Forecasting Related Time Series
— Appendices —

Ulrich K. Müller and Mark W. Watson
Department of Economics
Princeton University

This Draft: December 19, 2025

A Data

The state employment data are from the U.S. Bureau of Labor Statistics and are "Total Nonfarm Employees" in each state. The data were downloaded from the FRB-St.Louis FRED database. As examples, data for Alaska and Wyoming are the FRED series AKNA and WYNA.

The Euro-area industrial production data include data from the 16 countries: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden and the UK. These data are from OECD and were downloaded from FRED. As examples, data for Austria and the UK are the FRED series AUTPROINDMISMEI and GBRPROINDMISMEI.

The 17 sectoral price series for personal consumption expenditures in the U.S. are from the U.S. Bureau of Economic Analysis and are taken from the monthly NIPA tables 2.3.4 and 2.3.5. These contain a 16-sector decomposition of the PCE. Following Stock and Watson (2016), we decomposed the "housing and utilities" sector into the two sectors "housing-energy" and "housing excluding energy."

B Computation

B.1 General comments

The posteriors were obtained by standard MCMC with numerous Gibbs steps. Detailed descriptions of the steps for each model are in the following subsections. Here we provide some generic comments on the numerical implementation.

B.1.1 Posterior Draw from Linear Gaussian State Space System (SSS)

Consider a generic state space system

$$\begin{aligned} y_t &= h_t' s_t + w_t,\ w_t \sim \mathcal{N}(0, R_t) \\ s_t &= F_t s_{t-1} + u_t,\ u_t \sim \mathcal{N}(0, Q_t),\ t = 2, \ldots, T \\ s_1 &\sim \mathcal{N}(s_{1|0}, P_{1|0}) \end{aligned}$$

where y_t is a scalar. The evaluation of the log-likelihood $\sum_{t=1}^T \ell_t$ of $\{y_t\}_{t=1}^T$ after integrating out $\{s_t\}_{t=1}^T$ can be obtained by the following Kalman iterations for $t = 1, \ldots, T$:

1. $\ell_t = -\frac{1}{2}e_t^2/\omega_t^2 - \frac{1}{2}\log\omega_t^2$ with $e_t = y_t - h_t' s_{t|t-1}$, $\omega_t^2 = h_t' v_t + R_t$ and $v_t = P_{t|t-1}h_t$

2. $s_{t|t} = s_{t|t-1} + v_t e_t/\omega_t^2$

3. $P_{t|t} = P_{t|t-1} - v_t v_t'/\omega_t^2$

4. $s_{t+1|t} = F_{t+1}s_{t|t}$

5. $P_{t+1|t} = F_{t+1}P_{t|t}F_{t+1}' + Q_{t+1}$.

To generate a random draw $\{s_t^{**}\}_{t=1}^T$ from the posterior $\{s_t\}_{t=1}^T|\{y_t\}_{t=1}^T$, we employ the algorithm developed in Durbin and Koopman (2002). The following expressions are simple rearrangements of the formulas given there, optimized for a scalar measurement and computational efficiency: Set $s_{1|0}^* = 0$ and $s_1^* \sim \mathcal{N}(0, P_{1|0})$ and iterate for $t = 1, \ldots, T$

1. $b_t = v_t/\omega_t^2$ where $\omega_t^2 = h_t' v_t + R_t$ and $v_t = P_{t|t-1}h_t$

2. $a_t = (y_t - h_t' s_{t|t-1})/\omega_t^2$ and $a_t^* \sim (h_t'(s_t^* - s_{t|t-1}^*) + \varepsilon_t^*)/\omega_t^2$ with $\varepsilon_t^* \sim \mathcal{N}(0, R_t)$

3. $s_{t|t} = s_{t|t-1} + v_t a_t$, $s_{t|t}^* = s_{t|t-1}^* + v_t a_t^*$

4. $P_{t|t} = P_{t|t-1} - b_t v_t'$

5. $s_{t+1}^* \sim \mathcal{N}(F_{t+1}s_t^*, Q_{t+1})$

6. $s_{t+1|t} = F_{t+1}s_{t|t}$, $s_{t+1|t}^* = F_{t+1}s_{t|t}^*$

7. $P_{t+1|t} = F_{t+1}P_{t|t}F_{t+1}' + Q_{t+1}$

followed by an iteration $t = T, T-1, \ldots, 1$ with initial values $r_T = r_T^* = 0$

1. $r_{t-1} = F_t' r_t + h_t(a_t - b_t' F_t' r_t)$, $r_{t-1}^* = F_t' r_t^* + h_t(a_t^* - b_t' F_t' r_t^*)$

2. $s_t^{**} = s_{t|t-1} + P_{t|t-1}r_{t-1} + s_t^* - s_{t|t-1}^* - P_{t|t-1}r_{t-1}^*$

so the scalars $\{a_t\}_{t=1}^T, \{a_t^*\}_{t=1}^T$, vectors $\{b_t\}_{t=1}^T$, $\{s_{t|t-1}\}_{t=1}^T$, $\{s_t^*\}_{t=1}^T$, and matrices $\{P_{t|t-1}\}_{t=1}^T$ must be saved during the first set of iterations to generate the draws $\{s_t^{**}\}_{t=1}^T$.

If the posterior means $s_{t|T}$ and covariance matrices $P_{t|T}$ of $s_t|\{y_t\}_{t=1}^T$ are also needed, then they can be computed by adding to the above iteration $t = T, T-1, \ldots, 1$ the following three steps with initial value $G_T = 0$

3. $G_{t-1} = A_t - h_t x_t' - x_t h_t' + h_t h_t'(b_t' x_t + 1/\omega_t^2)$ with $A_t = F_t' G_t F_t$ and $x_t = A_t b_t$

4. $P_{t|T} = P_{t|t-1} - P_{t|t-1} G_{t-1} P_{t|t-1}$

5. $s_{t|T} = s_{t|t-1} + P_{t|t-1} r_{t-1}$

In applications of this algorithm, the vector h_t and the matrices $P_{1|0}$, Q_t and F_t are often sparse, which we exploit to further gain computational efficiency.

B.1.2 Posterior Draw from Gaussian Hierarchical Model with Conjugate Likelihood

Suppose $\theta_0 \sim \mathcal{N}(\mu_0, \Omega_0)$, $\theta_j|\theta_0 \sim \mathcal{N}(\theta_0, \Omega_j)$ and the log-likelihood of the observations is of the form $C - \frac{1}{2}\sum_{j=1}^n (Y_j - \theta_j)'\Sigma_j^{-1}(Y_j - \theta_j) + \log\det\Sigma_j$ for some conformable Y_j (as would be the case for independent observations $Y_j \sim \mathcal{N}(\theta_j, \Sigma_j)$). Elementary calculations show that the posterior of θ_0 is Gaussian $\mathcal{N}(V_0(\Omega_0^{-1}\mu_0 + \sum_{j=1}^n (\Sigma_j + \Omega_j)^{-1} Y_j), V_0)$ with $V_0^{-1} = \Omega_0^{-1} + \sum_{j=1}^n (\Sigma_j + \Omega_j)^{-1}$. Furthermore, conditional on θ_0, the posteriors for $\{\theta_j\}_{j=1}^n$ are independent Gaussian $\mathcal{N}(V_j \Sigma_j^{-1}(Y_j - \theta_0) + \theta_0, V_j)$ with $V_j^{-1} = \Omega_j^{-1} + \Sigma_j^{-1}$.

B.1.3 Draws from Posterior with Hierarchical Normal Prior

Suppose the prior for n scalar parameters is $\{\theta_j\}_{j=1}^n \sim \mathcal{HN}(m_{m_\theta}, v_{m_\theta}, m_{\ln(v_\theta)}, v_{\ln(v_\theta)})$, and the likelihood factors in θ_j. Then the standard way of generating draws from θ_j is to (i) condition on m_θ and v_θ and update θ_j from the likelihood information and the prior $\theta_j \sim \mathcal{N}(m_\theta, v_\theta)$ separately for $j = 1, \ldots, n$; (ii) condition on $\{\theta_j\}_{j=1}^n$ and update m_θ and v_θ.

This standard approach will not lead to a well-mixing chain when n is large, however. To see why, consider the extreme case where the likelihood is uninformative about θ_j. Then in the first step, $\theta_j \sim iid\mathcal{N}(m_\theta, v_\theta)$. By the law of large numbers, $n^{-1}\sum_{j=1}^n \theta_j \approx m_\theta$ and $n^{-1}\sum_{j=1}^n (\theta_j - m_\theta)^2 \approx v_\theta$, so in the second step, we recover nearly the same parameters (m_θ, v_θ) that we started with.

