nu^{old} + 0.5\zeta$, where ζ is standard normal, samples of ν can be obtained by means of the Metropolis Hastings algorithm.

Regarding the model with noise terms generated by a stochastic volatility model, let $\boldsymbol{\theta}_{sv}^{(m)}$ denote the parameters of the stochastic volatility model defined in Kastner (2019).

Algorithm 2 (MCMC estimation). Choose starting values for $\boldsymbol{\alpha}^{(0)}$, $\boldsymbol{\Sigma}_t^{(0)}$, and $\mathscr{Y}_T^{miss,(0)}$. For $m = 1, \ldots, M_0 + M_1$ we draw from the conditional posterior distributions:

(1) Sample $\boldsymbol{\theta}_{sv}^{(m)}$ and $\boldsymbol{\Sigma}_t^{(m)}$ from $\pi\left(\boldsymbol{\theta}_{sv}^{(m)}, \boldsymbol{\Sigma}_t^{(m)}|\boldsymbol{\alpha}^{(m-1)}, \mathscr{Y}_T^{obs}, \mathscr{Y}_T^{miss,(m-1)}\right)$.

(2) Sample $\boldsymbol{\alpha}^{(m)}$ from $\pi\left(\boldsymbol{\alpha}|\boldsymbol{\Sigma}_t^{(m)}, \mathscr{Y}_T^{obs}, \mathscr{Y}_T^{miss,(m-1)}\right)$.

(3) Sample $\mathscr{Y}_T^{miss,(m)}$ from $\pi\left(\mathscr{Y}_T^{miss}|\boldsymbol{\alpha}^{(m)}, \boldsymbol{\Sigma}_t^{(m)}, \mathscr{Y}_T^{obs}\right)$.

We discard the first M_0 draws (burn-in), which results in M_1 draws from the posterior distribution. Let $M = M_0 + M_1$.

Sampling Step 1 is fully implemented in the `factorstochvol`-package of Hosszejni and Kastner (2021). In the case where only inflation is modeled ($k = 1$, $k_c = 0$) the `stochvol`-package implemented in Kastner (2016) is applied.

Sampling Step 2 in Algorithm 2 is the same as in Algorithm 1, where $\boldsymbol{\Sigma}_t^{(m)}$ are samples from the sampler implemented by Hosszejni and Kastner (2021). When applying sampling Step 3 of Algorithm 2 we do not observe convergence of the sampler. In particular, larger volatility values also result in draws of $\boldsymbol{\alpha}^{(m)}$ which further on yields very extreme draws of $\mathscr{Y}_T^{miss,(m)}$. This effect builds up and results in a sampler which does not converge. Given the plausible assumption that

monthly gross domestic product does not fluctuate too much compared to the quarterly variance of GDP, we imposed a prior on the variance of the samples of the monthly GDP growth rates. In more detail, given the sample variance of quarterly GDP growth rates, $\widehat{\mathbb{V}}(\Delta \ln GDP_q)$, and the sample variance of our monthly GDP variables, that is $\widehat{\mathbb{V}}(y_{3t})$, we demand for

$$\frac{\widehat{\mathbb{V}}(y_{3t}) - \widehat{\mathbb{V}}(\Delta \ln GDP_q)}{\widehat{\mathbb{V}}(\Delta \ln GDP_q)} \leq 1.25 \ .$$

Then samples of $\mathscr{Y}_T^{miss,(m)}$ follow from Step 3 of Algorithm 1 and rejection sampling. Note that with $\mathscr{Y}_T^{miss,(m)}$ derived by linear interpolation, $\widehat{\mathbb{V}}(y_{3t}) \approx \widehat{\mathbb{V}}(\Delta \ln GDP_q)$. Hence, by the argument that monthly GDP does not fluctuate too much in relation to quarterly fluctuations, the prior imposed on $\mathscr{Y}_T^{miss,(m)}$ is not very restrictive, but it turned out that this step was necessary to obtain convergence.

B.1 Forward-Filtering-Backward-Sampling

To obtain the missing values of the monthly GDP growth rate variable and samples of the initial values (and therefore $\mathscr{Y}_T^{miss,(m)}$) we apply forward-filtering-backward-sampling proposed in Frühwirth-Schnatter (1994). First, we write our auto-regressive model (11) in state space form. We call this system the high-frequency system in the following, since all coordinates of $\mathbf{y}_t$ are assumed to be observed for $t \in \mathbb{Z}$, i.e. at monthly frequency. In the empirical data we observe a quarterly growth rate of GDP, e.g., for the first quarter, where the monthly GDP growth rate for January directly follows from the quarterly GDP growth rate (from January to March) minus the monthly growth rates sampled first for March and then for February, while the monthly GDP growth rates for March and February follow from samples obtained by means of forward-filtering-backward-sampling, which will be obtained in this section. To simplify the notation we skip the sampling index (m). First, we express the high-frequency model in state-space form:

$$
\begin{aligned}
\mathbf{y}_t &= \mathbf{S}_y \mathbf{x}_t \\
\mathbf{x}_{t+1} &= \mathbf{F}\mathbf{x}_{t+1} + \mathbf{G}\boldsymbol{\varepsilon}_t \ , \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}\left(\mathbf{0}_{nk+k_c}, \boldsymbol{\Sigma}_t\right) \ , \text{ where} \qquad (17) \\
\boldsymbol{\Sigma}_t &= \begin{cases} \lambda_t \boldsymbol{\Sigma} \ , \quad \lambda_t \text{ sampled from inverse gamma distribution} \\ \text{following from the stochastic volatility model (see equation (2))} \end{cases} \ , \\
\mathbf{S}_{iy} &:= \underbrace{\left(\mathbf{I}_k \ \mathbf{0}_{k\times k(p-1)}\right)}_{[k\times kp]} \ ,
\end{aligned}
$$

