Online Appendix: Macroeconomic Forecasting in a Multi-country Context

Yu Bai*[1], Andrea Carriero[2], Todd E. Clark[3], and Massimiliano Marcellino[4]

[1]Bocconi University

[2]Queen Mary University of London, and University of Bologna

[3]Federal Reserve Bank of Cleveland

[4]Bocconi University, IGIER, and CEPR

A Proof of Theorem 1

In the following proof, the notation $\sim$ indicates asymptotic equivalence. We say that a is asymptotically equivalent to b if $a/b = O(1)$.

As shown in equation (14) of Bitto and Frühwirth-Schnatter (2019), the marginal density for $\beta_j \sim \mathcal{NG}(\lambda, \kappa)$, given λ, κ, can be expressed as

$$\pi_{NG}(\beta_j) = \frac{(\sqrt{\lambda\kappa})^{\lambda+\frac{1}{2}}}{\sqrt{\pi}2^{\lambda-\frac{1}{2}}\Gamma(\lambda)}\left|\beta_j\right|^{\lambda-\frac{1}{2}} K_{\lambda-\frac{1}{2}}\left(\sqrt{\lambda\kappa}\left|\beta_j\right|\right),$$

where $K_p(\cdot)$ is the modified Bessel function of the second kind of index p. Let us first consider the concentration properties. If $\lambda > \frac{1}{2}$, then $\lambda - \frac{1}{2} > 0$. By 9.6.9 in Abramowitz and Stegun (1965), as $\left|\beta_j\right| \to 0$,

$$K_{\lambda-\frac{1}{2}}\left(\sqrt{\lambda\kappa}\left|\beta_j\right|\right) \sim \frac{1}{2}\Gamma(\lambda - \frac{1}{2})\left(\frac{1}{2}\sqrt{\lambda\kappa}\left|\beta_j\right|\right)^{\frac{1}{2}-\lambda}.$$

*Correspondence to: Yu Bai, Centre for Applied Research on International Markets, Money Banking and Regulation, Via Röntgen 1, Milan 20136, Italy. E-mail: yu.bai@bocconialumni.it

Then,

$$\pi_{NG}(\beta_j) \sim \frac{(\sqrt{\lambda\kappa})^{\lambda+\frac{1}{2}}}{\sqrt{\pi}2^{\lambda-\frac{1}{2}}\Gamma(\lambda)}\left|\beta_j\right|^{\lambda-\frac{1}{2}} \times \frac{1}{2}\Gamma(\lambda-\frac{1}{2})\Big(\frac{1}{2}\sqrt{\lambda\kappa}\left|\beta_j\right|\Big)^{\frac{1}{2}-\lambda}$$
$$= \frac{\sqrt{\lambda\kappa}}{\sqrt{\pi}}\frac{\Gamma(\lambda-\frac{1}{2})}{\Gamma(\lambda)} \times \frac{1}{2} = O(1).$$

If $0 < \lambda < \frac{1}{2}$, recall that

$$K_{\lambda-\frac{1}{2}}\Big(\sqrt{\lambda\kappa}\left|\beta_j\right|\Big) = \frac{1}{2}\pi\frac{I_{\frac{1}{2}-\lambda}(\sqrt{\lambda\kappa}\left|\beta_j\right|) - I_{\lambda-\frac{1}{2}}(\sqrt{\lambda\kappa}\left|\beta_j\right|)}{\sin((\lambda-\frac{1}{2})\pi)},$$

where $I_p(\cdot)$ is the modified Bessel function of the first kind with index p. By 9.6.7 in Abramowitz and Stegun (1965), as $\left|\beta_j\right| \to 0$,

$$I_{\frac{1}{2}-\lambda}(\sqrt{\lambda\kappa}\left|\beta_j\right|) \sim \frac{\Big(\frac{1}{2}\sqrt{\lambda\kappa}\left|\beta_j\right|\Big)^{\frac{1}{2}-\lambda}}{\Gamma(\frac{3}{2}-\lambda)}$$
$$I_{\lambda-\frac{1}{2}}(\sqrt{\lambda\kappa}\left|\beta_j\right|) \sim \frac{\Big(\frac{1}{2}\sqrt{\lambda\kappa}\left|\beta_j\right|\Big)^{\lambda-\frac{1}{2}}}{\Gamma(\frac{3}{2}-\lambda)}.$$

Thus,

$$\pi_{NG}(\beta_j) \sim \frac{(\sqrt{\lambda\kappa})^{\lambda+\frac{1}{2}}}{\sqrt{\pi}2^{\lambda-\frac{1}{2}}\Gamma(\lambda)}\left|\beta_j\right|^{\lambda-\frac{1}{2}} \times \frac{1}{2}\pi \times \frac{1}{\sin((\lambda-\frac{1}{2})\pi)} \times \left(\frac{\Big(\frac{1}{2}\sqrt{\lambda\kappa}\left|\beta_j\right|\Big)^{\frac{1}{2}-\lambda}}{\Gamma(\frac{3}{2}-\lambda)} - \frac{\Big(\frac{1}{2}\sqrt{\lambda\kappa}\left|\beta_j\right|\Big)^{\lambda-\frac{1}{2}}}{\Gamma(\frac{3}{2}-\lambda)}\right)$$
$$= C - \frac{(\sqrt{\lambda\kappa})^{2\lambda} \times (\frac{1}{2})^{2\lambda}\sqrt{\pi}}{\sin((\lambda-\frac{1}{2})\pi)\Gamma(\lambda)\Gamma(\lambda+\frac{1}{2})}\Big(\frac{1}{\left|\beta_j\right|}\Big)^{1-2\lambda}$$
$$= O\bigg(\Big(\frac{1}{\left|\beta_j\right|}\Big)^{1-2\lambda}\bigg).$$

Finally, if $\lambda = \frac{1}{2}$, by 9.6.8 in Abramowitz and Stegun (1965),

$$K_0\Big(\sqrt{\frac{1}{2}\kappa}\left|\beta_j\right|\Big) \sim -\log(\sqrt{\frac{1}{2}\kappa}\left|\beta_j\right|).$$

Then,

$$\pi_{NG}(\beta_j) \sim \frac{\sqrt{\frac{1}{2}\kappa}}{\pi} \times -\log(\sqrt{\frac{1}{2}\kappa}\left|\beta_j\right|) = O\Big(\frac{1}{\log(\left|\beta_j\right|)}\Big).$$

We now move to the asymptotic tail behavior. By 9.7.2 in Abramowitz and Stegun (1965), as $|\beta_j| \to \infty$,

$$K_{\lambda-\frac{1}{2}}\left(\sqrt{\lambda\kappa}|\beta_j|\right) \sim \sqrt{\frac{\pi}{2\sqrt{\lambda\kappa}|\beta_j|}}e^{-\sqrt{\lambda\kappa}|\beta_j|}.$$

Then,

$$\begin{aligned}\pi_{NG}(\beta_j) &\sim \frac{(\sqrt{\lambda\kappa})^{\lambda+\frac{1}{2}}}{\sqrt{\pi}2^{\lambda-\frac{1}{2}}\Gamma(\lambda)}|\beta_j|^{\lambda-\frac{1}{2}}\sqrt{\frac{\pi}{2}}(\sqrt{\lambda\kappa})^{-\frac{1}{2}}|\beta_j|^{-\frac{1}{2}}\exp\left(-\sqrt{\lambda\kappa}|\beta_j|\right)\\ &= O\left(\frac{|\beta_j|^{\lambda-1}}{\exp\left(\sqrt{\lambda\kappa}|\beta_j|\right)}\right),\end{aligned}$$

which completes the proof.

B Model specifications and priors

B.1 Country-specific VARs

The country-specific VAR(p) model — denoted the CVAR specification in the paper's results — is specified as

$$\begin{aligned} y_{i,t} &= c_i + \sum_{l=1}^{p} B_{i,l} y_{i,t-l} + u_{i,t} && (1)\\ u_{i,t} &= A_i^{-1} H_{i,t}^{0.5} \epsilon_{i,t}, \epsilon_{i,t} \sim i.i.d.\ N(0, I_G), && (2)\end{aligned}$$

where $i = 1, \ldots, N$, $t = 1, \ldots, T$, and the dimension of $y_{i,t}$, $u_{i,t}$ and $\epsilon_{i,t}$ is $G \times 1$. A_i^{-1} is a lower triangular matrix with diagonal elements equal to 1. $H_{i,t}$ is diagonal with generic j-th element $h_{ij,t}$ evolving as a random walk (RW):

$$\ln h_{ij,t} = \ln h_{ij,t-1} + e_{ij,t}, j = 1, \ldots, G, \quad (3)$$

where $e_{it} \sim N(0, \Phi_i)$ with a full covariance matrix Φ_i as in Primiceri (2005).

Letting $B_i = [c_i, B_{i,1}, \ldots, B_{i,p}]'$, the priors are specified as:

$$\begin{aligned} vec(B_i) &\sim N(0, \underline{\Omega}_{B_i})\\ vec(A_i) &\sim N(0, \underline{\Omega}_{A_i})\\ \Phi_i &\sim IW(Q_0, W_0).\end{aligned}$$

For the prior variances of the autoregressive coefficient matrices, we set them as in the Minnesota prior:

$$\underline{\Omega}_{B_{i,l}^{(mn)}} = \begin{cases} \frac{\lambda_1}{l^{\lambda_3}} \frac{1}{\sigma_n^2} & \text{for the coefficients on own lags} \\ \frac{\lambda_2}{l^{\lambda_3}} \frac{\sigma_m^2}{\sigma_n^2} & \text{for the coefficients on cross-variable lags} \\ \lambda_0 \sigma_m^2 & \text{for the intercept,} \end{cases} \tag{4}$$

where $m, n = 1, \ldots, G$. λ_1 measures the overall tightness to coefficients related to own lags. λ_2 is related to cross-variable shrinkage. We assume Gamma priors for these two hyperparameters: $\lambda_1 \sim \mathcal{G}(1, 0.04)$, $\lambda_2 \sim \mathcal{G}(1, 0.04^2)$. λ_3 determines the additional shrinkage for coefficients associated with higher order lags and is set to 2 (quadratic decay). The scale parameters σ_m^2, σ_n^2 are obtained from univariate AR(1) regressions. We elicit an uninformative prior for the intercept by setting $\lambda_0 = 100$. In the case of the free elements in the contemporaneous matrix A_i, we set the prior mean to 0 and the prior variance to be non-informative: $\underline{\Omega}_{A_i} = 10 \times I$. Finally, as in the previous section, we follow the literature and set a modestly informative prior for Φ: $\Phi \sim IW(Q_0, W_0)$, where Q_0, W_0 take very conservative values: $W_0 = 0.01 \times I$ and $Q_0 = G + 2$.[1]