A better approach, especially if the likelihood is not very informative, is to condition in the second step on the current "z-scores" $\{z_j\}_{j=1}^n$ with $z_j = \frac{\theta_j - m_\theta}{\sqrt{v_\theta}}$. Given the current value (m_θ^c, v_θ^c), a random walk Metropolis proposal $(m_\theta^p, \ln v_\theta^p)' \sim \mathcal{N}((m_\theta^c, \ln v_\theta^c)', \Lambda)$ then induces the values $\theta_j^p = m_\theta^p + \sqrt{v_\theta^p / v_\theta^c}(\theta_j^c - m_\theta^c)$, $j = 1, \ldots, n$. The acceptance probability for the proposal involves the probability of $(m_\theta^p, \ln v_\theta^p)$ relative to $(m_\theta^c, \ln v_\theta^c)$ (in the hierarchical normal model, computed from $m_\theta \sim \mathcal{N}(m_{m_\theta}, v_{m_\theta})$ and $\ln(v_\theta) \sim \mathcal{N}(m_{\ln(v_\theta)}, v_{\ln(v_\theta)})$), and the likelihood of $\{\theta_j^p\}_{j=1}^n$ relative to $\{\theta_j^c\}_{j=1}^n$, but it does not involve the prior $\theta_j \sim \mathcal{N}(m_\theta, v_\theta)$, as the z-scores are, by construction, equally likely for all values of (m_θ, v_θ). Of course, if the evaluation of the likelihood is computationally expensive, then performing this step is much slower than conditioning on $\{\theta_j\}_{j=1}^n$. But note that in the flat likelihood example, this alternative approach has excellent mixing properties, as it mixes as well as a random walk Metropolis chain that explores (m_θ, v_θ) under the only information that $m_\theta \sim \mathcal{N}(m_{m_\theta}, v_{m_\theta})$ and $\ln(v_\theta) \sim \mathcal{N}(m_{\ln(v_\theta)}, v_{\ln(v_\theta)})$.

B.1.4 Geweke (2004) Test

It is notoriously easy to make coding mistakes in posterior samplers. We tested (each component of) our code with the Geweke (2004) test, in the implementation described in Müller and Watson (2020).

B.1.5 Random Walk Metropolis Step Size

The algorithms involve random walk Metropolis draws. Let $\theta = (\theta_1, \ldots, \theta_k)$ be a vector valued parameter that is subject to a Metropolis step. If the current value is θ_c, the proposed value is drawn from $\theta_p \sim \mathcal{N}(\theta_c, \kappa_0^2 \tau \operatorname{diag}(\kappa_1^2 v_1, \ldots, \kappa_k^2 v_k))$, where v_j are the prior variances, $\tau \in \{n^{-1}, T^{-1}, (nT)^{-1}\}$ depending on the order of accumulation of information about θ, and the κ_j are positive constants. Note that in the presence of hierarchical priors, the prior vector $(v_1, \ldots, v_k)$ might itself be different from draw to draw. The constants κ_j are determined in the burn-in phase to approximately yield a 50% acceptance rate. In particular, we keep track of the $k+1$ acceptance rates $\hat{a}_i$, $i = 0, \ldots, k$ in the last 200 draws. Here $\hat{a}_0$ corresponds to the rate under the proposal above, and $\hat{a}_i$ for $i > 0$ is the rate for the proposal that moves only one element, $\mathcal{N}(\theta_c, \text{diag}(0, \dots, 0, \kappa_0^2\kappa_i^2\tau v_i, 0, \dots, 0))$. We then update $\{\kappa_i\}_{i=0}^k$ via

$$\kappa_i \leftarrow \left(\left(\left(\frac{\hat{a}_i}{1 - \hat{a}_i} \right)^{0.4} \wedge 3 \right) \vee 1/3 \right) \kappa_i$$

that is, the more $\hat{a}_i$ deviates from 50%, the larger the adjustment, but only up to a maximal increase or decrease by factor of 3 or 1/3, respectively. This process is repeated every 200 draws for a phase in the burn-in period. The initial baseline values of κ_i were also determined in this fashion, and then hard-coded and held constant across variations of the sample size (n, T), data sets and variations of the model (which more plausibly yields a reasonable acceptance rate due to the presence of the deterministic τ).

The process to update the κ_i is computationally costly, as it requires $2(k+1)$ evaluations of the likelihood, rather than just 2. To reduce this burden, we treat vectors of a parameter that corresponds to the $p = 12$ AR coefficients $\{\phi_{j,t,l}\}_{l=1}^p$ as one block with a single corresponding κ. For the common value of the AR coefficients $\{m_{\phi_l}\}_{l=1}^p$, the proposal variance is further multiplied by $\text{diag}(1, 2, \dots, p)$, with the idea that the likelihood has much more curvature for l small relative to the sharply decreasing prior variances. We similarly treat the q stochastic volatility components as one block and rescale the proposal variance by $\text{diag}(1, 2, \dots, q)$.

B.1.6 Burn-in and Number of Draws

We initialize the sampler with $u_{n+1,t} = 0$, $\mu_{j,t} = o_{j,t} = 0$ for $t = 1, \dots, T$ and $j = 1, \dots, n+1$ and all parameters are equal to the prior mean. For a sampler that generates N total usable MCMC draws, we use a burn-in period that consists of three phases: First, we take 200 draws that do not update any priors or other common components such as $u_{n+1,t}$, $\mu_{n+1,t}$ or $o_{n+1,t}$. Second, for $l = 1, \dots, N/3$ further draws that involve all steps, we set the Metropolis step size standard deviations equal to $5^{(1-3l/N)}$ of its baseline value (that is, κ_0 of the previous subsection is inflated by a factor $5^{(1-3l/N)}$). The idea is to allow the sampler to make larger movements initially to quickly approach values where the posterior is high. Finally, for another $N/3$ draws, we adjust the Metropolis step sizes as described in the last subsection.

We set $N = 1000$ when $n < 20$ and $N = 2000$ for the employment application for models I-VI, and a twice as large value of N for the RTS model. It takes about one minute to generate the $N = 4000$ draws for the employment application in the RTS model, including the burn-in, in a Fortran implementation on a 24 core workstation.

B.2 Details on the Gibbs Steps for each Model

The description of the algorithms reference the numbered equations in the paper. For convenience, we have reproduced those equations here.

$$\phi_i(L)(y_{i,t} - \mu_i) = \varepsilon_{i,t} \tag{1}$$

$$\{\theta_j\}_{j=1}^n \sim \mathcal{HN}(m_{m_\theta}, v_{m_\theta}, m_{\ln(v_\theta)}, v_{\ln(v_\theta)}) \tag{2}$$

$$\begin{aligned} y_{j,t} &= \mu_j + \omega u_{j,t} && (3)\\ u_{j,t} &= \sum_{l=1}^{p} \phi_{j,l} u_{j,t-l} + \epsilon_{j,t} && (4)\\ \epsilon_{j,t} &= \sigma_j \varepsilon_{j,t} && (5)\\ \varepsilon_{j,t} &\sim iid\mathcal{N}(0,1) && (6) \end{aligned}$$

$$\ln(\omega^2) \sim \mathcal{N}(0, \infty) \tag{7}$$

$$\mu_j \sim \mathcal{N}(0, \infty) \tag{8}$$

$$u_{j,-11:0}|(\sigma_j, \phi_j) \sim \mathcal{N}(0, \sigma_j^2 \Sigma(\phi_j)) \text{ with } \Sigma(\phi) = \Sigma_{AR}(c\phi) \tag{9}$$

$$\phi_{j,l} \sim \mathcal{N}(0, (0.2/l)^2) \tag{10}$$

$$\ln(\sigma_j^2) \sim \mathcal{N}(0, 0.3^2) \tag{11}$$

$$\{\phi_{j,l}\}_{j=1}^n \sim \mathcal{HN}(0, 0.5^2(0.2/l)^2, \ln((0.2/l)^2), 0.5^2) \tag{12}$$

$$\{\ln(\sigma_j^2)\}_{j=1}^n \sim \mathcal{HN}(0, 0.5^2, \ln(0.3^2), 0.5^2) \tag{13}$$

$$\varepsilon_{j,t} \sim \mathcal{T}(\nu_j) \tag{14}$$

$$\{\ln(\nu_j - 2)\}_{j=1}^n \sim \mathcal{HN}(\ln(12-2), 0.5^2, \ln(0.5^2), 0.5^2) \tag{15}$$