$\mathbf{S}_{cy}$ of dimension $k_c \times k_c p$ is obtained conformingly. Then,

$$
\mathbf{S}_y := \left(\begin{array}{c|cc} \mathbf{0}_{nk\times 1} & \mathbf{I}_n \otimes \mathbf{S}_{iy} & \mathbf{0} \\ \mathbf{0}_{k_c\times 1} & \mathbf{0} & \mathbf{S}_{cy} \end{array}\right) \in \mathbb{R}^{nk+k_c\times 1+nkp+k_c p} \ ,
$$

$$
\mathbf{F} := \left(\begin{array}{c|c|c} 1 & \mathbf{0}_{1\times nkp} & \mathbf{0}_{1\times nk_c} \\ \hline \underbrace{\begin{matrix} \mathbf{e}_{1p}\otimes\mathbf{a}_1 \\ \mathbf{e}_{1p}\otimes\mathbf{a}_2 \\ \vdots \\ \mathbf{e}_{1p}\otimes\mathbf{a}_n \end{matrix}}_{\in\mathbb{R}^{kpn\times 1}} & \underbrace{\mathbf{I}_n \otimes \begin{pmatrix} \mathbf{A}_{yy,1} & \cdots & \ldots & \mathbf{A}_{yy,p} \\ \mathbf{I}_k & \mathbf{0} & \ldots & \\ \mathbf{0} & \ddots & \mathbf{0} & \ldots \\ \ddots & \mathbf{0} & \mathbf{I}_k & \mathbf{0} \end{pmatrix}}_{\in\mathbb{R}^{knp\times knp}} & \underbrace{\mathbf{1}_n \otimes \begin{pmatrix} \mathbf{A}_{yc,1} & \ldots & \mathbf{A}_{yc,p} \\ \mathbf{0} & \ldots & \\ & \mathbf{0} & \ldots \\ & \mathbf{0} & \end{pmatrix}}_{\in\mathbb{R}^{knp\times k_c p}} \\ \hline \underbrace{\begin{matrix} \mathbf{e}_{1p}\otimes\mathbf{a}_c \end{matrix}}_{[k_c p\times 1]} & \mathbf{0}_{k_c\times k} \ \cdots \ \cdots \ \mathbf{0}_{k_c\times k} & \underbrace{\begin{pmatrix} \mathbf{A}_{c,1} & \ldots & \ldots & \mathbf{A}_{c,p} \\ \mathbf{I}_{k_c} & \mathbf{0} & \ldots & \\ \mathbf{0} & \ddots & \mathbf{0} & \ldots \\ \ddots & \mathbf{0} & \mathbf{I}_{k_c} & \mathbf{0}_{k_c\times k_c} \end{pmatrix}}_{\in\mathbb{R}^{k_c p\times k_c p}} \end{array}\right) \ ,
$$

$$
\mathbf{x}_t = \left(1, \mathbf{y}_{1t}^\top, \ldots, \mathbf{y}_{1t-p+1}^\top, \mathbf{y}_{2t}^\top, \ldots, \mathbf{y}_{2t-p+1}^\top, \ldots, \mathbf{y}_{nt}^\top, \ldots, \mathbf{y}_{nt-p+1}^\top, \mathbf{y}_{ct}^\top, \ldots, \mathbf{y}_{ct-p+1}^\top\right)^\top \in \mathbb{R}^{(nk+k_c)p+1}
$$

$$
\mathbf{G} = \underbrace{\left(\begin{array}{cc} \mathbf{0}_{1\times nk} & \mathbf{0}_{1\times k_c} \\ \underbrace{\mathbf{I}_n \otimes \left(\mathbf{e}_{1p}\otimes\mathbf{I}_k\right)}_{[nkp\times nk]} & \mathbf{0}_{nkp\times k_c} \\ \mathbf{0}_{k_c p\times nk} & \underbrace{\mathbf{e}_{1p}\otimes\mathbf{I}_{k_c}}_{[k_c p\times k_c]} \end{array}\right)}_{[npk+pk_c+1\times nk+k_c]} \ , \quad \text{and} \quad \mathbf{e}_{1p} = \underbrace{(1,0,\ldots,0)^\top}_{[p\times 1]} \ . \qquad (18)
$$

To cope with missing observations, we consider $\mathbf{w}_{it}$, $\mathbf{w}_{ct}$ and $\mathbf{w}_t$, while N denotes the sampling rate. In our application $N = 3$ and $\mathbf{w}_{it} = (\Delta \ln IP_{it}, Infl_{it}, \ln GDP_{it} - \ln GDP_{i,t-3})^\top$, $i = 1, \ldots, n$, $\mathbf{w}_{ct} = \mathbf{y}_{ct} = \left(\Delta \ln p_{gas,t}, \Delta \ln p_{oil,t}\right)^\top$, and $\mathbf{w}_t := \left(\mathbf{w}_{1t}^\top, \ldots, \mathbf{w}_{nT}^\top, \mathbf{w}_{ct}^\top\right)^\top$. Let $\mathbf{w}_t^+ = \mathbf{w}_t$ for $t \in N\mathbb{Z}$, while for $t \notin N\mathbb{Z}$ the missing observations are replaced by a normally distributed random variable with mean zero and covariance matrix $\mathbf{R}_t = \mathbf{Q}_t\mathbf{Q}_t^\top$ (see also Seong et al., 2013). The information sets generated by $\mathbf{w}_t$ and $\mathbf{w}_t^+$ are denoted by $\mathscr{Y}_t$ (the σ-field generated by $\mathbf{y}_s : \ 0 < s \leq t$) and $\mathscr{Y}_t^+$ (the σ-field generated by $\mathbf{w}_s^+ : \ 0 < s \leq t$), respectively.[15] This