For the country-specific VAR with hierarchical shrinkage (CVAR-H), we follow exactly the approach in Chan (2021). Following Chan, the reduced-form model (1) is expressed in structural form

$$A_i y_{i,t} = \tilde{c}_i + \sum_{l=1}^{p} \tilde{B}_{i,l} y_{i,t-l} + H_{i,t}^{0.5} \epsilon_{i,t},$$

where $\tilde{c}_i = A c_i$, $\tilde{B}_{i,l} = A B_{i,l}$, and the innovations $e_{ij,t}$ in (3) are assumed to be independent across variables (equation $j = 1, \ldots, G$ of the VAR for country i): $e_{ij,t} \sim N(0, \sigma_{e_{ij}}^2)$. The priors are specified as

$$\beta_{i,j} | \lambda_1, \lambda_2, \psi_{i,j}, C_{i,j} \sim N(0, 2\lambda_{i,j} \psi_{i,j} C_{i,j}),$$

where $\lambda_{i,j}$ equals λ_1 if $\beta_{i,j}$ are related to own lags but equals λ_2 for coefficients related to cross-variable lags. $C_{i,j}$ are specified according to

$$C_{i,l} = \left\{ \begin{array}{ll} \frac{1}{l^{\lambda_3}} \frac{1}{\sigma_n^2} & \text{for the coefficients on own lags} \\ \frac{1}{l^{\lambda_3}} \frac{\sigma_m^2}{\sigma_n^2} & \text{for the coefficients on cross-variable lags} \end{array} \right\}$$

and $\psi_{i,j}$ are assumed to follow a Gamma prior:

$$\psi_{i,j} \sim \mathcal{G}(\nu_\psi, \nu_\psi / 2),$$

with an additional hyper-prior on $\nu_\psi \sim \mathcal{G}(1, 1)$. For $\sigma_{e_{ij}}^2$, priors are assumed to be $\sigma_{e_{ij}}^2 \sim \mathcal{IG}(5, 0.04)$.

[1] See, e.g., D'Agostino et al. (2013) and Clark and Ravazzolo (2015).

In Section 6.4, we also consider a version of model (1)-(2) with hierarchical shrinkage and Horseshoe prior. Similarly to the definitions of Section 3, let β_c, β_{AR}, and β_o be the coefficients related to intercept, own lags, and cross-variable lags, and let $\beta_{i,j}$ be the jth elements in the coefficient block i, where $i = \{c, AR, o\}$. In this case, we replace the prior specification in (4) by assuming that $\beta_{i,j}$ follows (10) (in the main paper), where the global shrinkage parameter λ differs in each coefficient block.

B.2 Factor-augmented country-specific VARs

The factor-augmented country-specific VAR (CFAVAR) takes the form:

$$\begin{aligned}
\begin{bmatrix} y_{i,t} \\ F_t \end{bmatrix} &= c_i + \sum_{l=1}^{p} B_{i,l} \begin{bmatrix} y_{i,t-l} \\ F_{t-l} \end{bmatrix} + u_{i,t} \\
Y_t^* &= \Lambda F_t + \varepsilon_t \\
F_t &= \sum_{l=1}^{q} \Pi_l F_{t-l} + v_t, v_t \sim i.i.d.\ N(0, \Sigma_v),
\end{aligned}$$

where $Y_t^* = (y'_{1,t}, \ldots, y'_{i-1,t}, y'_{i+1,t}, y'_{N,t})'$ is the collection of foreign variables. F_t is an $r \times 1$ vector of weakly exogenous unobservable factors representing foreign information, which affect the variables in country i via the loadings $B^*_{i,l}, i = 1, \ldots, N, l = 1, \ldots, p$. Factors are estimated (recursively, as forecasting moves forward in time and the estimation sample expands) by principal components (see, e.g., Stock and Watson (2002a) and Stock and Watson (2002b)) and assumed to follow a VAR process with lag length q. In the VAR for $[y_{i,t}, F_t]$, the innovation vector $u_{i,t}$ includes the stochastic volatility structure previously indicated in the country-specific VAR's equation (2).

Priors for c_i and $B_{i,l}$ are specified in the same way as in the country-specific VARs. The same hyper-priors are imposed on (λ_1, λ_2), which are the overall tightness parameters on coefficients related to own lags and cross-variable lags. We specify the maximum number of factors and lag length to be $r^{\max} = 4$ and $q^{\max} = 4$, respectively. The number of factors is determined by the $IC2$ information criterion of Bai and Ng (2002), and the number of lags is determined by the Bayesian Information Criterion (BIC). The VAR for the factors is separately estimated by Bayesian methods with non-informative priors. Specifically, letting $\pi = vec([\Pi_1, \ldots, \Pi_q]')$, we specify $\pi \sim N(0, 100 \times I_{r^2 q})$. Following Korobilis (2016), $\hat{\Sigma}_v$ is fixed at the OLS estimate to streamline computations (it also eliminates the uncertainty associated with covariance matrix estimation).

B.3 Global VARs

A GVAR model consists of a number of country-specific equations that are combined to form a global model. Assuming that the global economy consists of $N + 1$ countries, in the first step, we estimate the following country-specific VARX model for every country $i = 0, 1, ..., N$:

$$y_{i,t} = c_i + \sum_{l=1}^{p} B_{i,l} y_{i,t-l} + \sum_{l=0}^{p^*} B^*_{i,l} y^*_{i,t-l} + u_{i,t}, \tag{5}$$

$$u_{i,t} = A_i^{-1} H_{i,t}^{0.5} \epsilon_{i,t}, \quad \epsilon_{i,t} \overset{i.i.d.}{\sim} N(0, I_G), \tag{6}$$

where $t = 1, ..., T$, $y_{i,t}$ is a $G \times 1$ vector of endogenous variables in country i, c_i is a $G \times 1$ vector of intercept terms, $B_{i,l}(l = 1, ..., p)$ denotes the $G \times G$ matrix of parameters associated with lagged endogenous variables and $B^*_{i,l}(l = 0, 1, ..., p^*)$ is the matrix of parameters associated with contemporaneous and lagged weakly exogenous variables. The weakly exogenous foreign variables $y^*_{i,t}$ are constructed as a weighted average of the endogenous variables in other countries:

$$y^*_{i,t} = \sum_{j=0}^{N} w_{i,j} y_{j,t} \tag{7}$$

and the weights satisfy the following two restrictions: $w_{i,i} = 0$ and $\sum_{j=0}^{N} w_{i,j} = 1$. Weights are constructed from standardized bilateral trade flows. The data are available from Mohaddes and Raissi (2020).

In the second step, $N + 1$ country-specific VARX models are stacked to form a global model, which is given by

$$G y_t = c + \sum_{q=1}^{Q} H_q y_{t-q} + u_t, \tag{8}$$

where $y_t = (y'_{1,t}, ..., y'_{N,t})'$, $Q = \max(p, p^*)$, and G and H_q are both $NG \times NG$ dimensional coefficient matrices. Details on how to solve the global model can be found in Pesaran et al. (2009) and Huber (2016).

Priors for c_i and $B_{i,l}$ are specified in the same way as in the CFAVAR. More specifically, c_i and $B_{i,l}$ follow the same specification as in (4). For the prior on the elements of $B^*_{i,l}$, means are set to zero and variances are defined as: $\lambda_4 \frac{\sigma_m^2}{\sigma_n^2}$, where σ_m^2, σ_n^2 are obtained from univariate AR(1) regressions. We assume a Gamma prior for $\lambda_4 \sim \mathcal{G}(1, 0.02^2)$. Both p and p^* are set to 4.

B.4 Multi-country VARs

B.4.1 Factor shrinkage approach

The factor shrinkage approach used with the CC specification relies on the VAR written in system form. We define $X_t = I_{NG} \otimes x'_t$, where $x_t = (1, Y'_{t-1}, \ldots, Y'_{t-p})'$, β_i is the $k \times 1$ vector containing coefficients related to each i, $k = NGp + 1$, and $\beta = (\beta'_1, \ldots, \beta'_N)'$ is the $NGk \times 1$ vector containing all coefficients. Write the

VAR as

$$Y_t = X_t\beta + u_t, \tag{9}$$

where $u_t \sim i.i.d.\ N(0, \Sigma_t)$.

Canova and Ciccarelli (2009) assume that the vector of coefficients β can be expressed as:

$$\beta = \sum_{i=1}^{F} \Xi_i \theta_i \tag{10}$$

where $\Xi = [\Xi_1, \ldots, \Xi_F]$ are known matrices and $\theta = (\theta_1', \ldots, \theta_F')'$ is a low dimensional vector ($\dim(\theta) < K$, where $K = kNG$) of unknown parameters, and $\theta_1, \ldots, \theta_F$ are mutually orthogonal.[2]

We consider the factorization used in Canova et al. (2007) and Canova and Ciccarelli (2013). We assume $F = 4$. θ_1 is a scalar factor that is common across all countries, $\theta_2' = (\theta_{2,1}, \ldots, \theta_{2,N})'$ is an $N \times 1$ vector of country-specific factors, $\theta_3' = (\theta_{3,1}, \ldots, \theta_{3,G})'$ is a $G \times 1$ vector of variable-specific factors and $\theta_4' = (\theta_{4,1}, \ldots, \theta_{4,p-1})'$ is a $(p-1) \times 1$ vector of lag-specific factors.[3] $\Xi_1, \ldots, \Xi_4$ are assumed to be known with elements associated with the corresponding original parameters equal to 1 and 0 otherwise. For example, consider a multi-country VAR model in (1) with $N = 2, G = 2, p = 1$. In this case, Ξ_1 is a 20×1 vector of ones, and Ξ_2 and Ξ_3 take the form:

$$\underset{20\times 2}{\Xi_2} = \begin{bmatrix} \iota_1 & 0 \\ \iota_1 & 0 \\ 0 & \iota_2 \\ 0 & \iota_2 \end{bmatrix}, \quad \underset{20\times 2}{\Xi_3} = \begin{bmatrix} \iota_3 & 0 \\ 0 & \iota_4 \\ \iota_3 & 0 \\ 0 & \iota_4 \end{bmatrix},$$

where $\iota_1 = (0,1,1,0,0)'$, $\iota_2 = (0,0,0,1,1)'$, $\iota_3 = (0,1,0,1,0)'$, and $\iota_4 = (0,0,1,0,1)'$. Thus, we can rewrite (9) as:

$$\begin{aligned} Y_t &= X_t\beta + u_t \\ &= X_t(\Xi\theta) + u_t = \tilde{X}_t\theta + u_t. \end{aligned} \tag{11}$$

In this case, $\dim(\theta) = N + G + p$. By construction, the $\tilde{X}_t$'s are linear combinations of the original right-hand-side variables in (9), and the parameterization above can capture comovement across lagged variables.