$$y_{j,t} = \mu_j + \omega(u_{j,t} + o_{j,t}) \tag{16}$$

$$o_{j,t} = \kappa_j \eta_{j,t} \text{ and } \eta_{j,t} \sim \mathcal{T}(\nu_j^o) \tag{17}$$

$$\{\ln(\kappa_j^2)\}_{j=1}^n \sim \mathcal{HN}(\ln(0.1^2), 0.5^2, \ln(0.3^2), 0.5^2) \tag{18}$$

$$\{\ln(\nu_j^o - 2)\}_{j=1}^n \sim \mathcal{HN}(\ln(4-2), 0.5^2, \ln(0.5^2), 0.5^2) \tag{19}$$

$$\epsilon_{j,t} = \sigma_{j,t}\varepsilon_{j,t} \tag{20}$$

$$\ln(\sigma_{j,t}^2) = \ln(\sigma_j^2) + \sum_{l=1}^{q} \varphi_{l,t}\xi_{j,l} \tag{21}$$

$$\{\xi_{j,l}\}_{j=1}^n | (m_{\xi_l}, v_\xi) \sim iid\mathcal{N}(m_{\xi_l}, v_\xi) \tag{22}$$

$$m_{\xi_l} \sim \mathcal{N}(0, 0.01^2) \tag{23}$$

$$\ln(v_\xi) \sim \mathcal{N}(\ln(0.01^2), 0.5^2) \tag{24}$$

$$y_{j,t} = \mu_{j,t} + \omega(u_{j,t} + o_{j,t}) \tag{25}$$

$$u_{j,t} = \sum_{l=1}^{p} \phi_{j,l,t}u_{j,t-l} + \epsilon_{j,t} \tag{26}$$

$$\mu_{j,t} \sim RW(\mu_j, \omega^2\gamma_{\mu_j}^2) \tag{27}$$

$$\phi_{j,l,t} \sim RW(\phi_{j,l}, \gamma_{\phi_{j,l}}^2) \tag{28}$$

$$\{\ln(\gamma_{\mu_j}^2)\}_{j=1}^n \sim \mathcal{HN}(\ln(0.005^2), 2^2, \ln(0.3^2), 0.5^2) \tag{29}$$

$$\{\ln(\gamma_{\phi_{j,l}}^2)\}_{j=1}^n \sim \mathcal{HN}(\ln((0.005/l)^2), 0.5^2, \ln(0.3^2), 0.5^2) \tag{30}$$

$$y_{j,t} = \mu_{j,t} + \mu_{n+1,t} + \omega(u_{j,t} + o_{j,t} + o_{n+1,t} + c_{j,t}) \tag{31}$$

$$c_{j,t} = \sum_{l=0}^{5} \lambda_{j,l,t}u_{n+1,t-l} \tag{32}$$

$$\lambda_{j,l,t} \sim RW(\lambda_{j,l}, \gamma_{\lambda_{j,l}}^2). \tag{33}$$

$$\ln(\sigma_{n+1}^2) \sim \mathcal{N}(\ln(0.2^2), 0.5^2) \tag{34}$$

$$\mu_{n+1} = 0 \text{ (a normalization)}, \quad \ln(\gamma_{\mu_{n+1}}^2) \sim \mathcal{N}(\ln(0.005^2), 0.5^2) \tag{35}$$

$$\ln(\kappa_{n+1}^2) \sim \mathcal{N}(\ln(0.1^2), 0.5^2) \tag{36}$$

$$\{\lambda_{j,0}\}_{j=1}^n \sim \mathcal{HN}(1, 0, \ln(0.2^2), 0.5^2) \tag{37}$$

$$\{\lambda_{j,l}\}_{j=1}^{n} \sim \mathcal{HN}(0, 0.5^2(0.05/l)^2, \ln((0.05/l)^2), 0.5^2) \text{ for } l > 0 \tag{38}$$

$$\{\ln(\gamma^2_{\lambda(j,l)})\}_{j=1}^{n} \sim \mathcal{HN}(\ln((0.0005/(l+1))^2), 0.5^2, \ln(0.3^2), 0.5^2) \tag{39}$$

In our description of the algorithms, we condition on all variables if not stated otherwise. To ease notation, we do not explicitly mention updates to $\epsilon_{j,t} = \sigma_{j,t}\varepsilon_{j,t}$ that arise from changes in $\varepsilon_{j,t}$ and $\sigma_{j,t}$.

B.2.1 Model I: Bayesian shrinkage

1. $\{\omega, \{\sigma_j\}_{j=1}^n\}$: Conditioning on the current value of $\{\ln \omega^2 + \ln \sigma_j^2\}_{j=1}^n$, draw $\ln \omega^2 | \{\ln \omega^2 + \ln \sigma_j^2\}_{j=1}^n$ from conjugate normal obtained from (11) and (7), then update $\{\sigma_j\}_{j=1}^n$ according to new ω. [The model depends on $\{\omega, \{\sigma_j\}_{j=1}^n\}$ only through the products $\{\omega\sigma_j\}_{j=1}^n$, so there is no additional contribution to the posterior.]

2. $\{\sigma_j, \mu_j, u_{j,-11:0}, \{\varepsilon_{j,t}\}_{t=1}^T\}$ looping over j:

   (a) Draw σ_j: RW Metropolis step with prior (11) and likelihood computed from Kalman filter with state $(\mu_j, u_{j,t-1}, u_{j,t-2}, \ldots, u_{j,t-12})$, measurement equation (3) and state evolution (4), initial state drawn from $\mu_j \sim \mathcal{N}(0, \infty)$ (approximated by using large but finite variance) and (9).

   (b) Draw $\{\mu_j, u_{j,-11:0}, \{\varepsilon_{j,t}\}_{t=1}^T\} | \sigma_j$: Kalman smoother draw from same SSS as in Step 2a.

3. $\{\phi_j, \{\varepsilon_{j,t}\}_{t=1}^T\}$ looping over j: Metropolis-Hastings step with proposal generated from Kalman smoother draw from SSS with state $\phi_j = (\phi_{j,1}, \ldots, \phi_{j,12})$, measurement equation (4) and initial state (10). The proposal ϕ_j^p is accepted over the current value ϕ_j^c with probability $1 \wedge \frac{L_9(\phi_j^p)}{L_9(\phi_j^c)}$, where $L_9(\phi_j)$ is the likelihood of (9). (This Metropolis step adjusts the Gibbs step for the initial values.)

B.2.2 Model II: Hierarchical priors

1. $\{\omega, m_{\ln \sigma^2}\}$: Conditioning on the current value of $\ln \omega^2 + m_{\ln \sigma^2}$, draw $\ln \omega^2 | \ln \omega^2 + m_{\ln \sigma^2}$ from conjugate normal obtained from (13) and (7), then update $m_{\ln \sigma^2}$ accordingly.

2. $\{m_{\ln \sigma^2}, v_{\ln \sigma^2}, \{\mu_j, u_{j,-11:0}, \sigma_j, \{\varepsilon_{j,t}\}_{t=1}^T\}_{j=1}^n\}$:

(a) Draw $\{m_{\ln \sigma^2}, v_{\ln \sigma^2}\}|\{(\ln \sigma_j^2 - m_{\ln \sigma^2})/\sqrt{v_{\ln \sigma^2}}\}_{j=1}^n$: Bivariate RW Metropolis step with prior (13) and likelihood computed from Kalman filter of Model I Step 2a applied to $j = 1, \ldots, n$. Update $\{\sigma_j^2\}_{j=1}^n$ if accepted.

(b) $\{\sigma_j, \mu_j, u_{j,-11:0}, \{\varepsilon_{j,t}\}_{t=1}^T\}_{j=1}^n | m_{\ln \sigma^2}, v_{\ln \sigma^2}$: Same as Model I Step 2, except that the prior $\ln \sigma_j^2 \sim iid\mathcal{N}(m_{\ln \sigma^2}, v_{\ln \sigma^2})$ is used in place of (11).

3. $\{\{m_{\phi_l}, v_{\phi_l}\}_{l=1}^{12}, \{\phi_j, \{\varepsilon_{j,t}\}_{t=1}^T\}_{j=1}^n\}$:

   (a) Draw $\{m_{\phi_l} v_{\phi_l}\}_{l=1}^{12} | \{\{(\phi_{j,l} - m_{\phi_l})/\sqrt{v_{\phi_l}}\}_{l=1}^{12}\}_{j=1}^n$: 24-dimensional RW Metropolis step, with prior (12) and likelihood computed from the product over $j = 1, \ldots, n$ of (9) and Kalman filters with state $\phi_j = (\phi_{j,1}, \ldots, \phi_{j,12})$ and measurement equations (4). Update $\{\phi_j\}_{j=1}^n$ if accepted.