[15] If a slow stock variable was considered, the subsequent steps to obtain a sampling distribution for $\mathbf{y}_t$ could be

results in the following state space form, where we consider one slow flow variable (GDP). For each country i the monthly GDP growth rate is contained in the third coordinate of $\mathbf{y}_{it}$.[16] Let us introduce the following notation:

$$\begin{aligned}
\mathbf{w}_t^+ &= \mathbf{H}_t\mathbf{x}_t + \mathbf{Q}_t\boldsymbol{\eta}_t \ , \quad \boldsymbol{\eta}_t \sim \mathcal{N}\left(\mathbf{0}_{nk+k_c}, \mathbf{I}_{nk+k_c}\right) \\
\mathbf{x}_{t+1} &= \mathbf{F}\mathbf{x}_{t+1} + \mathbf{G}\boldsymbol{\varepsilon}_t \ , \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}\left(\mathbf{0}_{nk+k_c}, \boldsymbol{\Sigma}_t\right) \ , \text{ where} \qquad (19) \\
\mathbf{H}_{it} = \mathbf{H}^{obs} &:= \underbrace{\begin{pmatrix} \mathbf{I}_{k-1} & \mathbf{0}_{k-1\times k^2} & \mathbf{0}_{k-1\times k(p-k-1)+1} \\ \mathbf{0}_{1\times k-1} & \mathbf{1}_{1\times k}\otimes\mathbf{e}_{1k}^\top & \mathbf{0}_{1\times k(p-k-1)+1} \end{pmatrix}}_{[k\times kp]} \text{ for } t\in N\mathbb{Z} \text{ and } \mathbf{e}_{1k} = (1,0,\ldots,0)^\top \in \mathbb{R}^{k\times 1} \\
\mathbf{H}_{it} = \mathbf{H}^{notobs} &:= \underbrace{\begin{pmatrix} \mathbf{I}_{k-1} & \mathbf{0}_{k-1\times k(p-1)+1} \\ \mathbf{0}_{1\times k-1} & \mathbf{0}_{1\times k(p-1)+1} \end{pmatrix}}_{[k\times kp]} \text{ for } t\notin N\mathbb{Z}
\end{aligned}$$

$\mathbf{H}_{ct}$ of dimension $k_c \times k_c p$ is obtained conformingly

$$\begin{aligned}
\mathbf{H}_t &= \begin{pmatrix} \mathbf{I}_n\otimes\mathbf{H}_{it} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{ct} \end{pmatrix} \\
\mathbf{Q}_{it} &= \mathbf{0}_{[k\times k]} \text{ for } t\in N\mathbb{Z} \ , \ \mathbf{Q}_{it} = \underbrace{(\mathbf{0}_{k\times k-1} \ \mathbf{e}_{kk})}_{[k\times k]} \text{ for } t\notin N\mathbb{Z}
\end{aligned}$$

where $\mathbf{e}_{kk} = (0,\ldots,0,1)^\top \in \mathbb{R}^{k\times 1}$.

Note that $\mathbf{Q}_{ct}$ of dimension $k_c \times k_c$ is obtained conformingly; in our application $\mathbf{Q}_{ct} = \mathbf{0}_{2\times 2}$.

$$\mathbf{Q}_t = \begin{pmatrix} \mathbf{I}_n\otimes\mathbf{Q}_{it} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{ct} \end{pmatrix} . \qquad (20)$$

Recall that those coordinates of $\mathbf{y}_t$ observed every period are called the fast variables, while the coordinates only observed at $N\mathbb{Z}$, $N>1$, are called the slow variables. Define $\mathbf{x}_{t|T} := \mathbb{E}\left(\mathbf{x}_{it}|\mathscr{Y}_T^+\right)$, $\boldsymbol{\Pi}_{t|T} := \mathbb{C}ov\left(\mathbf{x}_t\mathbf{x}_t^\top|\mathscr{Y}_T^+\right)$ and $\boldsymbol{\Pi}_{t,t-1|T} := \mathbb{C}ov\left(\mathbf{x}_t\mathbf{x}_{t-1}^\top|\mathscr{Y}_T^+\right)$, $\boldsymbol{\theta}$ denotes the model parameters (for the Kalman filter as well as the Kalman smoother, see, e.g., Shumway and Stoffer, 1982; Deistler and Scherrer, 2018). Let

$$\begin{aligned}
\mathbf{K}_t &:= \mathbf{F}\boldsymbol{\Pi}_{t|t-1}\mathbf{H}_t^\top\boldsymbol{\Sigma}_{t|t-1}^{-1} \ , \\
\boldsymbol{\Pi}_{t+1|t} &= \mathbb{V}\left(\mathbf{x}_{t+1}-\mathbf{x}_{t+1|t}\right) = \mathbf{F}\boldsymbol{\Pi}_{t|t-1}\mathbf{F}^\top + \mathbf{G}\boldsymbol{\Sigma}\mathbf{G}^\top - \mathbf{K}_t\boldsymbol{\Sigma}_{t|t-1}\mathbf{K}_t^\top \ , \\
\mathbf{x}_{t+1|t} &= \mathbf{F}\mathbf{x}_{t|t-1} + \mathbf{K}_t\left(\mathbf{w}_t^+ - \mathbf{w}_{t|t-1}^+\right) \ , \\
\mathbf{w}_{t+1|t}^+ &= \mathbf{H}_{t+1}\mathbf{x}_{t+1|t} \ , \\
\boldsymbol{\Sigma}_{t+1|t} &= \mathbb{V}\left(\mathbf{w}_{t+1}^+ - \mathbf{w}_{t+1|t}^+\right) = \mathbf{H}_{t+1}\boldsymbol{\Pi}_{t+1|t}\mathbf{H}_{t+1}^\top + \mathbf{Q}_{t+1}\mathbf{I}\mathbf{Q}_{t+1}^\top \ , \qquad (21)
\end{aligned}$$

where $\mathbb{V}(\cdot)$ denotes a variance. The system is started at some $\mathbf{x}_{1|0} = (1,\mathbf{x}_0)^\top$ and $\boldsymbol{\Pi}_{1|0} = diag\left(1, \mathbf{1}_{pnk+pk_c}^\top\right)$, such that $\mathbf{w}_{1|0}^+ = \mathbf{H}_1\mathbf{z}_{1|0}$ and $\boldsymbol{\Sigma}_{1|0} = \mathbf{H}_1\boldsymbol{\Pi}_{1|0}\mathbf{H}_1^\top + \tilde{\mathbf{Q}}_1\tilde{\mathbf{Q}}_1^\top$. Note that for

adapted in a straightforward way. In the following we mainly focus on our application, where one slow flow variable is considered.