To incorporate SV, we decompose Σ_t as $\Sigma_t = A^{-1}H_tA'^{-1}$, where A is lower diagonal with diagonal

[2] A more general form is $\beta = \sum_{i=1}^{F} \Xi_i\theta_i + e$, where $e \sim N(0, \Sigma \otimes \sigma^2 I)$ is an approximation error uncorrelated with u_t. However, most of the literature assumes an exact factorization ($\sigma^2 = 0$); see, for example, Canova et al. (2007); Canova and Ciccarelli (2009); Dées and Güntner (2017). Koop and Korobilis (2019) estimate σ^2 by a forgetting factor approach and find that it is very small (< 0.01). In some limited checks, we found that considering the approximation error harms forecasting performance.

[3] To avoid collinearity with θ_1, θ_4 can contain at most $p - 1$ components.

elements equal to 1, and the diagonal elements in H_t evolve according to (3).

We specify the priors for θ, A, and Φ as (independent), Normal, Normal, and Inverse Wishart, respectively:

$$\theta \sim N(0, \underline{\Omega}_\theta), \quad a \sim N(0, \underline{\Omega}_a), \quad \Phi \sim IW(Q_0, W_0), \tag{12}$$

where a denotes the vector of free elements in A. The prior mean for θ is set to zero, and the prior covariance matrix $\underline{\Omega}_\theta$ is assumed to be diagonal. Letting $\omega_{\theta_{i,j}}$ be the elements in $\underline{\Omega}_\theta$ associated with the jth elements in θ_i, where $i = 1, \dots, 4$, then

$$\underline{\omega}_{\theta_{i,j}} = \left\{ \begin{array}{ll} \sum_{m=1}^{NG} \sigma_m^2 & i = 1,2,3 \\ \frac{\sum_{m=1}^{NG} \sigma_m^2}{l^2}, & i = 4, l = 2, \dots, p \end{array} \right. .$$

The prior mean for a is set to 0, and the prior variance is set to $\underline{\Omega}_a = 10 \times I$. Q_0, W_0 are specified as $Q_0 = NG + 2$, $W_0 = 0.01 \times I$.

B.4.2 Prior specifications for other models

For the approach in Angelini et al. (2019) and the hierarchical shrinkage considered in this paper, the prior for free elements in A is assumed to be Normal with zero mean and variance equal to $10 \times I_{NG}$. The prior for Φ takes the form $\Phi \sim IW(Q_0, W_0)$, and Q_0, W_0 are specified as $Q_0 = NG + 2$, $W_0 = 0.01 \times I$.

For the prior in (4), σ_i^2, σ_j^2 are obtained from univariate AR(1) regressions. The prior for the intercept is assumed to be uninformative by setting the prior variance equal to $100 \times \sigma_i^2$, where σ_i^2 is again from a univariate AR(1) regression. The hyper-priors on overall shrinkage parameters are specified in the same way as in country-specific VARs. For the additional hyperparameter λ_4 controlling the tightness for coefficients related to cross-variable lags for foreign countries, we use a prior of $\lambda_4 \sim \mathcal{G}(1, 0.02^2)$.

For the SSSS prior, we follow Korobilis (2016) exactly. For (7) and (8) (in the main paper) , we set $\xi_{ij}^2 = \tau_{ij}^2 = 4$ and $\underline{c}^{\mathrm{DI}} = \underline{c}^{\mathrm{CSH}} = 0.0025$. The priors for indicators are specified as

$$\gamma_{ij}^{\mathrm{DI}} \sim Bernoulli(\pi_{ij}^{\mathrm{DI}}), \ \ \pi_{ij}^{\mathrm{DI}} \sim \mathcal{B}(1,1)$$
$$\gamma_{ij}^{\mathrm{CSH}} \sim Bernoulli(\pi_{ij}^{\mathrm{CSH}}), \ \ \pi_{ij}^{\mathrm{CSH}} \sim \mathcal{B}(1,1).$$

For the Horseshoe prior, no more prior specifications are needed. For the Normal-Gamma prior, recall that we specify $a^\omega \sim \mathcal{E}(b)$ and $\kappa^2 \sim \mathcal{G}(d_1, d_2)$. We set b equal to the number of coefficients in each block and elicit a non-informative prior for κ^2 by setting $d_1 = d_2 = 0.01$. For the Normal-Gamma-Gamma prior, recall that $2a \sim \mathcal{B}(\alpha_a, \beta_a)$, $2c \sim \mathcal{B}(\alpha_c, \beta_c)$, and we set $\alpha_a = \alpha_c = 2, \beta_a = \beta_c = 1$.

B.5 Univariate models

For AR(p)-SV models applied to each scalar output growth or interest rate variable, generally denoted y_t, we have

$$
\begin{aligned}
y_t &= c + \sum_{\ell=1}^{p} \rho_\ell y_{t-\ell} + u_t, \\
u_t &= h_t^{0.5} v_t, \quad v_t \overset{i.i.d.}{\sim} N(0,1), \\
\log h_t &= \log h_{t-1} + e_t, \quad e_t \overset{i.i.d.}{\sim} N(0, \sigma_e^2).
\end{aligned}
$$

As in Clark and Ravazzolo (2015), lag length is set to 2 for output growth and 4 for the interest rate. Letting $\theta = (c, \rho_1, \ldots, \rho_p)'$, we specify the following priors:

$$
\theta \sim N(0, V), \quad \sigma_e^2 \sim IG(v_h, S_h), \quad \log h_0 \sim N(a_0, b_0).
$$

V is assumed to be diagonal with elements equal to $\frac{\theta_1}{\ell^{\theta_2}}$, $\ell = 1, \ldots, p$, for autoregressive coefficients and $100 \times \hat{\sigma}_y^2$ for the intercept. θ_1 is set to 0.04, θ_2 is set to 2, and $\hat{\sigma}_y^2$ is obtained from a univariate AR(1) regression. We use a modestly informative prior for σ_e^2 to control the time variation by setting v_h equal to 2 and S_h to 0.04. For the prior on initial conditions, we set $a_0 = 0$ and $b_0 = 10$.

For the UCSV model, we have

$$
\begin{aligned}
y_t &= \tau_t + \varepsilon_t^y, \quad \varepsilon_t^y \sim N(0, e^{h_t}), \\
\tau_t &= \tau_{t-1} + \varepsilon_t^\tau, \quad \varepsilon_t^\tau \sim N(0, e^{g_t}), \\
h_t &= h_{t-1} + \varepsilon_t^h, \quad \varepsilon_t^h \sim N(0, \omega_h^2), \\
g_t &= g_{t-1} + \varepsilon_t^g, \quad \varepsilon_t^g \sim N(0, \omega_g^2),
\end{aligned}
$$

with initial conditions τ_0, h_0 and g_0 as unknown parameters. We can rewrite the above UCSV model in the non-centered parameterization:

$$
\begin{aligned}
y_t &= \tau_t + e^{\frac{1}{2}(h_0 + \omega_h \tilde{h}_t)} \tilde{\varepsilon}_t^y, \\
\tau_t &= \tau_{t-1} + e^{\frac{1}{2}(g_0 + \omega_g \tilde{g}_t)} \tilde{\varepsilon}_t^\tau, \\
\tilde{h}_t &= \tilde{h}_{t-1} + \tilde{\varepsilon}_t^h, \\
\tilde{g}_t &= \tilde{g}_{t-1} + \tilde{\varepsilon}_t^g,
\end{aligned}
$$

where $\tilde{h}_0 = \tilde{g}_0 = 0$ and $\tilde{\varepsilon}_t^y$, $\tilde{\varepsilon}_t^\tau$, $\tilde{\varepsilon}_t^h$, and $\tilde{\varepsilon}_t^g$ are all *i.i.d.* $N(0,1)$. We assume Normal priors for all model parameters: $\omega_h \sim N(0, 0.2^2)$, $\omega_g \sim N(0, 0.2^2)$, $h_0 \sim N(0, 10)$, $g_0 \sim N(0, 10)$, and $\tau_0 \sim N(0, 10)$.

C Algorithms

C.1 Algorithms for VARs with Minnesota-type prior

For all the country-specific VARs, country-specific factor-augmented VARs, global VARs, and multi-country VARs with Minnesota prior, the MCMC samplers follow almost exactly the steps in Carriero et al. (2022), but an additional step is needed to update prior tightness parameters. We highlight three issues related to the sampler, and refer interested readers to Appendix A.3 in their paper for other details.

Step 1: Update $\beta|\cdot$. We update the coefficients equation by equation, as in the corrected triangular algorithm in Carriero et al. (2022). Details can be found in Appendix C.5.

Step 2: Update $\lambda_i|\cdot$, $i = 1, 2, 4$. Let S_{λ_i}, $i = 1, 2, 4$, be the collection of all indexes such that parameters associated with the overall shrinkage parameters belong to this set. It can easily be shown that, with a Gamma prior, $\lambda_i \sim \mathcal{G}(1, c_i)$, conditional posteriors follow a Generalized Inverse Gaussian distribution:

$$\lambda_i|\cdot \sim \mathcal{GIG}\Big(1 - \frac{\dim(S_{\lambda_i})}{2}, 2c_i, \sum_{(i,j)\in S_{\lambda_i}} \frac{\beta_{i,j}^2}{2C_{i,j}}\Big).$$

The density of $x \sim \mathcal{GIG}(p, a, b)$ is given by $f(x) \propto x^{p-1} \exp\big(-(ax + b/x)/2\big)$. dim denotes the dimension of the set, and $C_{i,j}$ are the prior local variance parameters (the elements in (4) without an overall shrinkage parameter).