   (b) $\{\phi_j, \{\varepsilon_{j,t}\}_{t=1}^T\} | \{m_{\phi_l}, v_{\phi_l}\}_{l=1}^{12}$ looping over j: Metropolis-Hastings step with proposal generated from Kalman smoother draw from SSS of Step 3a with initial state $\phi_j \sim \mathcal{N}((m_{\phi_1}, \ldots, m_{\phi_{12}}), \operatorname{diag}(v_{\phi_1}, \ldots, v_{\phi_{12}}))$. The proposal ϕ_j^p is accepted over the current value ϕ_j^c with probability $(1 \wedge \frac{L_9(\phi_j^p)}{L_9(\phi_j^c)})$, where $L_9(\phi_j)$ is the likelihood of (9).

4. $\{\{m_{\phi_l}\}_{l=1}^{12}, \{\phi_j, \{\varepsilon_{j,t}\}_{t=1}^T\}_{j=1}^n\}$: Ignoring the likelihood contribution from (9), the log-likelihood is quadratic in ϕ_j. We can thus generate a Metropolis-Hastings proposal of $\{m_{\phi_l}\}_{l=1}^{12}$ and $\{\phi_j\}_{j=1}^n$ using the algorithm in Section B.1.2, and accept it with probability $1 \wedge \frac{\prod_{j=1}^n L_9(\phi_j^p)}{\prod_{j=1}^n L_9(\phi_j^c)}$ in obvious notation. [This step is not needed given Step 3, but it improves mixing.]

B.2.3 Model III: Student-t innovations

Let $S_{j,t}$ be independent draws of $\nu_j/\chi^2_{\nu_j}$. Then $\varepsilon_{j,t} \sim \mathcal{T}(\nu_j)$ from (14) can be represented as

$$\varepsilon_{j,t} = \sqrt{S_{j,t}} z_{j,t}, \quad z_{j,t} \sim iid\mathcal{N}(0, 1).$$

The sampler for Model III (and those below) treats $\{\{S_{j,t}\}_{t=1}^T\}_{j=1}^n$ as an additional unobserved component (which we condition on if not explicitly part of a block), so that after conditioning, we recover a Gaussian model for $\varepsilon_{j,t}|S_{j,t}$ and thus $u_{j,t}|S_{j,t}, \phi_j, \sigma_j^2$.

1. $\{m_{\ln(\nu-2)}, v_{\ln(\nu-2)}, m_{\ln \sigma^2}, v_{\ln \sigma^2}, \{\nu_j, \sigma_j^2, \{S_{j,t}\}_{t=1}^T\}_{j=1}^n\}$:

(a) $\{m_{\ln(\nu-2)}, v_{\ln(\nu-2)}, m_{\ln\sigma^2}, v_{\ln\sigma^2}\}|\{(\ln(\nu_j - 2) - m_{\ln(\nu-2)})/\sqrt{v_{\ln(\nu-2)}}, (\ln\sigma_j^2 - m_{\ln\sigma^2})/\sqrt{v_{\ln\sigma^2}}\}_{j=1}^n$: Four dimensional RW Metropolis step with priors (13), (15) and likelihood computed from student-t density (14). Update $\{\nu_j, \sigma_j^2\}_{j=1}^n$ if accepted.

(b) $\{\ln(\nu_j - 2), \sigma_j^2\}_{j=1}^n|m_{\ln(\nu-2)}, v_{\ln(\nu-2)}, m_{\ln\sigma^2}, v_{\ln\sigma^2}$: Looping over j, bivariate RW Metropolis step with priors (13), (15) and likelihood computed from (14).

(c) $\{\{S_{j,t}\}_{t=1}^T\}_{j=1}^n|m_{\ln(\nu-2)}, v_{\ln(\nu-2)}, m_{\ln\sigma^2}, v_{\ln\sigma^2}, \{\nu_j, \sigma_j^2\}_{j=1}^n$: Looping over j, draw $S_{j,t}$ independently from $(\nu_j + \varepsilon_{j,t}^2)/\chi^2_{\nu_j+1}$, $t = 1, \ldots, T$.

[Given that $m_{\ln\sigma^2}, v_{\ln\sigma^2}, \{\sigma_j^2\}$ are also updated in Step 2 below, we could keep them fixed in this step, but it improves mixing to exploit the relatively weaker informativeness of the student-t likelihood to update the variance parameters in both steps.]

2. Perform Steps 1, 3 and 4 of Model II, except that in the SSS, the variance of the measurement equation is now given by $\sigma_j^2 S_{j,t}$.

B.2.4 Model IV: Additive outliers

In analogy to Model III, $\eta_{j,t} \sim \mathcal{T}(\nu_j^o)$ can be represented as

$$\eta_{j,t} = \sqrt{S_{j,t}^o} z_{j,t}^o, \qquad z_{j,t}^o \sim iid\mathcal{N}(0,1)$$

where $S_{j,t}^o$ are independent draws of $\nu_j^o/\chi^2_{\nu_j^o}$. The sampler for Model IV (and those below) treats $\{\{S_{j,t}^o\}_{t=1}^T\}_{j=1}^n$ as an additional unobserved component (which we condition on if not explicitly part of a block), so that after conditioning, we recover a Gaussian model for $\eta_{t,j}|S_{j,t}^o$. To ease notation, we do not explicitly mention updates to $o_{j,t} = \kappa_{j,t}\eta_{j,t}$ that arise from changes in $\eta_{j,t}$ and $\kappa_{j,t}$.

1. $\{\omega, m_{\ln\kappa^2}, m_{\ln\sigma^2}\}$: Conditioning on the current values of $\ln\omega^2 + m_{\ln\sigma^2}$ and $\ln\omega^2 + m_{\ln\kappa^2}$, draw $\ln\omega^2|(\ln\omega^2 + m_{\ln\sigma^2}, \ln\omega^2 + m_{\ln\kappa^2})$ from conjugate normal obtained from (13), (18) and (7), then update $(m_{\ln\sigma^2}, m_{\ln\kappa^2})$ accordingly.

2. $\{m_{\ln\kappa^2}, v_{\ln\kappa^2}, m_{\ln\sigma^2}, v_{\ln\sigma^2}, \{\mu_j, u_{j,-11:0}, \sigma_j, \kappa_j, \{\eta_{j,t}, \varepsilon_{j,t}\}_{t=1}^T\}_{j=1}^n\}$:

   (a) Draw $\{m_{\ln\kappa^2}, v_{\ln\kappa^2}, m_{\ln\sigma^2}, v_{\ln\sigma^2}\}|\{(\ln\sigma_j^2 - m_{\ln\sigma^2})/\sqrt{v_{\ln\sigma^2}}, (\ln\kappa_j^2 - m_{\ln\kappa^2})/\sqrt{v_{\ln\kappa^2}}\}_{j=1}^n$: 4-dimensional RW Metropolis step with priors (13), (18) and likelihood computed from Kalman filter with state $(\mu_j, u_{j,t-1}, u_{j,t-2}, \ldots, u_{j,t-12})$, measurement equation (16), state evolution (4), and initial state $\mu_j \sim \mathcal{N}(0, \infty)$ (approximated by using large but finite variance) and (9). Update $\{\sigma_j, \kappa_j\}_{j=1}^n$ if accepted.

(b) $\{\ln \sigma_j^2, \ln \kappa_j^2\}_{j=1}^n | m_{\ln \kappa^2}, v_{\ln \kappa^2}, m_{\ln \sigma^2}, v_{\ln \sigma^2}$: Looping over j, bivariate RW Metropolis step with prior (15), (18) and likelihood computed from same Kalman filter as in Step 2a.

(c) $\{\mu_j, u_{j,-11:0}, \{\eta_{j,t}, \varepsilon_{j,t}\}_{t=1}^T\} | m_{\ln \kappa^2}, v_{\ln \kappa^2}, m_{\ln \sigma^2}, v_{\ln \sigma^2}, \{\sigma_j, \kappa_j\}_{j=1}^n$: Looping over j, draw from Kalman smoother from same SSS as in Step 2a.