[16] Matrices $\mathbf{H}_{it}$ and $\mathbf{Q}_{it}$ are constructed for the case when only one slow variable is considered, namely GDP growth. This (flow) variable is the third coordinate of $\mathbf{w}_{it}$.

variables which cannot be observed, denoted $w_{jt}^{+,notobs}$, $\mathbb{E}(w_{jt}^{+,notobs}) = 0$, $\mathbb{V}(w_{jt}^{+,notobs} - w_{jt|t-1}^{+,notobs}) = \mathbb{V}(w_{jt}^{+,notobs}) = 1$, and $\mathbb{C}ov(w_{jt}^{+,notobs}, y_{it}) = 0$, for all $i \neq j$ and all t, where $\mathbb{C}ov(\cdot,\cdot)$ denotes the covariance matrix. For those time points where $\mathbf{w}_t^+ = \mathbf{w}_t$ we get $\mathbb{V}(\mathbf{w}_t - \mathbf{w}_{t|t-1}) = \mathbf{H}_t\mathbf{\Pi}_{t|t-1}\mathbf{H}_t^\top$. In the following forecasts we consider $\mathbf{w}_{t+h}$ (and not $\mathbf{w}_{t+h}^+$), where $\mathbf{H}_{t+h} = \begin{pmatrix} \mathbf{I}_n \otimes \mathbf{H}^{obs} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{ct} \end{pmatrix}$ for all $h > 0$. The h-step ahead forecasts, $h \geq 1$, follow from

$$\begin{aligned}
\mathbf{x}_{t+h|t} &= \mathbf{F}\mathbf{x}_{t+h-1|t} \ , \\
\mathbf{\Pi}_{t+h|t} &= \mathbb{V}\left(\mathbf{x}_{t+h} - \mathbf{x}_{t+h|t}\right) = \mathbf{F}\mathbf{\Pi}_{t+h-1|t}\mathbf{F}^\top + \mathbf{G}\mathbf{\Sigma}\mathbf{G}^\top \ , \\
\mathbf{w}_{t+h|t} &= \mathbf{H}_{t+h}\mathbf{x}_{t+h|t} \ , \\
\mathbf{\Sigma}_{t+h|t} &= \mathbb{V}\left(\mathbf{w}_{t+h} - \mathbf{w}_{t+h|t}\right) = \mathbf{H}_{t+h}\mathbf{\Pi}_{t+h|t}\mathbf{H}_{t+h}^\top \\
&\quad \text{such that for } h = 0 \text{ we get} \\
\mathbf{x}_{t|t} &= \mathbf{x}_{t|t-1} + \mathbf{\Pi}_{t|t-1}\mathbf{H}_t^\top\mathbf{\Sigma}_{t|t-1}^{-1}\left(\mathbf{w}_t^+ - \mathbf{w}_{t|t-1}^+\right) \ , \\
\mathbf{\Pi}_{t|t} &= \mathbb{V}\left(\mathbf{x}_t - \mathbf{x}_{t|t}\right) = \mathbf{\Pi}_{t|t-1} - \mathbf{\Pi}_{t|t-1}\mathbf{H}_t^\top\mathbf{\Sigma}_{t|t-1}^{-1}\mathbf{H}_t\mathbf{\Pi}_{t|t-1} \ . \qquad (22)
\end{aligned}$$

Note that $\mathbf{x}_{t|t} = \mathbb{E}\left(\mathbf{x}_t|\mathscr{Y}_t^+\right)$ and $\mathbf{\Pi}_{t|t} = \mathbb{V}\left(\mathbf{x}_t - \mathbf{x}_{t|t}\right) = \mathbb{V}\left(\mathbf{x}_t - \mathbb{E}\left(\mathbf{x}_t|\mathscr{Y}_t^+\right)\right)$. For the fast variables we get $\mathbf{x}_{jt|t} = \mathbb{E}\left(\mathbf{x}_{jt}|\mathscr{Y}_t^+\right) = x_{jt} = w_{jt}$, that is the conditional expectation is the actual observation of the variable j, while for the slow variables we get $\mathbf{x}_{jt|t} = \mathbb{E}\left(\mathbf{x}_{jt}|\mathscr{Y}_t^+\right)$ which is obtained by the above recursions. The lagged coordinates contained in $\mathbf{x}_t$ follow from these terms in a deterministic way. Only those elements of $\mathbf{\Pi}_{t|t}$ referring to covariances of slow variables are non-zero. This directly follows from the properties of conditional expectation. In addition,

$$\begin{aligned}
\mathbf{x}_{t|t} &= \mathbf{x}_{t|t-1} + \mathbf{\Pi}_{t|t-1}\mathbf{H}_t^\top\mathbf{\Sigma}_{t|t-1}^{-1}\left(\mathbf{w}_t^+ - \mathbf{w}_{t|t-1}^+\right) \\
&= \mathbf{x}_{t|t-1} + \mathbf{\Pi}_{t|t-1}\mathbf{H}_t^\top\left(\mathbf{H}_t\mathbf{\Pi}_{t|t-1}\mathbf{H}_t^\top\right)^{-1}\left(\mathbf{w}_t^+ - \mathbf{w}_{t|t-1}^+\right) \\
&= \left(1, w_{1t}^+, w_{2t}^+, x_{3t|t-1}, w_{4t}^+, w_{5t}^+, x_{6t|t-1}, \ldots, w_{(n-1)k+1,t}^+, w_{(n-1)*k+2,t}^+, x_{nk,t|t-1}, w_{c1t}^+, w_{c2t}^+\right)^\top \\
&= \left(1, y_{1t}, y_{2t}, x_{3t|t-1}, y_{4t}, y_{5t}, x_{6t|t-1}, \ldots, y_{(n-1)k+1,t}, y_{(n-1)*k+2,t}, x_{nk,t|t-1}, y_{c1t}, y_{c2t}\right)^\top \ . \qquad (23)
\end{aligned}$$