Step 3: Update the volatility. For the volatility estimation, let $\tilde{u}_t = Au_t$ denote the rescaled residuals. The elements of $\tilde{u}_t$ obey the following process:

$$\ln \tilde{u}_{ij,t}^2 = \ln h_{ij,t} + \ln \epsilon_{ij,t}^2, \quad i = 1, \ldots, N, j = 1, \ldots, G.$$

So, together with state equation (2), we have a non-linear and non-Gaussian state space system. To get the volatility estimates, we use the KSC algorithm, first introduced in Kim et al. (1998) and detailed for VAR models in Del Negro and Primiceri (2015). We use a 10-state mixture of Normals to approximate the distribution of non-Gaussian errors $\ln \epsilon_{ij,t}^2$. The details of approximation are provided in Table 1 of Omori et al. (2007).

Step 4: Update the free elements in A. This can be done with the equation-by-equation approach of Cogley and Sargent (2005) or with the joint approach of Chan (2017). For the latter, letting a denote the free elements in A, it can be shown that a can be interpreted as the coefficients from the regression:

$$u_t = K_t a + e_t, e_t \sim N(0, D_t),$$

where $D_t = diag(h_{1,t}, \ldots, h_{NG,t})$, and K_t is given as

$$K_t = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & 0 \\ -u_{1t} & 0 & 0 & 0 & 0 & \cdots & \cdots & \vdots \\ 0 & -u_{1t} & -u_{2t} & 0 & 0 & \cdots & \cdots & \vdots \\ \vdots & & & \ddots & \ddots & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & -u_{1t} & \cdots & -u_{(NG-1)t} \end{bmatrix}.$$

This permits drawing a jointly. Given the prior $a \sim N(0, \underline{\Omega}_a)$, the posterior is also Gaussian $a|\beta, h, \Phi, Y \sim N(\overline{\mu}_a, \overline{\Omega}_a)$, where

$$\overline{\Omega}_a = (\underline{\Omega}_a^{-1} + K'H^{-1}K)^{-1}$$
$$\overline{\mu}_a = \overline{\Omega}_a K'H^{-1}u.$$

This algorithm can be more efficient than the equation-by-equation approach, because a is updated jointly. However, the band matrix K_t does not have a fixed bandwidth (the number of non-zeros elements increases with model size). Thus, letting n denote the number of variables in the model, the complexity of this algorithm is still $O(n^3)$, and the estimation quickly becomes computationally demanding as the model size increases. Accordingly, for country-specific models, which are small ($n = N = 3$), we use this algorithm to update a. But for multi-country models, which are large ($n = NG = 21$), we use the algorithm of Cogley and Sargent (2005) to draw a equation by equation.

Step 5: Update $\Phi|\cdot$. Since we elicit a conditionally conjugate prior, the conditional posterior takes the same form, which can be shown to be:

$$\Phi|\cdot \sim IW\Big(Q_0 + T, W_0 + \sum_{t=1}^{T} \big(\log(h_t) - \log(h_{t-1})\big)\big(\log(h_t) - \log(h_{t-1})\big)'\Big).$$

C.2 Algorithm for multi-country VAR with factor shrinkage

Most of the steps of the algorithm for the CC specification follow from the previous section, except that we have to adapt step 1's treatment of the VAR's coefficients. With the transformation, we see that given $\theta \sim N(0, \underline{\Omega}_\theta)$, the conditional posterior $\theta|Y, a, h, \Phi$ is multivariate Normal, $N(\overline{\mu}_\theta, \overline{\Omega}_\theta)$, with moments:

$$\overline{\Omega}_\theta = (\underline{\Omega}_\theta^{-1} + \tilde{Z}'\tilde{\Sigma}^{-1}\tilde{Z})^{-1}$$
$$\overline{\mu}_\theta = \overline{\Omega}_\theta \tilde{Z}'\tilde{\Sigma}^{-1}Y,$$

where $Y, \tilde{Z}$ are stacked versions of $Y_t, \tilde{Z}_t$ and $\tilde{\Sigma} = diag(\Sigma_1, \ldots, \Sigma_T)$.

C.3 Algorithms for multi-country VARs with hierarchical shrinkage

As in Algorithm 1 in the main text, the MCMC estimation involves 5 steps. The only new step compared to above is to update the prior variance parameters and associated hyperparameters. We provide details of the conditional posterior distributions for these parameters. In Section 6.4, we also estimate country-specific VAR-SV specifications with hierarchical shrinkage and Horseshoe prior. The algorithm follows exactly the ones described below.

First, consider the Horseshoe prior:

$$\beta_j|\omega_j^2 \sim N(0, \omega_j^2), \;\; \omega_j^2|\gamma_j^2 \sim \mathcal{G}(\frac{1}{2}, \gamma_j^2), \;\; \gamma_j^2 \sim \mathcal{G}(\frac{1}{2}, \lambda),$$

and $\lambda \sim C^{+}(0, 1)$. It follows from straightforward calculation that

$$\omega_j^2|\cdot \sim \mathcal{GIG}(0, 2\gamma_j^2, \beta_j^2),$$

where $\mathcal{GIG}(p, a, b)$ denotes the Generalized Inverse Gaussian distribution with pdf given by $f(x) \propto x^{p-1}\exp\big(-(ax + b/x)/2\big)$. For the conditional posterior of $\gamma_j^2|\cdot$, since the Gamma distribution is conjugate for the Gamma likelihood function, we have that

$$\gamma_j^2|\cdot \sim \mathcal{G}(1, \lambda + \omega_j^2).$$

Updating $\lambda|\cdot$ follows the same steps as above since the prior admits the hierarchical representation: $\lambda \sim \mathcal{G}(\frac{1}{2}, \xi^2)$, $\xi^2 \sim \mathcal{G}(\frac{1}{2}, 1)$.

Second, consider the Normal-Gamma prior:

$$\beta_j|\omega_j^2 \sim N(0, \omega_j^2), \;\; \omega_j^2 \sim \mathcal{G}(a^{\omega}, \frac{a^{\omega}\kappa^2}{2}),$$

and $a^{\omega} \sim \mathcal{E}(b)$ and $\kappa^2 \sim \mathcal{G}(d_1, d_2)$. It follows similarly as in the Horseshoe prior that

$$\omega_j^2|\cdot \sim \mathcal{GIG}(a^{\omega} - 0.5, a^{\omega}\kappa^2, \beta_j^2).$$

The conditional posterior for $a^{\omega}|\cdot$ is not available in closed form. We use adaptive Random Walk Metropolis-Hastings algorithms as in Roberts and Rosenthal (2009) with acceptance probability given by

$$\min\left\{1, \frac{p(a^{\omega,\text{new}})a^{\omega,\text{new}}}{p(a^{\omega})a^{\omega}} \prod_j \frac{p(\beta_j|a^{\omega,\text{new}}, \kappa^2)}{p(\beta_j|a^{\omega}, \kappa^2)}\right\},$$

where the marginal prior is given by

$$p(\beta_j|a^\omega, \kappa^2) = \frac{(\sqrt{a^\omega \kappa^2})^{a^\omega + \frac{1}{2}}}{\sqrt{\pi} 2^{a^\omega - \frac{1}{2}} \Gamma(a^\omega)} \left|\beta_j\right|^{a^\omega - \frac{1}{2}} K_{a^\omega - \frac{1}{2}}\left(\sqrt{a^\omega \kappa^2} \left|\beta_j\right|\right),$$

and $K(\cdot)$ denotes a modified Bessel function of the second kind. At each iteration i, a new value $a^{\omega,\text{new}}$ is proposed according to

$$\log a^{\omega,\text{new}} = \log a^\omega + \epsilon_j, \quad \epsilon_j \sim N(0, \sigma^{2(i)}_{\psi_j}). \tag{13}$$

The variance of the increments is fixed at 1 for the first 50 iterations, and then updated by

$$\log \sigma^{2(i+1)}_{a^\omega} = \log \sigma^{2(i)}_{a^\omega} + \frac{1}{i^q}(\hat{\alpha} - \alpha^*), \tag{14}$$

where $\hat{\alpha}$ is the estimated acceptance probability of current draws and α^* is the desired acceptance probability. The parameter q controls the degree of vanishing adaption, which is necessary to make the adaptive algorithm valid.[4] This algorithm leads to an average acceptance rate that converges to α^*. Following Griffin and Brown (2017), we set $q = 0.55, \alpha^* = 0.3$. Then updating $\kappa^2|\cdot$ is quite straightforward since it again follows a Gamma distribution:

$$\kappa^2|\cdot \sim \mathcal{G}\Big(Ma^\omega + d_1, d_2 + a^\omega \sum_j \omega_j^2\Big),$$

where M denotes the number of parameters in each block.

Finally, consider the Normal-Gamma-Gamma prior:

$$\beta_j|\tau_j^2, \lambda_j^2 \sim N\left(0, \phi \frac{\tau_j^2}{\lambda_j^2}\right), \;\; \tau_j^2 \sim \mathcal{G}(a, 1), \;\; \lambda_j^2 \sim \mathcal{G}(c, 1),$$

where $\phi = 2c/(a\kappa^2)$, $2a \sim \mathcal{B}(\alpha_a, \beta_a)$, $2c \sim \mathcal{B}(\alpha_c, \beta_c)$, and $\kappa^2|a, c \sim F(2a, 2c)$. We proceed as in Cadonna et al. (2020). As we use marginalized distributions in each step to improve sampling efficiency, the steps described below are not interchangeable.