3. $\{m_{\ln(\nu^o-2)}, v_{\ln(\nu^o-2)}, m_{\ln \kappa^2}, v_{\ln \kappa^2}, \{\nu_j^o, \kappa_j, \{S_{j,t}^o\}_{t=1}^T\}_{j=1}^n\}$:

   (a) $\{m_{\ln(\nu^o-2)}, v_{\ln(\nu^o-2)}, m_{\ln \kappa^2}, v_{\ln \kappa^2}\} | \{(\ln(\nu_j^o - 2) - m_{\ln(\nu^o-2)})/\sqrt{v_{\ln(\nu^o-2)}}, (\ln \kappa_j^2 - m_{\ln \kappa^2})/\sqrt{v_{\ln \kappa^2}}\}_{j=1}^n$: Four dimensional RW Metropolis step with priors (18), (19) and likelihood computed from (17). Update $\{\nu_j^o, \kappa_j^2\}_{j=1}^n$ if accepted.

   (b) $\{\ln(\nu_j^o - 2), \kappa_j\}_{j=1}^n | m_{\ln(\nu^o-2)}, v_{\ln(\nu^o-2)}, m_{\ln \kappa^2}, v_{\ln \kappa^2}$: Looping over j, bivariate Metropolis step with priors (18), (19) and student-t likelihood (17).

   (c) $\{\{S_{j,t}^o\}_{t=1}^T\}_{j=1}^n | m_{\ln(\nu^o-2)}, v_{\ln(\nu^o-2)}, m_{\ln \kappa^2}, v_{\ln \kappa^2} \{\nu_j^o, \kappa_j\}_{j=1}^n$: Looping over j, draw $S_{j,t}^o$ independently from $(\nu_j^o + \eta_{j,t}^2)/\chi^2_{\nu_j^o+1}$, $t = 1, \ldots, T$.

4. Perform Step 1 of Model III and Steps 3-4 of Model II, except that in the SSS, the variance of the measurement equation is given by $\sigma_j^2 S_{j,t}$.

B.2.5 Model V: Time varying volatility

1. $\{m_{\ln \kappa^2}, v_{\ln \kappa^2}, m_{\ln \sigma^2}, v_{\ln \sigma^2}, \{m_{\xi_l}\}_{l=1}^q, v_\xi, \{\mu_j, u_{j,-11:0}, \sigma_j, \{\xi_{j,l}\}_{l=1}^q, \kappa_j, \{\eta_{j,t}, \varepsilon_{j,t}\}_{t=1}^T\}_{j=1}^n\}$:

   (a) Draw $(m_{\ln \kappa^2}, v_{\ln \kappa^2}, m_{\ln \sigma^2}, v_{\ln \sigma^2}, \{m_{\xi_l}\}_{l=1}^q, v_\xi) | \{(\ln \sigma_j^2 - m_{\ln \sigma^2})/\sqrt{v_{\ln \sigma^2}}, (\ln \kappa_j^2 - m_{\ln \kappa^2})/\sqrt{v_{\ln \kappa^2}}, \{(\xi_{j,l} - m_{\xi_l})/\sqrt{v_\xi}\}_{l=1}^q\}_{j=1}^n$: $(5+q)$-dimensional RW Metropolis step with prior (13), (18), (23)-(24) and likelihood computed from Kalman filter from same SSS as in Step 2 of Model IV (except that the measurement equation now has variance $\omega^2 \sigma_{j,t}^2 S_{j,t} + \omega^2 \kappa_j^2 S_{j,t}^o$). Update $\{\sigma_j, \kappa_j, \{\xi_{j,l}\}_{l=1}^q\}_{j=1}^n$ if accepted.

(b) $\{\ln\sigma_j^2, \ln\kappa_j^2, \{\xi_{j,l}\}_{l=1}^q\}_{j=1}^n | m_{\ln\kappa^2}, v_{\ln\kappa^2}, m_{\ln\sigma^2}, v_{\ln\sigma^2}, \{m_{\xi_l}\}_{l=1}^q, v_\xi$: Looping over j, $q+2$-dimensional RW Metropolis step with prior (13), (18), (22) and likelihood computed from same Kalman filter as in Step 1a.

(c) $\{\mu_j, u_{j,-11:0}, \{\eta_{j,t}, \varepsilon_{j,t}\}_{t=1}^T\} | m_{\ln\kappa^2}, v_{\ln\kappa^2}, m_{\ln\sigma^2}, v_{\ln\sigma^2}, \{m_{\xi_l}\}_{l=1}^q, v_\xi, \{\sigma_j, \kappa_j, \{\xi_{j,l}\}_{l=1}^q\}_{j=1}^n$: Looping over j, Kalman smoother draw from same SSS as in Step 1a.

2. $\{m_{\ln(\nu-2)}, v_{\ln(\nu-2)}, m_{\ln\sigma^2}, v_{\ln\sigma^2}, \{m_{\xi_l}\}_{l=1}^q, v_\xi, \{\nu_j, \sigma_j^2, \{\xi_{j,l}\}_{l=1}^q, \{S_{j,t}\}_{t=1}^T\}_{j=1}^n\}$:

(a) $(m_{\ln(\nu-2)}, v_{\ln(\nu-2)}, m_{\ln\sigma^2}, v_{\ln\sigma^2}, \{m_{\xi_l}\}_{l=1}^q, v_\xi) | \{(\ln(\nu_j - 2) - m_{\ln(\nu-2)})/\sqrt{v_{\ln(\nu-2)}}, (\ln\sigma_j^2 - m_{\ln\sigma^2})/\sqrt{v_{\ln\sigma^2}}, \{(\xi_{j,l} - m_{\xi_l})/\sqrt{v_\xi}\}_{l=1}^q\}_{j=1}^n$: Six dimensional RW Metropolis step with priors (13), (15), (23)-(24) and likelihood computed from student-t density (14). Update $\{\nu_j, \sigma_j^2, \{\xi_{j,l}\}_{l=1}^q\}_{j=1}^n$ if accepted.

(b) $\{\ln(\nu_j - 2), \sigma_j^2, \{\xi_{j,l}\}_{l=1}^q\}_{j=1}^n | m_{\ln(\nu-2)}, v_{\ln(\nu-2)}, m_{\ln\sigma^2}, v_{\ln\sigma^2}, \{m_{\xi_l}\}_{l=1}^q, v_\xi$: Looping over j, three dimensional RW Metropolis step with priors (13), (15), (22) and likelihood computed from (14).

(c) $\{\{S_{j,t}\}_{t=1}^T\}_{j=1}^n | m_{\ln(\nu-2)}, v_{\ln(\nu-2)}, m_{\ln\sigma^2}, v_{\ln\sigma^2}, \{m_{\xi_l}\}_{l=1}^q, v_\xi, \{\nu_j, \sigma_j^2, \{\xi_{j,l}\}_{l=1}^q\}_{j=1}^n$: Looping over j, draw $S_{j,t}$ independently from $(\nu_j + \varepsilon_{j,t}^2)/\chi^2_{\nu_j+1}$, $t = 1, \ldots, T$.

3. Perform Steps 1 and 3 of Model IV and Steps 3-4 of Model II, except that in the SSS, the variance of the measurement equation is given by $\sigma_{j,t}^2 S_{j,t}$.

B.2.6 Model VI: Time varying conditional mean parameters

1. $\{\omega, m_{\ln\kappa^2}, m_{\ln\sigma^2}, m_{\ln\gamma_\mu^2}\}$: Conditioning on the current values of $\ln\omega^2 + m_{\ln\sigma^2}$, $\ln\omega^2 + m_{\ln\kappa^2}$ and $\ln\omega^2 + m_{\ln\gamma_\mu^2}$, draw $\ln\omega^2 | \ln\omega^2 + m_{\ln\sigma^2}, \ln\omega^2 + m_{\ln\kappa^2}, \ln\omega^2 + m_{\ln\gamma_\mu^2}$ from conjugate normal obtained from (13), (18), (29) and (7), then update $(m_{\ln\sigma^2}, m_{\ln\kappa^2}, m_{\ln\gamma_\mu^2})$ accordingly.