For those coordinates where w_{jt} is an observed fast variable, the conditional expectation given the past and the current observations is simply y_{jt} (this follows again from the properties of conditional expectation). The variables $x_{jt|t-1}$ $(= \mathbb{E}\left(\mathbf{x}_{jt}|\mathscr{Y}_t^+\right)$ as mentioned before) follow from the above recursion (23). The Kalman-smoothing equations for $t = T-1, \ldots, 2, 1$ are

$$\begin{aligned}
\mathbf{B}_{t+1} &= \mathbf{\Pi}_{t|t-1}\left(\mathbf{F}^\top - \mathbf{H}_t^\top\mathbf{K}_t^\top\right)\mathbf{\Pi}_{t+1|t}^{-1} \\
\mathbf{x}_{t|T} &= \mathbf{x}_{t|t} + \mathbf{B}_{t+1}\left(\mathbf{x}_{t+1|T} - \mathbf{x}_{t+1|t}\right) \\
\mathbf{\Pi}_{t|T} &= \mathbf{\Pi}_{t|t} + \mathbf{B}_{t+1}\left(\mathbf{\Pi}_{t+1|T} - \mathbf{\Pi}_{t+1|t}\right)\mathbf{B}_{t+1}^\top \ . \qquad (24)
\end{aligned}$$

Since $x_{jt|t} = x_{jt}$ for the fast variables, also $x_{jT|t} = x_{jt}$ and the corresponding variance terms in

$\mathbf{\Pi}_{t|T}$ are zero. Since $x_{jt|t} = x_{jT|t}$ the rows of $\mathbf{B}_{t+1}$ referring to fast variables have to be zero.

Assuming that the noise terms conditional on $\mathbf{x}_0$ are normally distributed, it follows from Frühwirth-Schnatter (1994) or Frühwirth-Schnatter (2006)[p. 419] that the missing values can be recursively drawn from a (degenerated) normal distribution with mean vector $\bar{\mathbf{x}}_{t|T}$ and covariance matrix $\bar{\mathbf{\Pi}}_{t|T}$, $t = T, T-1, \ldots, 1, 0$. In the following we slightly adapt the proof of Frühwirth-Schnatter (1994) to sample $\mathbf{y}_t = \mathbf{S}_y \mathbf{x}_t$; $\mathbf{S}_y$ is a $nk + k_c \times 1 + nkp + k_c p$ selector matrix (see also (17)) where $\mathbf{S}_y \mathbf{S}_y^\top = \mathbf{I}_{nk+k_c}$.[17] By the Bayes theorem we get

$$\pi\left(\mathbf{x}_t | \mathbf{y}_{t+1}, \ldots, \mathbf{y}_T, \mathscr{Y}_t^+, \boldsymbol{\theta}\right) \propto \pi\left(\mathbf{S}_y \mathbf{x}_{t+1} | \mathbf{x}_t, \boldsymbol{\theta}\right) \pi\left(\mathbf{x}_t | \mathscr{Y}_t^+, \boldsymbol{\theta}\right) \tag{25}$$

For $t \notin N\mathbb{Z} + 1$, the last density $\pi\left(\mathbf{x}_t | \mathscr{Y}_t^+, \boldsymbol{\theta}\right)$ is a normal density with mean vector $\mathbf{x}_{t|t}$ and covariance matrix $\mathbf{\Pi}_{t|t}$. The density $\pi\left(\mathbf{S}_y \mathbf{x}_{t+1} | \mathbf{x}_t, \boldsymbol{\theta}\right) = \pi\left(\mathbf{y}_{t+1} | \mathbf{x}_t, \boldsymbol{\theta}\right)$ is a normal density with mean vector $\mathbf{S}_y \mathbf{F} \mathbf{x}_t$ and covariance matrix $\mathbf{S}_y \mathbf{G} \mathbf{\Sigma}_t \mathbf{G}^\top \mathbf{S}_y^\top$. Since $\mathbf{x}_{t+1}$ contains $\mathbf{y}_{t+1}, \mathbf{y}_t, \ldots, \mathbf{y}_{t-p+2}$, $[\mathbf{x}_{t+1}]_{(1+(nk+k_c):1+p(nk+k_c))}$ deterministically follows from $[\mathbf{x}_t]_{(2:1+(p-1)(nk+k_c))}$. Hence, $\mathbf{x}_{t+1}$ follows a singular normal distribution with mean vector $\mathbf{F}\mathbf{x}_t$ and covariance matrix $\mathbf{G}\mathbf{\Sigma}_t\mathbf{G}^\top$. From the appendix in Frühwirth-Schnatter (1994) we know that by "completing the square" in the corresponding state space model we arrive at a normal distribution with a mean vector of the form $\bar{\mathbf{x}}_{t|T} = \left(\mathbf{I} - \mathbf{B}_{t+1}\mathbf{F}\right)\mathbf{x}_{t|t} + \mathbf{B}_{t+1}\mathbf{x}_{t+1}$ and a covariance matrix of the form $\bar{\mathbf{P}}_{t|T} = \left(\mathbf{I} - \mathbf{B}_{t+1}\mathbf{F}\right)\mathbf{\Pi}_{t|t}$. For our application this result and the relationship between $\mathbf{x}_t$ and $\mathbf{y}_t, \ldots, \mathbf{y}_{t-N+1}$ shows that $\mathbf{y}_t$ follows a normal distribution with mean vector $\bar{\mathbf{y}}_{t|T}$ and covariance matrix $\mathbf{S}_y \bar{\mathbf{P}}_{t|T} \mathbf{S}_y^\top$. Hence, for $t \notin N\mathbb{Z} + 1$, samples of $\mathbf{y}_t$ follow from:

$$\begin{aligned}
\mathbf{y}_t | \mathbf{y}_{t+1}, \ldots, \mathbf{y}_T, \mathscr{Y}_T^+ &\sim \mathcal{N}\left(\bar{\mathbf{y}}_{t|T}, \mathbf{S}_y \bar{\mathbf{P}}_{t|T} \mathbf{S}_y^\top\right) \text{ , where} \\
\bar{\mathbf{y}}_{t|T} &= \mathbf{S}_y\left(\mathbf{I} - \mathbf{B}_{t+1}\mathbf{F}\right)\mathbf{x}_{t|t} + \mathbf{B}_{t+1}\mathbf{S}_y^T \mathbf{y}_{t+1} \\
&= \mathbf{S}_y \mathbf{x}_{t|t} + \mathbf{S}_y \mathbf{B}_{t+1} \mathbf{S}_y^\top \left(\mathbf{y}_{t+1} - \mathbf{S}_y \mathbf{F} \mathbf{x}_{t|t}\right) = \mathbf{S}_y \mathbf{x}_{t|t} + \mathbf{S}_y \mathbf{B}_{t+1} \mathbf{S}_y^\top \left(\mathbf{y}_{t+1} - \mathbf{S}_y \mathbf{x}_{t+1|t}\right) \text{ ,} \\
\bar{\mathbf{P}}_{t|T} &= \left(\mathbf{I} - \mathbf{B}_{t+1}\mathbf{F}\right)\mathbf{\Pi}_{t|t} \text{ ,} \\
\mathbf{B}_{t+1} &= \mathbf{\Pi}_{t|t}\mathbf{F}^\top \left(\mathbf{F}_t \mathbf{\Pi}_{t|t} \mathbf{F}^\top + \mathbf{G}\mathbf{\Sigma}_t\mathbf{G}^\top\right)^{-1} \text{ ,}
\end{aligned} \tag{26}$$

where in our application $\mathbf{x}_t$ directly follows from $\mathbf{y}_t, \ldots, \mathbf{y}_{t-N+1}$. By plugging in terms obtained above and additional calculations we get

$$\begin{aligned}
\bar{\mathbf{P}}_{t|T} &= \left(\mathbf{I} - \mathbf{B}_{t+1}\mathbf{F}\right)\mathbf{\Pi}_{t|t} = \mathbf{\Pi}_{t|t} - \mathbf{B}_{t+1}\mathbf{F}\mathbf{\Pi}_{t|t} \\
&= \mathbf{\Pi}_{t|t} - \mathbf{\Pi}_{t|t}\mathbf{F}^\top \underbrace{\left(\mathbf{F}_t \mathbf{\Pi}_{t|t} \mathbf{F}^\top + \mathbf{G}\mathbf{\Sigma}_t\mathbf{G}^\top\right)^{-1}}_{=\mathbf{\Pi}_{t+1|t}^{-1} \text{by (22)}} \mathbf{F}\mathbf{\Pi}_{t|t} \\
&= \mathbf{\Pi}_{t|t} - \mathbf{\Pi}_{t|t}\mathbf{F}^\top \mathbf{\Pi}_{t+1|t}^{-1} \mathbf{\Pi}_{t+1|t} \mathbf{\Pi}_{t+1|t}^{-1\top} \mathbf{F}\mathbf{\Pi}_{t|t}^\top \\
&= \mathbf{\Pi}_{t|t} - \mathbf{B}_{t+1}\mathbf{\Pi}_{t+1|t}\mathbf{B}_{t+1}^\top = \mathbf{\Pi}_{t|t} - \mathbf{B}_{t+1}\mathbb{V}\left(\mathbf{x}_{t+1} - \mathbf{x}_{t+1|t}\right)\mathbf{B}_{t+1}^\top \text{ .}
\end{aligned} \tag{27}$$

Thus, for those periods t where $t \notin N\mathbb{Z} + 1$, samples of $\mathbf{y}_t$ follow from (26). Since $\mathbf{\Pi}_{t|t}$ is a sparse matrix, we sample from a singular normal distribution.

Finally, for $t \in N\mathbb{Z} + 1$, $\mathbf{y}_{t+1}, \ldots, \mathbf{y}_{t+N-1}$ and $\mathbf{w}_{t+N-1}^+ = \mathbf{w}_{t+N-1}$ allow to calculate y_{jt} by

[17] To simplify notation we often do not distinguish between samples obtained by means of (26) and the random variables $\mathbf{y}_t$.

means of $y_{jt} = w_{j,t+N-1} - y_{jt+1} - \cdots - y_{jt+N-1}$ for all slow coordinates j. Hence, in formal terms in this case the conditional distribution of $\mathbf{y}_t|\mathbf{y}_{t+1}, \ldots \mathbf{y}_T, \mathscr{Y}_t^+$ is a Dirac distribution with point mass on the observed fast variables and $y_{jt} = w_{j,t+N-1} - \mathbf{y}_{jt+1} - \cdots - \mathbf{y}_{jt+N-1}$ for the slow coordinates j. Hence, for the periods $t \in N\mathbb{Z}+1$ and slow variables with index j (that is, for the first month of the corresponding quarter in our application), we get $x_{jt|T} = \mathbb{E}\left(x_{jt}|\mathscr{Y}_T^+, \mathbf{x}_{t+1}, \mathbf{x}_{t+2}, \ldots\right) = y_{j,t+N-1}^+ - y_{j,t+1} - \cdots - y_{j,t+N-1}$, for those periods $t = s+1$, $t \in N\mathbb{Z}$, after s where $\mathbf{w}_s = \mathbf{w}_s^+$. The variance of this term is zero. The lagged coordinates contained in $\mathbf{x}_t$ follow from these terms in a deterministic way.

For our application this implies: We observe a quarterly growth rate of GDP, e.g., for the first quarter. The monthly GDP growth rates for March and February follow from samples obtained by means of (26). The monthly GDP for January directly follows from the quarterly GDP growth rate (from January to March) minus the monthly growth rates sampled first for March and then for February.