Step a: Update $a|\cdot$. Use the prior $p(\beta_j|\lambda_j^2, a, c)$, marginalized *w.r.t.* τ_j^2, to draw $a|\cdot$ via an adaptive Random Walk Metropolis-Hastings algorithm on $z = \log(a/(0.5 - a))$. The variance of the increments is updated as in the Normal-Gamma case. At each iteration m, letting a^* be the candidate draw and $a^{(m-1)}$ be

[4]This means that the variances of increments are fixed as $i \to \infty$. Two conditions are provided in equations (1.1) and (1.2) of Roberts and Rosenthal (2009). The condition in equation (1.2) in their paper is generally satisfied provided that ψ is bounded above.

the previous draw, the acceptance probability is given by

$$\min\left\{1, \frac{q_a(a^*)}{q_a(a^{(m-1)})}\right\}, \quad q_a(a) = p(a|\cdot)a(0.5-a).$$

Letting m be the number of parameters in each block, $\log q_a(a)$ is given by

$$\begin{aligned}\log q_a(a) = a\Big(&-m\log 2 + \frac{m}{2}\log\kappa^2 - \frac{m}{2}\log c + \frac{1}{2}\sum_j \log\lambda_j^2 + \frac{1}{2}\sum_j \log\beta_j^2\Big)\\
&+ \frac{5}{4}m\log a + m\frac{a}{2}\log a - m\log\Gamma(a+1)\\
&+ \sum_j \log K_{a-\frac{1}{2}}\Big(\beta_j\sqrt{\lambda_j^2\kappa^2 a/c}\Big)\\
&- \log\mathcal{B}(a,c) + a\Big(\log a + \log\big(\frac{\kappa^2}{2c}\big)\Big) - \log a - (a+c)\log\Big(1 + \frac{a\kappa^2}{2c}\Big)\\
&+ (\alpha_a - 1)\log(2a) - (\beta_a - 1)\log(1-2a)\\
&+ \log a + \log(0.5-a).\end{aligned}$$

<u>*Step b:*</u> Update $\tau_j^2|\cdot$. This step is simple, as the conditional posterior is again $\mathcal{GIG}$:

$$\tau_j^2|\cdot \sim \mathcal{GIG}\Big(a - \frac{1}{2}, 2, \frac{\lambda_j^2\beta_j^2}{\phi}\Big).$$

<u>*Step c:*</u> Update $c|\cdot$. Use the prior $p(\beta_j|\tau_j^2, a, c)$, marginalized *w.r.t.* λ_j^2, to draw $c|\cdot$ via an adaptive Random Walk Metropolis-Hastings algorithm on $z = \log(c/(0.5-c))$. The variance of the increments is updated as in the Normal-Gamma case. At each iteration m, letting c^* be the candidate draw and $c^{(m-1)}$ be the previous draw, the acceptance probability is given by

$$\min\left\{1, \frac{q_c(c^*)}{q_c(c^{(m-1)})}\right\}, \quad q_c(c) = p(c|\cdot)c(0.5-c).$$

Letting m be the number of parameters in each block, $\log q_c(c)$ is given by

$$\begin{aligned}\log q_c(c) &= m\log\Gamma(c+0.5) - m\log\Gamma(c+1) + \frac{m}{2}\log c\\ &- (c+0.5)\Big(\sum_j \log(4c\tau_j^2 + \beta_j^2\kappa^2 a) - \sum_j \log(4c\tau_j^2)\Big)\\ &- \log\mathcal{B}(a,c) - (a-1)\log c - (a+c)\log(1+\frac{a\kappa^2}{2c})\\ &+ (\alpha_c - 1)\log(2c) + (\beta_c - 1)\log(1-2c)\\ &+ \log c + \log(0.5-c).\end{aligned}$$

Step d: Update $\lambda_j^2|\cdot$. This step is simple; the conditional posterior is $\mathcal{G}$:

$$\lambda_j^2|\cdot \sim \mathcal{G}\Big(\frac{1}{2} + c, \frac{\beta_j^2}{2\phi\tau_j^2} + 1\Big).$$

Step e: Update $\kappa^2|\cdot$. Notice that the prior of κ^2 admits the following hierarchical representation: $\kappa^2|a \sim \mathcal{G}(a, d_2)$, $d_2|a, c \sim \mathcal{G}(c, \frac{2c}{a})$. Then updating $\kappa^2|\cdot$ involves first sampling from

$$d_2|\cdot \sim \mathcal{G}(a + c, \kappa^2 + \frac{2c}{a}),$$

then sampling from (m is the number of parameters in each block)

$$\kappa^2|\cdot \sim \mathcal{G}(\frac{m}{2} + a, \frac{a}{4c}\sum_j \frac{\lambda_j^2}{\tau_j^2}\beta_j^2 + d_2).$$

C.4 Corrected triangular algorithm

Consider an n-variable reduced-form $VAR(p)$ model as in Carriero et al. (2022):

$$y_t = \Pi' x_t + A^{-1}\Lambda_t^{0.5}\epsilon_t, \quad \epsilon_t \overset{i.i.d.}{\sim} N(0, I_n),$$

where $t = 1, \ldots, T$, x_t is an $(np+1)\times 1$ dimensional vector containing the lags of y_t and an intercept, $\Pi = (\Pi_0, \Pi_1, \ldots, \Pi_p)'$ is an $(np+1)\times n$ matrix of coefficients, A^{-1} is a unit lower triangular matrix, and $\Lambda_t^{0.5}$ is diagonal with the log of the generic j-th element following a random walk process.

Defining $\tilde{y}_t = Ay_t$ with generic j-th element $\tilde{y}_{j,t} = y_{j,t} + a_{j,1}y_{1,t} + \cdots + a_{j,j-1}y_{j-1,t}$, consider the triangular representation of the system:

$$\tilde{y}_t = A\Pi' x_t + \Lambda_t^{0.5}\epsilon_t = A(x_t'\Pi)' + \Lambda_t^{0.5}\epsilon_t,$$

which can be expressed as the following system of equations:

$$\begin{aligned}
\tilde{y}_{1,t} &= x_t'\pi^{(1)} + \lambda_{1,t}^{0.5}\epsilon_{1,t} \\
\tilde{y}_{2,t} &= a_{2,1}x_t'\pi^{(1)} + x_t'\pi^{(2)} + \lambda_{2,t}^{0.5}\epsilon_{2,t} \\
\tilde{y}_{3,t} &= a_{3,1}x_t'\pi^{(1)} + a_{3,2}x_t'\pi^{(2)} + x_t'\pi^{(3)} + \lambda_{3,t}^{0.5}\epsilon_{3,t} \\
&\vdots \\
\tilde{y}_{n,t} &= a_{n,1}x_t'\pi^{(1)} + \cdots + a_{n,n-1}x_t'\pi^{(n-1)} + x_t'\pi^{(n)} + \lambda_{n,t}^{0.5}\epsilon_{n,t},
\end{aligned}$$

where $\pi^{(j)}$ denotes the coefficients of the j-th equation. Clearly, $\pi^{(j)}$ appears not only in equation j but also in equations $j+1$ through n. Letting $z_{j+l,t} = \tilde{y}_{j+l,t} - \sum_{i\neq j,i=1}^{j+l} a_{j+l,i}x_t'\pi^{(i)}$, for $l = 0, ..., n-j$, and $a_{i,i} = 1$, consider the following system of equations:

$$\begin{aligned}
z_{j,t} &= x_t'\pi^{(j)} + \lambda_{j,t}^{0.5}\epsilon_{j,t} \\
z_{j+1,t} &= a_{j+1,j}x_t'\pi^{(j)} + \lambda_{j+1,t}^{0.5}\epsilon_{j+1,t} \\
&\vdots \\
z_{n,t} &= a_{n,j}x_t'\pi^{(j)} + \lambda_{n,t}^{0.5}\epsilon_{n,t}.
\end{aligned}$$

Then, using the above triangular representation, the full conditional posterior of $\pi^{(j)}|\cdot$ follows immediately from standard Bayesian linear regression results (assuming that prior means are zero):

$$\pi^{(j)}|\cdot \sim N\left(\overline{\mu}_{\pi^{(j)}}, \overline{\Omega}_{\pi^{(j)}}\right),$$

where

$$\begin{aligned}
\overline{\Omega}_{\pi^{(j)}}^{-1} &= \underline{\Omega}_{\pi^{(j)}}^{-1} + \sum_{i=j}^{n} a_{i,j}^2 \sum_{t=1}^{T} \frac{1}{\lambda_{i,t}} x_t x_t' \\
\overline{\mu}_{\pi^{(j)}} &= \overline{\Omega}_{\pi^{(j)}} \times \Big(\sum_{i=j}^{n} a_{i,j} \sum_{t=1}^{T} \frac{1}{\lambda_{i,t}} x_t z_{i,t} \Big),
\end{aligned}$$

with $a_{i,i} = 1$.

C.5 Algorithms for SSSS prior

The algorithms described in Appendix A.3 of Korobilis (2016) can be easily extended to our case with SV. Only step 1 has to be modified. In particular, let $Y = (y_1 \ \cdots \ y_T)'$, $x_t = (1, y_{t-1}')'$, and $X = (x_1 \ \cdots \ x_T)'$, and write the model as

$$Y = XB + U,$$

where $U = (u_1 \ \cdots \ u_T)'$. The sampler involves the following steps:

Step a: Update $\text{vec}(B)|\cdot$. It can be shown that

$$\text{vec}(B)|\cdot \sim N(\Gamma \times \mu_B, D_B),$$

where

$$D_B = \Big(V + \sum_{t=1}^{T} (\Sigma_t^{-1} \otimes x_t' x_t)\Big)^{-1}, \ \ \mu_B = D_B\Big(\text{vec}(\sum_{t=1}^{T} x_t y_t' \Sigma_t^{-1})\Big).$$

The diagonal matrix V contains prior variances; details of constructing the indicator matrix Γ can be found in Korobilis (2016).

Steps b,c,d,e: These follow exactly as in steps 2,3,4,5 in Korobilis (2016).

Steps f,g,h: Update free elements in A, stochastic volatility, and related parameters. These steps follow the corresponding steps used for the multi-country VAR with the Minnesota-type prior.

C.6 Algorithms for country-specific VAR with hierarchical shrinkage

We follow exactly the algorithms described in Chan (2021). Estimation for the intercept, autoregressive coefficients, free elements in A, and stochastic volatility is very similar to the algorithms used in this paper. It is worth mentioning that, as in Chan (2021), the model has been first transformed to structural form, and then estimation is performed equation by equation. For hyperparameters related to the Normal-Gamma prior, since a slightly different parameterization is used there, the updating of hyperparameters is slightly different. The conditional posterior for $\psi_{i,j}|\cdot$ is also $\mathcal{GIG}$, but with a slightly different parameterization. An independent Metropolis-Hastings algorithm is used to update $\nu_\psi|\cdot$. We refer the reader to Section 4 and Appendix B in that paper for more details.