2. $\{m_{\ln\kappa^2}, v_{\ln\kappa^2}, m_{\ln\sigma^2}, v_{\ln\sigma^2}, \{m_{\xi_l}\}_{l=1}^q, v_\xi, m_{\ln\gamma_\mu^2}, v_{\ln\gamma_\mu^2}, \{\mu_j, u_{j,-11:0}, \sigma_j, \{\xi_{j,l}\}_{l=1}^q, \gamma_{\mu(j)}^2, \kappa_j, \{\eta_{j,t}, \varepsilon_{j,t}\}_{t=1}^T\}_{j=1}^n\}$:

(a) Draw $\{m_{\ln\kappa^2}, v_{\ln\kappa^2}, m_{\ln\sigma^2}, v_{\ln\sigma^2}, \{m_{\xi_l}\}_{l=1}^q, v_\xi, m_{\ln\gamma_\mu^2}, v_{\ln\gamma_\mu^2}\} |$ $\{(\ln\sigma_j^2 - m_{\ln\sigma^2})/\sqrt{v_{\ln\sigma^2}}, (\ln\kappa_j^2 - m_{\ln\kappa^2})/\sqrt{v_{\ln\kappa^2}}, \{(\xi_{j,l} - m_{\xi_l})/\sqrt{v_\xi}\}_{l=1}^q, (\ln\gamma_{\mu(j)}^2 - m_{\ln\gamma_\mu^2})/\sqrt{v_{\ln\gamma_\mu^2}}\}_{j=1}^n$: $(7 + q)$-dimensional RW Metropolis step with prior (13), (18), (23), (29) and likelihood computed from Kalman filter with state $(\mu_{j,t}, u_{j,t-1}, u_{j,t-2}, \ldots, u_{j,t-12})$, measurement equation (25), state evolution (26) and $\mu_{j,t}|\mu_{j,t-1} \sim \mathcal{N}(\mu_{j,t-1}, \omega^2\gamma^2_{\mu(j)})$, and initial state $\mu_{j,0} \sim \mathcal{N}(0, \infty)$ (approximated by using large but finite variance) and (9). Update $\{\sigma_j, \kappa_j, \{\xi_{j,l}\}_{l=1}^q, \gamma^2_{\mu(j)}\}_{j=1}^n$ if accepted.

(b) $\{\ln\sigma_j^2, \ln\kappa_j^2, \{\xi_{j,l}\}_{l=1}^q, \gamma^2_{\mu(j)}\}_{j=1}^n | m_{\ln\kappa^2}, v_{\ln\kappa^2}, m_{\ln\sigma^2}, v_{\ln\sigma^2}, \{m_{\xi_l}\}_{l=1}^q, v_\xi, m_{\ln\gamma^2_\mu}, v_{\ln\gamma^2_\mu}$: Looping over j, $q+3$ dimensional RW Metropolis step with prior (15), (18), (22), (29) and likelihood computed from same Kalman filter as in Step 2a.

(c) $\{\mu_j, u_{j,-11:0}, \{\eta_{j,t}, \varepsilon_{j,t}\}_{t=1}^T\} | m_{\ln\kappa^2}, v_{\ln\kappa^2}, m_{\ln\sigma^2}, v_{\ln\sigma^2}, \{m_{\xi_l}\}_{l=1}^q, v_\xi, m_{\ln\gamma_\mu}, v_{\ln\gamma_\mu}$, $\{\sigma_j, \kappa_j, \{\xi_{j,l}\}_{l=1}^q, \gamma^2_{\mu(j)}\}_{j=1}^n$: Looping over j, draw from Kalman smoother from same SSS as in Step 2a.

3. $\{\{m_{\phi_l}, v_{\phi_l}, m_{\ln(\gamma^2_{\phi(l)})}, v_{\ln(\gamma^2_{\phi(l)})}\}_{l=1}^{12}, \{\{\gamma^2_{\phi(j,l)}\}_{l=1}^{12}, \{\{\phi_{j,l,t}\}_{l=1}^{12}, \varepsilon_{j,t}\}_{t=1}^T\}_{j=1}^n\}$:

   (a) Draw $\{m_{\phi_l}, v_{\phi_l}, m_{\ln(\gamma^2_{\phi(l)})}, v_{\ln(\gamma^2_{\phi(l)})}\}_{l=1}^{12} | \{\{(\phi_{j,l,1} - m_{\phi_l})/\sqrt{v_{\phi_l}}\}_{l=1}^{12}, \{(\ln(\gamma^2_{\phi(j,l)}) - m_{\ln(\gamma^2_{\phi(l)})})/\sqrt{v_{\ln(\gamma^2_{\phi(l)})}}\}_{l=1}^{12}\}_{j=1}^n$: 48-dimensional RW Metropolis step with prior (12), (30) and likelihood computed from (9) and Kalman filter with state evolution $\phi_{j,t,l}|\phi_{j,t-1,l} \sim \mathcal{N}(\phi_{j,t-1,l}, \gamma^2_{\phi(j,l)})$, measurement equation (26) and initial state $\phi_{j,1,l}$, $l = 1, \ldots, 12$. Update $\{\{\phi_{j,l,1}, \gamma^2_{\phi(j,l)}\}_{l=1}^{12}\}_{j=1}^n$ if accepted.

   (b) $\{\ln\gamma^2_{\phi(j,l)}\}_{l=1}^{12} | \{m_{\phi_l}, v_{\phi_l}, m_{\ln(\gamma^2_{\phi(l)})}, v_{\ln(\gamma^2_{\phi(l)})}, \phi_{j,l,1}\}_{l=1}^{12}$ looping over j: 12-dimensional RW Metropolis step with prior (30) and likelihood computed from same SSS as in Step 3a.

   (c) $\{\{\phi_{j,l,t}\}_{l=1}^{12}, \varepsilon_{j,t}\}_{t=1}^T | \{m_{\phi_l}, v_{\phi_l}, m_{\ln(\gamma^2_{\phi(l)})}, v_{\ln(\gamma^2_{\phi(l)})}, \phi_{j,l,1}, \ln\gamma^2_{\phi(j,l)}\}_{l=1}^{12}$ looping over j: Kalman smoother draw from same SSS as in Step 3a.

4. $\{\{\phi_{j,l,t}\}_{l=1}^{12}, \varepsilon_{j,t}\}_{t=1}^T$ looping over j: Metropolis-Hastings step with proposal generated from Kalman smoother draw from SSS of Step 3a, except that initial state is $\phi_{j,1} \sim \mathcal{N}((m_{\phi_1}, \ldots, m_{\phi_{12}}), \text{diag}(v_{\phi_1}, \ldots, v_{\phi_{12}}))$. The proposal ϕ_j^p is accepted over the current value ϕ_j^c with probability $1 \wedge \frac{L_9(\phi_j^p)}{L_{??}(\phi_j^c)}$, where $L_9(\phi_j)$ is likelihood of (9).

5. $\{\{m_{\phi_l}\}_{l=1}^{12}, \{\{\phi_{j,l,t}\}_{l=1}^{12}, \varepsilon_{j,t}\}_{t=1}^T\}_{j=1}^n$: Conditional on $m_\phi = (m_{\phi_1}, \ldots, m_{\phi_{12}})'$, ignoring the likelihood contribution from (9), and integrating out $\{\{\phi_{j,l,t}\}_{l=1}^{12}\}_{t=2}^T$, the log-likelihood for $\{y_{j,t}\}_{t=1}^T$ is quadratic in $\phi_{j,1} = (\phi_{j,1,1}, \ldots, \phi_{j,12,1})'$ with mean $\tilde{\phi}_{j,1}$ and variance $\tilde{P}_j$

that could be computed by the Kalman smoother by initializing the SSS of Step 3a with a diffuse initial state. Given the Gaussian prior (12) $m_\phi \sim \mathcal{N}(0, \Omega_0)$ with $\Omega_0 = \text{diag}(v_{m_{\phi_1}}, \ldots, v_{m_{\phi_{12}}})$ and $\phi_j | m_\phi \sim \mathcal{N}(m_\phi, \Omega_1)$ with $\Omega_1 = \text{diag}(v_{\phi_1}, \ldots, v_{\phi_{12}})$, we could thus generate a Metropolis-Hastings proposal for m_ϕ and $\{\phi_{j,1}\}_{j=1}^n$ using the algorithm in Section B.1.2, that is $m_\phi \sim \mathcal{N}(V_0 \sum_{j=1}^n (\Omega_1 + \tilde{P}_j)^{-1} \tilde{\phi}_{j,1}, V_0)$ with $V_0^{-1} = \Omega_0^{-1} + \sum_{j=1}^n (\Omega_1 + \tilde{P}_j)^{-1}$ and $\phi_{j,1} | m_\phi \sim \mathcal{N}(V_j \tilde{P}_j^{-1} (\tilde{\phi}_{j,1} - m_\phi) + m_\phi, V_j)$ with $V_j^{-1} = \Omega_1^{-1} + \tilde{P}_j^{-1}$. Furthermore, this proposal can be extended to a proposal for $\{\{\{\phi_{j,l,t}\}_{l=1}^{12}\}_{t=1}^T\}_{j=1}^n$ by taking draws from $\{\{\phi_{j,l,t}\}_{l=1}^{12}\}_{t=2}^T$ for $j = 1, \ldots, n$ via the SSS described in Step 3c. The proposal would then be accepted with probability $(1 \wedge \frac{\prod_{j=1}^n L_9(\{\phi_{j,l,1}^p\}_{l=1}^{12})}{\prod_{j=1}^n L_9(\{\phi_{j,l,1}^c\}_{l=1}^{12})})$ in obvious notation.