B.2 Convergence and Mixing

This section analyzes the convergence and mixing properties of our Bayesian sampler for the models with five variables. We consider the M_1 posterior draws $\theta_j^{(m)}$, $m = M_0+1, \ldots, M$, $M = M_0 + M_1$, for each country $i = 1, \ldots, n$. In the following $M_0 = 2,000$ and $M_1 = 8,000$.

Mixing of the Chain: To investigate the mixing behavior of the chain we derive the effective sample size $\widehat{M}_j^{eff}$ as, e.g., defined in Gelman et al. (2013)[Chapter 11.5]; here the `coda` package in `R` was applied.

Table 3 presents the average effective sample sizes (i.e., we take the sample mean of the effective sample sizes obtained for the corresponding parameters contained in $\boldsymbol{\alpha}$, $vech(\boldsymbol{\Sigma})$, etc.). In the Case 1, where the noise terms follow a t-distribution, we consider the parameter subvectors $\boldsymbol{\alpha}$, $vech(\boldsymbol{\Sigma})$, $\boldsymbol{\lambda}$, and ν. For $\boldsymbol{\varepsilon}_t$ generated by a stochastic volatility model (Case 2) we consider the parameter subvector $\boldsymbol{\alpha}$ and samples of the volatility matrix $\boldsymbol{\Sigma}_t$ at $t = T$ (too keep the amount of MCMC output to be stored low we only store the last value of the volatility process $(vech\left(\boldsymbol{\Sigma}_t\right))_{t=0,1,\ldots,T}$). For some parameters the effective sample size is larger than M_1 which can be explained by the estimation of the long run covariance matrix to obtain $\widehat{M}_j^{eff}$. In both cases we observe for the parameters $\boldsymbol{\alpha}$, the volatility parameters, and ν, that the average effective sample size is larger than 700 based on 8,000 MCMC draws. Only for $\boldsymbol{\lambda}$ the effective sample size remained relatively low, which can be explained by the relatively high persistence of the samples of λ_t. When applying the stochastic volatility model we observe that the average effective sample size is at least 5,000.

Convergence: Since the main focus of this paper is on forecasting, we run the Bayesian sampler with different *seeds* and compare all the fan charts (= distributions of Bayesian point forecasts) for the inflation forecasts by means of visual inspection. Here we observed that the fan charts strongly overlap also for different seeds. Therefore, we can conclude that the Bayesian sampler has sufficient convergence and mixing behaviour.

C Unobserved components model with stochastic volatility

We use the unobserved components model with stochastic volatility described in Kroese et al. (2014) as one of the benchmark models in the forecast evaluation. The observable variable, in our

	Case 1: t-distributed noise				Case 2: stochastic vol.	
	$\boldsymbol{\alpha}$	$\boldsymbol{\Sigma}$	$\boldsymbol{\lambda}$	ν	$\boldsymbol{\alpha}$	$\boldsymbol{\Sigma}_T$
Austria	1163	2012	101	1962	6261	7297
Belgium	719	1959	92	1962	5087	7023
Germany	864	2060	101	1820	5611	6824
Finland	1034	2030	120	1962	6134	6829
Italy	843	1998	67	2051	8048	7906
Slovakia	1096	2001	159	1962	5258	6918

Table 3: Average effective sample size $\widehat{M}_j^{eff}$.

case $Infl_t$, depends on the unobserved component, τ_t, and a stochastic volatility term,

$$Infl_t = \tau_t + e^{\tilde{h}_t/2}\epsilon_t, \tag{28}$$

where $\epsilon_t \sim iid\ N(0,1)$. The unobserved component and log-volatility, $\tilde{h}_t$, both follow a random walk, that is

$$\begin{aligned} \tau_t &= \tau_{t-1} + u_t, \\ \tilde{h}_t &= \tilde{h}_{t-1} + \nu_t, \end{aligned} \tag{29}$$

where $u_t \sim iid\ N(0,\omega_\tau^2)$ and $\nu_t \sim iid\ N(0,\omega_{\tilde{h}}^2)$. The state equations are initialized with $\tau_1 \sim N(\tau_0, V_\tau)$ and $\tilde{h}_1 \sim N(\tilde{h}_0, V_{\tilde{h}})$, where $\tau_1 = \tilde{h}_1 = 0$ and $V_\tau = V_{\tilde{h}} = 9$. We assume independent inverse-gamma priors for ω_τ^2 and $\omega_{\tilde{h}}^2$, namely

$$\begin{aligned} \omega_\tau^2 &\sim \mathcal{IG}(\tilde{\alpha}_\tau, \tilde{\lambda}_\tau)\ , \\ \omega_h^2 &\sim \mathcal{IG}(\tilde{\alpha}_{\tilde{h}}, \tilde{\lambda}_h), \end{aligned} \tag{30}$$

with $\tilde{\alpha}_\tau = \tilde{\alpha}_{\tilde{h}} = 10$ and $\tilde{\lambda}_\tau = 0.25^2(\tilde{\alpha}_\tau - 1)$ and $\tilde{\lambda}_{\tilde{h}} = 0.2^2(\tilde{\alpha}_{\tilde{h}} - 1)$. The stochastic volatility model is estimated by auxiliary mixture sampling, where the appropriate Gaussian mixture is chosen as proposed by Kim et al. (1998). We use the code UCSV.R of Kroese et al. (2014) to obtain 10,000 posterior draws after a burn-in of 2,000 draws for each estimation. For further details see Kroese et al. (2014).

References

Antolin-Diaz, J., Drechsel, T., and Petrella, I. (2021). Advances in nowcasting economic activity: Secular trends, large shocks and new data. Technical report, CEPR Discussion Paper No. DP15926.

Armendariz, S., Geis, A., Myrvoda, A., G., D., Chen, C., Kirabaeva, K., E., M., Minnett, D., Parry, I., Tim, T., and von Thadden-Kostopoulos, S. (2023). Kingdom of the Netherlands - the Netherlands. Selected issues. *IMF Country Report*, 107.