C.7 Algorithms for univariate models

We use the algorithms as described in Clark and Ravazzolo (2015) to estimate AR(p)-SV models. The steps to draw intercept and autoregressive parameters follow from standard linear regression results. To estimate stochastic volatility and related parameters, we follow the procedures described in Section 7.1 in Chan (2017). For the UCSV model, we estimate it in non-centered parameterization and then transform back to the centered parameterization to perform predictive simulation. Estimation details can be found in Appendix B in Chan (2018) and in Section 7.2 in Chan (2017).

D Additional empirical results

Table 1: Comparison of HS-CSH and baseline HS: descriptive statistics for all horizons

	All horizons		$h \leqslant 6$		$h > 6$			All horizons		$h \leqslant 6$		$h > 6$	
Output growth	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS	**Inflation**	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS
Mean	-1.229	-1.999	-0.623	-1.236	-1.836	-2.761	Mean	-0.204	0.087	-1.018	-0.759	0.610	0.933
Median	-1.053	-2.029	-0.610	-1.183	-2.209	-3.580	Median	0.325	0.264	0.287	0.077	0.349	0.639
Min	-6.251	-7.402	-3.926	-5.022	-6.251	-7.402	Min	-8.772	-7.086	-8.772	-7.086	-6.593	-5.164
Max	3.282	3.078	2.128	2.284	3.282	3.078	Max	10.556	10.025	3.519	4.223	10.556	10.025
% > 0	32.143	28.571	35.714	28.571	28.571	28.571	% > 0	57.143	55.952	54.762	52.381	59.524	59.524
$\%p <= 0.05$	0	2.381	0	2.381	0	2.381	$\%p <= 0.05$	0	1.190	0	2.381	0	0
Interest rate	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS							
Mean	7.951	6.192	3.960	2.191	11.942	10.193							
Median	6.169	5.192	3.180	1.804	10.788	8.174							
Min	-4.416	-5.618	-4.416	-5.618	0.337	0.383							
Max	28.668	23.514	21.308	14.492	28.668	23.514							
% > 0	90.476	86.905	80.952	73.810	100	100							
$\%p <= 0.05$	8.333	10.714	2.381	2.381	14.286	19.048							

Notes: ”HS-CSH” is the multi-country VAR model in which all the parameters related to CSH restrictions follow the same Horseshoe prior specification. The table provides summary statistics for the performance of this alternative model compared to the multi-country HS specification. Descriptive statistics include average, median, minimum, maximum, percentage of cases in which gains are above 0, and the percentage gains in which the forecasts from the competing models are statistically different according to the Diebold-Mariano (1995) test with fixed-smoothing asymptotics as in Coroneo and Iacone (2020).

Table 2: Comparison of HS-A and baseline HS: descriptive statistics for all horizons

	All horizons		$h \leqslant 6$		$h > 6$			All horizons		$h \leqslant 6$		$h > 6$	
Output growth	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS	**Inflation**	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS
Mean	-1.592	-3.194	-1.386	-2.642	-1.798	-3.746	Mean	-2.180	-1.956	-2.282	-2.190	-2.079	-1.721
Median	-1.626	-3.109	-1.160	-2.488	-2.434	-4.376	Median	-2.037	-1.909	-1.173	-1.290	-2.533	-3.450
Min	-7.806	-12.719	-7.806	-10.600	-7.287	-12.719	Min	-12.187	-10.419	-12.187	-10.419	-9.568	-8.276
Max	4.116	3.688	2.804	3.197	4.116	3.688	Max	9.258	10.280	3.404	4.276	9.258	10.280
% > 0	33.333	28.571	33.333	28.571	33.333	28.571	% > 0	28.571	25	28.571	21.429	28.571	28.571
%p <= 0.05	1.190	4.762	2.381	4.762	0	4.762	%p <= 0.05	0	0	0	0	0	0
Interest rate	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS							
Mean	10.697	10.033	4.370	3.699	17.023	16.367							
Median	9.769	10.128	4.496	3.532	17.011	16.445							
Min	-7.071	-10.052	-7.071	-10.052	6.479	7.464							
Max	25.761	26.416	17.567	15.201	25.761	26.416							
% > 0	90.476	86.905	80.952	73.810	100	100							
%p <= 0.05	21.429	38.095	9.524	23.810	33.333	52.381							

Notes: "HS-A" is the multi-country VAR model in which all the parameters follow the same Horseshoe prior specification. The table provides summary statistics for the performance of this alternative model compared to the multi-country HS specification. Descriptive statistics include average, median, minimum, maximum, percentage of cases in which gains are above 0, and the percentage gains in which the forecasts from the competing models are statistically different according to the Diebold-Mariano (1995) test with fixed-smoothing asymptotics as in Coroneo and Iacone (2020).

Table 3: Comparison of HS-E and baseline HS: descriptive statistics for all horizons

	All horizons		$h \leqslant 6$		$h > 6$			All horizons		$h \leqslant 6$		$h > 6$	
Output growth	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS	**Inflation**	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS
Mean	-1.154	-2.286	-1.212	-2.101	-1.095	-2.470	Mean	-1.756	-2.309	-2.332	-2.680	-1.181	-1.939
Median	-1.454	-2.461	-1.105	-1.768	-1.608	-2.767	Median	-1.109	-2.087	-1.109	-2.370	-1.036	-1.991
Min	-7.244	-10.132	-7.244	-9.228	-6.265	-10.132	Min	-13.790	-15.398	-13.790	-15.398	-9.156	-10.805
Max	4.161	3.987	3.126	3.642	4.161	3.987	Max	11.063	9.459	3.316	4.634	11.063	9.459
% > 0	33.333	28.571	38.095	28.571	28.571	28.571	% > 0	26.190	26.190	28.571	30.952	23.810	21.429
%p <= 0.05	0	1.190	0	2.381	0	0	%p <= 0.05	0	3.571	0	7.143	0	0
Interest rate	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS							
Mean	11.280	10.700	4.059	3.313	18.501	18.087							
Median	10.568	10.984	2.596	2.770	16.977	17.372							
Min	-7.634	-11.672	-7.634	-11.672	6.737	7.870							
Max	34.537	28.239	25.968	19.593	34.537	28.239							
% > 0	83.333	82.143	66.667	64.286	100	100							
%p <= 0.05	13.095	20.238	2.381	9.524	23.810	30.952							

Notes: "HS-E" is the multi-country VAR model in which all the parameters in each equation follow the same Horseshoe prior specification. The table provides summary statistics for the performance of this alternative model compared to the multi-country HS specification. Descriptive statistics include average, median, minimum, maximum, percentage of cases in which gains are above 0, and the percentage gains in which the forecasts from the competing models are statistically different according to the Diebold-Mariano (1995) test with fixed-smoothing asymptotics as in Coroneo and Iacone (2020).

Table 4: Comparison of Horseshoe priors with and without SV: descriptive statistics for all horizons

	All horizons		$h \leqslant 6$		$h > 6$			All horizons		$h \leqslant 6$		$h > 6$	
Output growth	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS	**Inflation**	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS
Mean	-1.852	-3.445	-0.893	-2.525	-2.811	-4.364	Mean	-22.974	-19.720	-11.820	-12.712	-34.127	-26.727
Median	-1.962	-3.434	-1.123	-2.603	-2.722	-4.309	Median	-16.394	-15.737	-7.538	-8.186	-24.002	-19.492
Min	-9.097	-11.549	-6.032	-8.837	-9.097	-11.549	Min	-98.999	-69.770	-56.606	-52.922	-98.999	-69.770
Max	5.197	4.580	5.197	4.580	2.362	0.835	Max	4.018	0.548	4.018	0.548	-7.184	-7.248
$\%>0$	28.571	19.048	38.095	28.571	19.048	9.524	$\%>0$	4.762	1.190	9.524	2.381	0	0
$\%p <= 0.05$	3.571	5.952	0	2.381	7.143	9.524	$\%p <= 0.05$	15.476	15.476	16.667	16.667	14.286	14.286
Interest rate	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS							
Mean	-9.429	-13.515	-12.149	-20.332	-6.709	-6.698							
Median	-7.592	-6.062	-8.375	-14.525	-7.590	-2.534							
Min	-57.608	-73.394	-57.608	-73.394	-27.100	-47.395							
Max	10.402	14.686	9.970	11.987	10.402	14.686							
$\%>0$	26.190	28.571	21.429	14.286	30.952	42.857							
$\%p <= 0.05$	11.905	28.571	16.667	42.857	7.143	14.286							

Notes: The table provides summary statistics for the performance of the alternative model with SV compared to the multi-country HS specification with SV. Descriptive statistics include average, median, minimum, maximum, percentage of cases in which gains are above 0, and the percentage gains in which the forecasts from the competing models are statistically different according to the Diebold-Mariano (1995) test with fixed-smoothing asymptotics as in Coroneo and Iacone (2020).

Table 5: Comparison of Horseshoe priors with expanding versus rolling windows: descriptive statistics for all horizons

	All horizons		$h \leqslant 6$		$h > 6$			All horizons		$h \leqslant 6$		$h > 6$	
Output growth	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS	**Inflation**	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS
Mean	1.095	0.342	1.229	0.707	0.960	-0.023	Mean	3.268	1.872	3.989	2.046	2.546	1.698
Median	0.849	0.160	1.102	0.528	0.759	0.002	Median	3.006	1.788	3.533	2.128	2.742	1.208
Min	-1.632	-1.930	-1.632	-1.930	-0.688	-1.772	Min	-1.479	-3.063	-1.318	-2.326	-1.479	-3.063
Max	7.256	4.464	7.256	4.464	3.948	3.023	Max	10.905	5.850	10.905	5.495	7.575	5.850
$\%>0$	83.333	54.762	80.952	59.524	85.714	50	$\%>0$	92.857	85.714	95.238	88.095	90.476	83.333
$\%p <= 0.05$	3.571	4.762	7.143	9.524	0	0	$\%p <= 0.05$	0	0	0	0	0	0
Interest rate	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS							
Mean	1.928	1.330	0.539	-0.077	3.317	2.738							
Median	1.435	0.461	-0.338	0.043	3.237	1.713							
Min	-6.350	-6.095	-6.350	-6.095	-2.769	-4.450							
Max	12.168	12.503	9.226	9.878	12.168	12.503							
$\%>0$	59.524	57.143	47.619	50	71.429	64.286							
$\%p <= 0.05$	10.714	4.762	7.143	0	14.286	9.524							

Notes: The table provides summary statistics for the performance of the HS model estimated with a rolling approach relative to the paper's baseline recursive approach. Descriptive statistics include average, median, minimum, maximum, percentage of cases in which gains are above 0, and the percentage gains in which the forecasts from the competing models are statistically different according to the Diebold-Mariano (1995) test with fixed-smoothing asymptotics as in Coroneo and Iacone (2020).