The implementation differs from this conceptually straightforward approach, as the Kalman smoother with diffuse initial state is numerically unstable. Instead, we apply Kalman smoothers with initial state $\phi_{j,1} \sim \mathcal{N}(0, \Omega_1)$ to obtain the smoothed state $\hat{\phi}_{j,1}$ with smoothed covariance matrix $\hat{P}_j$. From Section B.1.2, $\hat{P}_j = (\Omega_1^{-1} + \tilde{P}_j^{-1})^{-1} = V_j$ and $\hat{\phi}_{j,1} = \hat{P}_j \tilde{P}_j^{-1} \tilde{\phi}_{j,1}$. Thus $(\Omega_1 + \tilde{P}_j)^{-1} \tilde{\phi}_{j,1} = \Omega_1^{-1} \hat{\phi}_{j,1}$ and, applying the Woodbury matrix identity, $(\Omega_1 + \tilde{P}_j)^{-1} = \Omega_1^{-1} - \Omega_1^{-1} \hat{P}_j \Omega_1^{-1}$, so the proposal for m_ϕ becomes $m_\phi \sim \mathcal{N}(V_0 \sum_{j=1}^n \Omega_1^{-1} \hat{\phi}_{j,1}, V_0)$ with $V_0^{-1} = \Omega_0^{-1} + \sum_{j=1}^n (\Omega_1^{-1} - \Omega_1^{-1} \hat{P}_j \Omega_1^{-1})$. Finally, to draw the proposals for $\phi_{j,1} | m_\phi$, we exploit that $\hat{P}_j \tilde{P}_j^{-1} (\tilde{\phi}_{j,1} - m_\phi) = \hat{\phi}_{j,1} - \hat{P}_j \tilde{P}_j^{-1} m_\phi = \hat{\phi}_{j,1} - m_\phi + \hat{P}_j \Omega_1^{-1} m_\phi$, where the last equality uses again the Woodbury identity, so $\phi_{j,1} | m_\phi \sim \mathcal{N}(\hat{\phi}_{j,1} + \hat{P}_j \Omega_1^{-1} m_\phi, \hat{P}_j)$.

This step is repeated 3 times for $n \leq 20$, and 5 times for $n > 20$. This is computationally efficient, since acceptance is rare, and most of the computational effort is in the calculation of the Kalman smoothers and the subsequent matrix manipulations, which only need to be performed once.

[This step is not needed given Steps 3-4, but it improves mixing.]

6. Perform Step 2 of Model V and Step 3 of Model IV.

B.2.7 RTS Model

1. $\{\omega, m_{\ln \kappa^2}, m_{\ln \sigma^2}, m_{\ln \gamma_\mu^2}, \ln(\sigma_{n+1}^2), \ln \kappa_{n+1}^2, \ln \gamma_{\mu(n+1)}^2\}$: Conditioning on the current values of $\ln \omega^2 + m_{\ln \sigma^2}$, $\ln \omega^2 + m_{\ln \kappa^2}$, $\ln \omega^2 + m_{\ln \gamma_\mu^2}$ and $\ln \omega^2 + \ln(\sigma_{n+1}^2)$, draw $\ln \omega^2 | \ln \omega^2 + m_{\ln \sigma^2}, \ln \omega^2 + m_{\ln \kappa^2}, \ln \omega^2 + m_{\ln \gamma_\mu^2}, \ln \omega^2 + \ln(\sigma_{n+1}^2)$ from conjugate

normal obtained from (34), (13), (18), (29), (7), (35) and (36), then update $(m_{\ln\sigma^2}, m_{\ln\kappa^2}, m_{\ln\gamma_\mu^2}, \ln(\sigma^2_{n+1}), \ln\kappa^2_{n+1}, \ln\gamma^2_{u(n+1)})$ accordingly.

2. $\{\kappa^2_{n+1}, \{\{\eta_{j,t}\}_{t=1}^T\}_{j=1}^{n+1}\}$: Conditioning on the current values of $\{\{\kappa_{n+1}\eta_{n+1,t} + \kappa_{j,t}\eta_{j,t}\}_{t=1}^T\}_{j=1}^n$

   (a) Draw $\kappa^2_{n+1}|\{\{\kappa_{n+1}\eta_{n+1,t}+\kappa_{j,t}\eta_{j,t}\}_{t=1}^T\}_{j=1}^n$ by RW Metropolis step based on likelihood of $\{\{\kappa_{n+1}\eta_{n+1,t} + \kappa_{j,t}\eta_{j,t}\}_{t=1}^T\}_{j=1}^n$ induced by $\eta_{j,t} \sim \mathcal{N}(0, S^o_{j,t})$ independently across $j = 1, \ldots, n+1$ and $t = 1, \ldots, T$.

   (b) $\{\eta_{j,t}\}_{j=1}^{n+1}|\kappa^2_{n+1}, \{\kappa_{n+1}\eta_{n+1,t} + \kappa_{j,t}\eta_{j,t}\}_{j=1}^n$ looping over t: Draw $\eta_{n+1,t}$ from conjugate Gaussian posterior implied by the likelihood of Step 2a, and update $\{\eta_{j,t}\}_{j=1}^n$ accordingly.

3. $\{\gamma^2_{\mu(n+1)}, \{\mu_{n+1,t}\}_{t=1}^T, \{\{\mu_{j,t}\}_{t=1}^T\}_{j=1}^n\}$:

   (a) Draw $\gamma^2_{\mu(n+1)}|\{\{\mu_{n+1,t} + \mu_{j,t}\}_{t=1}^T\}_{j=1}^n$ by RW Metropolis step based on likelihood of $\{\{\Delta\mu_{n+1,t} + \Delta\mu_{j,t}\}_{t=1}^T\}_{j=1}^n$ induced by $\Delta\mu_{j,t} \sim \mathcal{N}(0, \gamma^2_{\mu(j)})$ independently across $j = 1, \ldots, n+1$ and $t = 2, \ldots, T$.

   (b) $\mu_{n+1,t}|\gamma_{\mu(n+1)}, \{\mu_{n+1,t}+\mu_{j,t}\}_{j=1}^n$ looping over t: Draw $\mu_{n+1,t}$ from conjugate Gaussian posterior implied by the likelihood of Step 3a, and update $\{\mu_{j,t}\}_{j=1}^n$ accordingly.

4. $\{\{m_{\lambda_l}, v_{\lambda_l}, m_{\ln(\gamma^2_{\lambda(l)})}, v_{\ln(\gamma^2_{\lambda(l)})}\}_{l=0}^5, \{\{\lambda_{j,l,t}, \gamma^2_{\lambda(j,l)}\}_{l=0}^5, \{\varepsilon_{j,t}\}_{t=1}^T\}_{j=1}^n\}$:

   (a) Draw $\{m_{\lambda_l}, v_{\lambda_l}, m_{\ln(\gamma^2_{\lambda(l)})}, v_{\ln(\gamma^2_{\lambda(l)})}\}_{l=0}^5|\{\{(\ln(\gamma^2_{\lambda(j,l)}) - m_{\ln(\gamma^2_{\lambda(l)})})/\sqrt{v_{\ln(\gamma^2_{\lambda(l)})}}\}_{l=0}^5\}_{j=1}^n$: 24-dimensional RW Metropolis step with prior (39) and likelihood computed from n Kalman filters with state $(u_{j,t-1}, \ldots, u_{j,t-12}, \lambda_{j,0,t}, \ldots \lambda_{j,5,t})$, evolution $\lambda_{j,t}|\lambda_{j,t-1} \sim \mathcal{N}(\lambda_{j,t-1}, \text{diag}(\gamma^2_{\lambda(0)}, \ldots, \gamma^2_{\lambda(5)}))$, measurement $\sum_{l=1}^{12}\phi_{j,t,l}u_{j,t-l} + \sum_{l=0}^5 \lambda_{j,l,t}u_{n+1,t-l} + \epsilon_{j,t}$ and initial state $(u_{j,0}, \ldots, u_{j,-11})$, $\lambda_{j,1} \sim \mathcal{N}((m_{\lambda(0)}, \ldots, m_{\lambda(5)})', \text{diag}(v_{\lambda(0)}, \ldots, v_{\lambda(5)}))$. Update $\{\{\gamma^2_{\lambda(j,l)}\}_{l=0}^5\}_{j=1}^n$ if accepted. [Only the last 23 components actually move, since m_{λ_0} is fixed at 1 and $v_{\lambda_0} = 0$.]