Banbura, M., Leiva-Leon, D., and Menz, J.-O. (2021). Do inflation expectations improve model-based inflation forecasts? Technical report, Banco de Espana Working Paper.

Bitto, A. and Frühwirth-Schnatter, S. (2019). Achieving shrinkage in a time-varying parameter model framework. *Journal of Econometrics*, 210(1):75 – 97. Annals Issue in Honor of John Geweke "Complexity and Big Data in Economics and Finance: Recent Developments from a Bayesian Perspective"'.

Bobeica, E. and Hartwig, B. (2023). The covid-19 shock and challenges for inflation modelling. *International Journal of Forecasting*, 39(1):519–539.

Carriero, A., Clark, T. E., Marcellino, M., and Mertens, E. (2022). Addressing covid-19 outliers in bvars with stochastic volatility. *Review of Economics and Statistics*, pages 1–38.

Chan, J. C. (2013). Moving average stochastic volatility models with application to inflation forecast. *Journal of Econometrics*, 176(2):162–172.

Chan, J. C. and Hsiao, C. Y. (2014). *Estimation of Stochastic Volatility Models with Heavy Tails and Serial Dependence*, chapter 6, pages 155–176. John Wiley & Sons, Ltd.

Clark, T. E. (2011). Real-time density forecasts from bayesian vector autoregressions with stochastic volatility. *Journal of Business & Economic Statistics*, 29(3):327–341.

Clark, T. E., Huber, F., Koop, G., Marcellino, M., and Pfarrhofer, M. (2023). Tail forecasting with multivariate bayesian additive regression trees. *International Economic Review*, 64(3):979–1022.

Clark, T. E. and Ravazzolo, F. (2015). Macroeconomic forecasting performance under alternative specifications of time-varying volatility. *Journal of Applied Econometrics*, 30(4):551–575.

Deistler, M. and Scherrer, W. (2018). *Modelle der Zeitreihenanalyse*. Mathematik Kompakt. Springer International Publishing.

Frühwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, 15(2):183–202.

Frühwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models*. Springer Series in Statistics. Springer.

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., and Rubin, D. (2013). *Bayesian Data Analysis, Third Edition*. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis.

Geweke, J. (1993). Bayesian treatment of the independent student-t linear model. *Journal of Applied Econometrics*, 8(S1):S19–S40.

Geweke, J., Koop, G., van Dijk, H., and van Dijk, H. (2011). *The Oxford Handbook of Bayesian Econometrics*. Oxford Handbooks in Economics. OUP Oxford.

Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378.

Hosszejni, D. and Kastner, G. (2021). Modeling univariate and multivariate stochastic volatility in r with stochvol and factorstochvol. *Journal of Statistical Software*, 100(12):1–34.

Jonckheere, J. (2022). Energy prices and inflation: it's complicated. *National Bank of Belgium - Blog*.

Kastner, G. (2016). Dealing with stochastic volatility in time series using the r package stochvol. *Journal of Statistical Software*, 69(5):1–30.

Kastner, G. (2019). Sparse bayesian time-varying covariance estimation in many dimensions. *Journal of Econometrics*, 210(1):98–115. Annals Issue in Honor of John Geweke "Complexity and Big Data in Economics and Finance: Recent Developments from a Bayesian Perspective".

Kilian, L. and Lütkepohl, H. (2017). *Structural Vector Autoregressive Analysis*. Themes in Modern Econometrics. Cambridge University Press.

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with arch models. *The Review of Economic Studies*, 65(3):361–393.

Koop, G. and Korobilis, D. (2019). Forecasting with high-dimensional panel vars. *Oxford Bulletin of Economics and Statistics*, 81(5):937–959.

Koop, G. and Korobilis, D. (2021). *Bayesian Multivariate Time Series Methods for Empirical Macroeconomics*. University of Strathclyde, 27 September 2009.

Kroese, D. P., Chan, J. C., et al. (2014). *Statistical Modeling and Computation*. Springer.

Krüger, F., Clark, T. E., and Ravazzolo, F. (2017). Using entropic tilting to combine bvar forecasts with external nowcasts. *Journal of Business & Economic Statistics*, 35(3):470–485.

Lenza, M. and Primiceri, G. E. (2022). How to estimate a vector autoregression after march 2020. *Journal of Applied Econometrics*, 37(4):688–699.

Lütkepohl, H. (2006). *New Introduction to Multiple Time Series Analysis*. Springer, Berlin Heidelberg.

Martin, G. M., Frazier, D. T., Maneesoonthorn, W., Loaiza-Maya, R., Huber, F., Koop, G., Maheu, J., Nibbering, D., and Panagiotelis, A. (2024). Bayesian forecasting in economics and finance: A modern review. *International Journal of Forecasting*, 40:811–839.

Pesaran, H. and Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics Letters*, 58(1):17 – 29.

Proietti, T. and Giovannelli, A. (2021). Nowcasting monthly GDP with big data: A model averaging approach. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 184(2):683–706.

Schorfheide, F. and Song, D. (2021). Real-time forecasting with a (standard) mixed-frequency var during a pandemic. Technical report, National Bureau of Economic Research.

Seong, B., Ahn, S. K., and Zadrozny, P. A. (2013). Estimation of vector error correction models with mixed-frequency data. *Journal of Time Series Analysis*, 34(2):194–205.

Shumway, R. H. and Stoffer, D. S. (1982). An approach to time series smoothing and forecoasting using the EM algorithm. *Journal of Time Series Analysis*, 3(4):253–264.

Stock, J. H. and Watson, M. W. (2007). Why has u.s. inflation become harder to forecast? *Journal of Money, Credit and banking*, 39:3–33.

Vaden, T., Majava, A., M., K. J., and Eronen, J. T. (2022). Energy Without Russia:The Case of Finland. *Friedrich-Ebert-Stiftung, Country Report*.