Table 6: Comparison with univariate models with HS baseline featuring SV: descriptive statistics for all horizons

	All horizons		$h \leqslant 6$		$h > 6$			All horizons		$h \leqslant 6$		$h > 6$	
Output growth	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS	**Inflation**	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS
Mean	0.535	0.732	0.290	0.396	0.779	1.067	Mean	1.279	3.874	0.252	0.927	2.306	6.822
Median	-0.513	-0.461	-0.342	-0.406	-0.750	-0.744	Median	2.022	3.073	1.740	1.213	2.572	7.117
Min	-3.642	-4.655	-3.616	-3.813	-3.642	-4.655	Min	-13.484	-7.475	-13.484	-7.475	-11.329	-3.670
Max	7.571	10.119	5.646	6.972	7.571	10.119	Max	13.247	17.693	6.875	7.681	13.247	17.693
%> 0	38.095	40.476	40.476	42.857	35.714	38.095	%> 0	70.238	77.381	59.524	64.286	80.952	90.476
$\%p <= 0.05$	13.095	14.286	11.905	7.143	14.286	21.429	$\%p <= 0.05$	0	8.333	0	0	0	16.667
Interest rate	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS							
Mean	8.982	7.799	8.442	7.370	9.522	8.227							
Median	10.504	10.207	8.651	8.779	14.056	16.663							
Min	-19.264	-28.167	-15.896	-22.757	-19.264	-28.167							
Max	28.698	29.027	24.232	25.927	28.698	29.027							
%> 0	78.571	63.095	85.714	69.048	71.429	57.143							
$\%p <= 0.05$	11.905	25	9.524	23.810	14.286	26.190							

Notes: This table presents descriptive statistics on comparisons of forecasting performance for the multi-country VAR-SV model with the Horseshoe prior (the paper's HS specification) relative to univariate models with SV. For output growth and the interest rate, we use an AR(p)-SV model, with $p = 2$ for output growth and $p = 4$ for the interest rate. For inflation, we use an unobserved component model with SV, as in Chan (2018). Descriptive statistics include average, median, minimum, maximum, percentage of cases in which gains are above 0, and the percentage gains in which the forecasts from the competing models are statistically different according to the Diebold-Mariano (1995) test with fixed-smoothing asymptotics as in Coroneo and Iacone (2020).

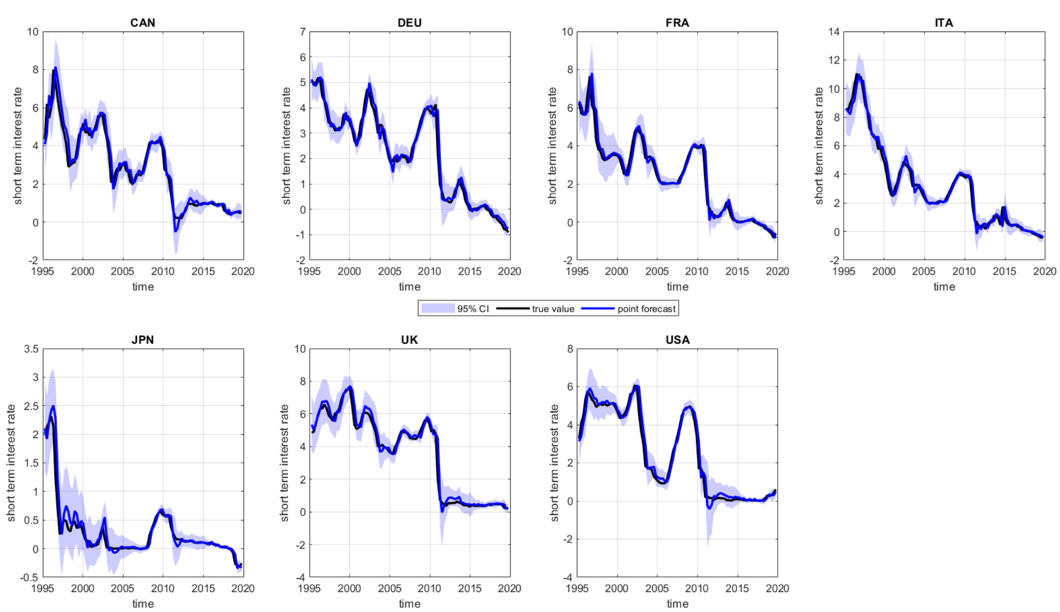

Figure 1: The figures present 1-step-ahead short-term interest rate forecasts for all G7 countries. The blue line and shaded areas are point forecasts and associated 95 percent forecast intervals. The black line shows the true values.

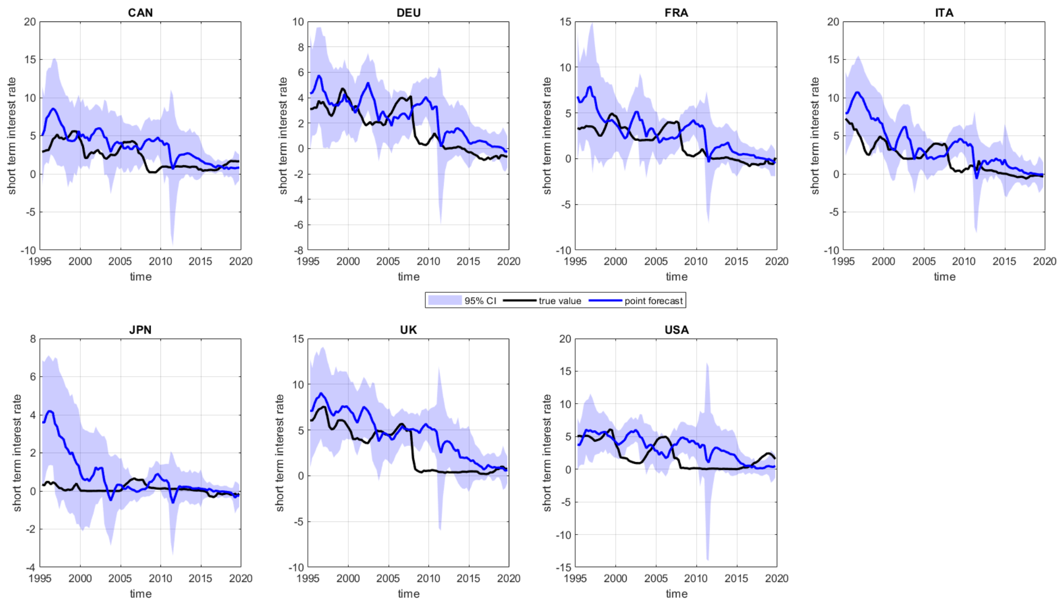

Figure 2: The figures present 12-steps-ahead short-term interest rate forecasts for all G7 countries. The blue line and shaded areas are point forecasts and associated 95 percent forecast intervals. The black line shows the true values.

Table 7: Directional forecast: 1-step-ahead changes in output growth

	CAN	DEU	FRA	ITA	JPN	UK	USA
HS	-0.380	6.715	2.508	5.870	7.556	8.174	9.225
	(0.648)	(0.000)	(0.006)	(0.000)	(0.000)	(0.000)	(0.000)
CVAR	-2.183	0.421	-2.928	-2.307	-2.318	-1.668	-3.797
	(0.986)	(0.337)	(0.998)	(0.990)	(0.990)	(0.952)	(0.999)

Notes: This table presents test statistics and associated p-values for directional predictive performance of 1-step-ahead changes in output growth from multi-country VAR-SV model with Horseshoe prior and single-country VAR-SV benchmark. The test statistics are computed according to equation (6) in Pesaran and Timmermann (1992).

Table 8: Comparison with hierarchical country-specific VARs featuring SV: HS prior versus Chan (2021), descriptive statistics for all horizons

	All horizons		$h \leqslant 6$		$h > 6$			All horizons		$h \leqslant 6$		$h > 6$	
Output growth	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS	**Inflation**	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS
Mean	1.778	2.238	1.525	1.924	2.032	2.552	Mean	2.944	3.462	1.520	2.027	4.369	4.896
Median	1.947	1.876	1.412	1.737	1.951	2.556	Median	2.174	2.881	0.924	2.124	4.364	4.747
Min	-3.041	-3.547	-3.041	-3.547	-1.975	-0.621	Min	-2.281	-1.249	-2.281	-1.249	-1.583	-0.724
Max	6.223	6.509	6.223	6.509	5.940	6.200	Max	9.802	10.208	6.099	6.701	9.802	10.208
%> 0	79.762	84.524	83.333	85.714	76.190	83.333	%> 0	80.952	86.905	76.190	83.333	85.714	90.476
$\%p <= 0.05$	17.857	14.286	11.905	11.905	23.810	16.667	$\%p <= 0.05$	1.190	22.619	2.381	11.905	0	33.333
Interest rate	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS							
Mean	9.936	9.228	8.174	7.606	11.699	10.851							
Median	11.195	10.659	8.529	8.511	14.053	13.219							
Min	-0.939	-0.742	-0.939	-0.742	1.082	-0.365							
Max	20.896	21.690	20.896	17.718	20.696	21.690							
%> 0	97.619	90.476	95.238	95.238	100	85.714							
$\%p <= 0.05$	30.952	30.952	11.905	14.286	50	47.619							

Notes: This table presents descriptive statistics on comparisons of forecasting performance for the hierarchical shrinkage in the country-specific VAR-SV model with the Horseshoe prior (the paper's HS specification) relative to the hierarchical shrinkage with the Normal-Gamma prior as in Chan (2021). Descriptive statistics include average, median, minimum, maximum, percentage of cases in which gains are above 0, and the percentage gains in which the forecasts from the competing models are statistically different according to the Diebold-Mariano (1995) test with fixed-smoothing asymptotics as in Coroneo and Iacone (2020).

Table 9: Comparison of country-specific VAR-SV and baseline multi-country VAR-SV with HS: descriptive statistics for all horizons

	All horizons		$h \leqslant 6$		$h > 6$			All horizons		$h \leqslant 6$		$h > 6$	
Output growth	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS	**Inflation**	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS
Mean	1.134	1.341	1.537	1.758	0.731	0.924	Mean	1.429	2.286	0.825	1.189	2.034	3.382
Median	0.775	1.152	1.064	1.201	0.617	0.968	Median	1.964	2.600	1.951	2.344	1.964	3.497
Min	-4.553	-3.130	-4.553	-2.116	-2.354	-3.130	Min	-13.569	-10.703	-13.569	-10.703	-11.219	-9.106
Max	5.461	5.745	5.461	5.745	3.937	5.032	Max	13.926	15.588	7.518	8.836	13.926	15.588
%> 0	72.619	73.810	85.714	80.952	59.524	66.667	%> 0	78.571	75	73.810	69.048	83.333	80.952
$\%p <= 0.05$	3.571	8.333	7.143	16.667	0	0	$\%p <= 0.05$	4.762	15.476	0	4.762	9.524	26.190
Interest rate	RMSFE	CRPS	RMSFE	CRPS	RMSFE	CRPS							
Mean	1.627	1.764	-0.253	0.435	3.506	3.093							
Median	1.030	3.074	-1.630	2.123	3.445	4.184							
Min	-12.774	-16.290	-11.779	-15.435	-12.774	-16.290							
Max	19.525	20.786	18.605	20.046	19.525	20.786							
%> 0	55.952	60.714	38.095	57.143	73.810	64.286							
$\%p <= 0.05$	23.810	33.333	7.143	23.810	40.476	42.857							

Notes: This table presents descriptive statistics on comparisons of forecasting performance for the country-specific VAR-SV model and multi-country VAR-SV model with the Horseshoe prior (the paper's HS specification). Descriptive statistics include average, median, minimum, maximum, percentage of cases in which gains are above 0, and the percentage gains in which the forecasts from the competing models are statistically different according to the Diebold-Mariano (1995) test with fixed-smoothing asymptotics as in Coroneo and Iacone (2020).

Table 10: Loss function levels for the benchmark CVAR specification

	RMSFE				CRPS			
Output growth	$h = 1$	$h = 4$	$h = 8$	$h = 12$	$h = 1$	$h = 4$	$h = 8$	$h = 12$
CAN	2.319	2.712	2.637	2.647	1.245	1.466	1.419	1.402
DEU	3.592	3.595	3.569	3.470	1.832	1.805	1.798	1.764
FRA	1.630	2.068	2.116	2.144	0.892	1.115	1.134	1.149
ITA	2.539	2.985	2.964	2.935	1.335	1.580	1.553	1.515
JPN	4.185	4.167	4.156	4.251	2.208	2.187	2.159	2.261
UK	2.070	2.542	2.509	2.534	1.080	1.311	1.282	1.287
USA	2.335	2.548	2.594	2.542	1.266	1.364	1.393	1.367
Inflation	$h = 1$	$h = 4$	$h = 8$	$h = 12$	$h = 1$	$h = 4$	$h = 8$	$h = 12$
CAN	1.870	1.722	1.816	1.809	0.998	1.006	1.060	1.085
DEU	1.140	1.277	1.377	1.376	0.663	0.743	0.814	0.794
FRA	1.118	1.410	1.439	1.456	0.618	0.783	0.848	0.875
ITA	0.934	1.498	1.690	1.777	0.503	0.830	0.946	1.005
JPN	1.652	1.797	1.846	1.858	0.891	0.978	1.015	1.023
UK	0.982	1.215	1.384	1.358	0.542	0.701	0.778	0.813
USA	2.151	2.227	2.223	2.193	0.988	1.102	1.185	1.160
Interest rate	$h = 1$	$h = 4$	$h = 8$	$h = 12$	$h = 1$	$h = 4$	$h = 8$	$h = 12$
CAN	0.474	1.327	2.130	2.588	0.231	0.704	1.187	1.502
DEU	0.323	1.068	1.795	2.202	0.162	0.589	1.083	1.374
FRA	0.416	1.298	2.083	2.445	0.192	0.675	1.182	1.419
ITA	0.461	1.502	2.518	3.190	0.234	0.777	1.393	1.817
JPN	0.177	0.740	1.342	1.577	0.068	0.278	0.536	0.676
UK	0.418	1.233	1.884	2.289	0.189	0.630	1.012	1.295
USA	0.353	1.188	2.076	2.644	0.173	0.648	1.205	1.588

References

Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. US Government Printing Office, Washington, 1965.

Elena Angelini, Magdalena Lalik, Michele Lenza, and Joan Paredes. Mind the gap: a multi-country BVAR benchmark for the eurosystem projections. *International Journal of Forecasting*, 35(4):1658–1668, 2019. doi: 10.1016/j.ijforecast.2018.12.004.

Jushan Bai and Serena Ng. Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221, 2002. doi: 10.1111/1468-0262.00273.

Angela Bitto and Sylvia Frühwirth-Schnatter. Achieving shrinkage in a time-varying parameter model framework. *Journal of Econometrics*, 210(1):75–97, 2019. doi: 10.1016/j.jeconom.2018.11.006.

Annalisa Cadonna, Sylvia Frühwirth-Schnatter, and Peter Knaus. Triple the Gamma — a unifying shrinkage prior for variance and variable selection in sparse state space and TVP models. *Econometrics*, 8(2): 1–36, 2020. doi: 10.3390/econometrics8020020.

Fabio Canova and Matteo Ciccarelli. Estimating multicountry VAR models. *International Economic Review*, 50(3):929–959, 2009. doi: 10.1111/j.1468-2354.2009.00554.x.

Fabio Canova and Matteo Ciccarelli. Panel vector autoregressive models: A survey. In *VAR Models in Macroeconomics – New Developments and Applications: Essays in Honor of Christopher A. Sims*, volume 32 of *Advances in Econometrics*, pages 205–246. Emerald Group Publishing Limited, December 2013. doi: 10.1108/S0731-9053(2013)0000031006.

Fabio Canova, Matteo Ciccarelli, and Eva Ortega. Similarities and convergence in G-7 cycles. *Journal of Monetary Economics*, 54(3):850–878, 2007. doi: 10.1016/j.jmoneco.2005.10.022.

Andrea Carriero, Joshua Chan, Todd E. Clark, and Massimiliano Marcellino. Corrigendum to: Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors. *Journal of Econometrics*, forthcoming, 2022. doi: 10.1016/j.jeconom.2021.11.010.

Joshua C.C. Chan. Notes on Bayesian macroeconometrics. 2017. URL `http://joshuachan.org/papers/BayesMacro.pdf`.

Joshua C.C. Chan. Specification tests for time-varying parameter models with stochastic volatility. *Econometric Reviews*, 37(8):807–823, 2018. doi: 10.1080/07474938.2016.1167948.

Joshua C.C. Chan. Minnesota-type adaptive hierarchical priors for large Bayesian VARs. *International Journal of Forecasting*, 37(3):1212–1226, 2021. doi: 10.1016/j.ijforecast.2021.01.002.

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

Timothy Cogley and Thomas J. Sargent. Drifts and volatilities: Monetary policies and outcomes in the post WWII US. *Review of Economic Dynamics*, 8(2):262–302, 2005. doi: 10.1016/j.red.2004.10.009.

Laura Coroneo and Fabrizio Iacone. Comparing predictive accuracy in small samples using fixed-smoothing asymptotics. *Journal of Applied Econometrics*, 35(4):391–409, 2020. doi: 10.1002/jae.2756.

Antonello D'Agostino, Luca Gambetti, and Domenico Giannone. Macroeconomic forecasting and structural change. *Journal of Applied Econometrics*, 28(1):82–101, 2013. doi: 10.1002/jae.1257.

Stéphane Dées and Jochen Güntner. Forecasting inflation across euro area countries and sectors: A panel VAR approach. *Journal of Forecasting*, 36(4):431–453, 2017. doi: 10.1002/for.2444.

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

Jim Griffin and Phil Brown. Hierarchical shrinkage priors for regression models. *Bayesian Analysis*, 12 (1):135–159, 2017. doi: 10.1214/15-BA990.

Florian Huber. Density forecasting using Bayesian global vector autoregressions with stochastic volatility. *International Journal of Forecasting*, 32(3):818–837, 2016. doi: 10.1016/j.ijforecast.2015.12.008.

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

Gary Koop and Dimitris Korobilis. Forecasting with high-dimensional panel VARs. *Oxford Bulletin of Economics and Statistics*, 81(5):937–959, 2019. doi: 10.1111/obes.12303.

Dimitris Korobilis. Prior selection for panel vector autoregressions. *Computational Statistics and Data Analysis*, 101:110–120, 2016. doi: 10.1016/j.csda.2016.02.011.

Kamiar Mohaddes and Mehdi Raissi. Compilation, revision, and updating of the global VAR (GVAR) database, 1979q2-2019q4. 2020. doi: 10.17863/CAM.56762.

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

M. Hashem Pesaran and Allan Timmermann. A simple nonparametric test of predictive performance. *Journal of Business and Economic Statistics*, 10(4):461–465, 1992. doi: 10.1080/07350015.1992.10509922.

M. Hashem Pesaran, Til Schuermann, and L. Vanessa Smith. Forecasting economic and financial variables with global VARs. *International Journal of Forecasting*, 25(4):642–675, 2009. doi: 10.1016/j.ijforecast.2009.08.007.

Giorgio E. Primiceri. Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies*, 72(3):821–852, 2005. doi: 10.1111/j.1467-937X.2005.00353.x.

Gareth O. Roberts and Jeffrey S. Rosenthal. Examples of adaptive MCMC. *Journal of Computational and Graphical Statistics*, 18(2):349–367, 2009. doi: 10.1198/jcgs.2009.06134.

James H. Stock and Mark W. Watson. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97(460):1167–1179, 2002a. doi: 10.1198/016214502388618960.

James H. Stock and Mark W. Watson. Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, 20(2):147–162, 2002b. doi: 10.1198/073500102317351921.