   (b) $\{\ln\gamma^2_{\lambda(j,l)}\}_{l=0}^5|\{m_{\lambda_l}, v_{\lambda_l}, m_{\ln(\gamma^2_{\lambda(l)})}, v_{\ln(\gamma^2_{\lambda(l)})}\}_{l=0}^5$ looping over j: 6-dimensional RW Metropolis step with prior (39) and likelihood computed same SSS as in Step 3a.

(c) $\{\{\lambda_{j,l,t}\}_{l=0}^{5}, \varepsilon_{j,t}\}_{t=1}^{T} | \{m_{\lambda_l}, v_{\lambda_l}, m_{\ln(\gamma^2_{\lambda(l)})}, v_{\ln(\gamma^2_{\lambda(l)})}, \lambda_{j,l,1}, \ln\gamma^2_{\lambda(j,l)}\}_{l=0}^{5}$ looping over j: Kalman smoother draw from same SSS as in Step 3.

5. $\{m_{\ln\kappa^2}, v_{\ln\kappa^2}, m_{\ln\sigma^2}, v_{\ln\sigma^2}, \{m_{\xi_l}\}_{l=1}^{q}, v_\xi, m_{\ln\gamma^2_\mu}, v_{\ln\gamma^2_\mu}, \{\mu_j, u_{j,-11:0}, \sigma_j, \{\xi_{j,l}\}_{l=1}^{q}, \gamma^2_{\mu(j)}, \kappa_j, v_{j,t}, \{\eta_{j,t}, \varepsilon_{j,t}\}_{t=1}^{T}\}_{j=1}^{n}\}$: Same as Step 2 of Model VI, except that in Step 2a, there is an additional contribution to the posterior from $\xi_{l,n+1} \sim \mathcal{N}(m_{\xi_l}, v_\xi)$, $l = 1, \ldots, q$.

6. $\{\{m_{\phi_l}, v_{\phi_l}, m_{\ln(\gamma^2_{\phi(l)})}, v_{\ln(\gamma^2_{\phi(l)})}\}_{l=1}^{12}, \{\{\gamma^2_{\phi(j,l)}\}_{l=1}^{12}, \{\{\phi_{j,l,t}\}_{l=1}^{12}, \varepsilon_{j,t}\}_{t=1}^{T}\}_{j=1}^{n+1}\}$: Same as Step 3 of Model VI, except that there are now $n+1$ processes.

7. $\{\sigma^2_{n+1}, \{\xi_{n+1,l}\}_{l=1}^{q}, \{u_{j,-11:0}, \{\varepsilon_{j,t}\}_{t=1}^{T}\}_{j=1}^{n+1}\}$: Conditional on $\{\sigma^2_{n+1}, \{\xi_{n+1,l}\}_{l=1}^{q}\}$, (26) and (9) imply an a priori mean-zero Gaussian distribution for $(u_{n+1,-11}, \ldots, u_{n+1,T})'$ with a band diagonal precision matrix with bandwidth 13. Furthermore, with $u^0_{j,t} = (y_{j,t} - \mu_{j,t} - \mu_{n+1,t} - \omega(o_{j,t} + o_{n+1,t}))/\omega = u_{j,t} + \sum_{l=0}^{5} \lambda_{j,l,t} u_{n+1,t-l}$, $u^0_{j,t} - \sum_{l=1}^{12} \phi_{j,l,t} u^0_{j,t-l} = \sum_{l=0}^{5} \lambda_{j,l,t} u_{n+1,t-l} - \sum_{l=1}^{12} \phi_{j,l,t} \sum_{\ell=0}^{5} \lambda_{j,\ell,t-l} u_{n+1,t-\ell-l} + \epsilon_{j,t}$ for $t = 1, \ldots, T$, $j = 1, \ldots, n$ are independent Gaussian measurements of a linear combination of 18 consecutive elements of $u_{n+1,t}$. The precision matrix of the Gaussian posterior of $\{u_{n+1,t}\}_{t=-11}^{T} | \{\{u^0_{j,t}\}_{t=-11}^{T}\}_{j=1}^{n}$ is thus band diagonal with bandwidth 18, and specialized linear algebra routines that exploit this band diagonal structure can be employed to efficiently generate a draw from the posterior. See for details. The same matrix calculations also enable us to efficiently compute the likelihood of $\{\{u^0_{j,t}\}_{t=-11}^{T}\}_{j=1}^{n}$ as a function of $\{\ln\sigma^2_{n+1}, \{\xi_{n+1,l}\}_{l=1}^{q}\}$ without conditioning on $\{u_{n+1,t}\}_{t=-11}^{T}$, which we employ in a RW Metropolis step that precedes the draw from $\{u_{n+1,t}\}_{t=-11}^{T} | \{\{u^0_{j,t}\}_{t=-11}^{T}\}_{j=1}^{n}$.

8. $\{\kappa^2_{n+1}, \{\eta_{n+1,t}\}_{t=1}^{T}, \{\{\varepsilon_{j,t}\}_{t=1}^{T}\}_{j=1}^{n}\}$: With $u^0_{j,t} = (y_{j,t} - \mu_{j,t} - \mu_{n+1,t} - \omega(c_{j,t} + o_{j,t}))/\omega = u_{j,t} + o_{n+1,t}$, $u^0_{j,t} - \sum_{l=1}^{12} \phi_{j,l,t} u^0_{j,t-l} = o_{n+1,t} - \sum_{l=1}^{12} \phi_{j,l,t} o_{n+1,t-l} + \epsilon_{j,t}$ for $t = 1, \ldots, T$, $j = 1, \ldots, n$ with $o_{n+1,t} = \kappa_{n+1}\eta_{n+1,t}$ are independent Gaussian measurements of a linear combination of 13 consecutive elements of $o_{n+1,t}$. Furthermore, conditional on κ^2_{n+1}, $\{o_{n+1,t}\}_{t=1}^{T}$ has a diagonal prior precision matrix. We can thus use the same approach as in Step 7 to update κ^2_{n+1} and $\{\eta_{n+1,t}\}_{t=1}^{T}$ given $\{\{u^0_{j,t}\}_{t=1}^{T}\}_{j=1}^{n}$. [Step 8 is not necessary given Step 2, but it improves mixing.]

9. $\{\gamma^2_{\mu(n+1)}, \{\mu_{n+1,t}\}_{t=2}^{T}, \{\{\varepsilon_{j,t}\}_{t=2}^{T}\}_{j=1}^{n}\}$: With $u^0_{j,t} = (y_{j,t} - \mu_{j,t} - \omega(o_{j,t} + o_{n+1,t} + c_{j,t}))/\omega = u_{j,t} + \mu_{n+1,t}/\omega$, $u^0_{j,t} - \sum_{l=1}^{12} \phi_{j,l,t} u^0_{j,t-l} = \mu_{n+1,t}/\omega - \sum_{l=1}^{12} \phi_{j,l,t} \mu_{n+1,t-l}/\omega + \epsilon_{j,t}$ for $t =$

$2, \ldots, T$, $j = 1, \ldots, n$ are independent Gaussian measurements of a linear combination of 13 consecutive elements of $\mu_{n+1,t}/\omega$. Furthermore, conditional on $\gamma^2_{\mu(n+1)}$, $\{\mu_{n+1,t}\}_{t=2}^T$ has a band diagonal precision matrix with bandwidth 2. We can thus use the same approach as in Step 7 to update $\gamma^2_{\mu(n+1)}$ and $\{\mu_{n+1,t}\}_{t=2}^T$ given $\{\{u^0_{j,t}\}_{t=2}^T\}_{j=1}^n$. [Step 9 is not necessary given Step 3, but it improves mixing.]

10. $\{m_{\ln(\nu-2)}, v_{\ln(\nu-2)}, m_{\ln\sigma^2}, v_{\ln\sigma^2}, \{m_{\xi_l}\}_{l=1}^q, v_\xi, \{\nu_j, \sigma^2_j, \{\xi_{j,l}\}_{l=1}^q, \{S_{j,t}\}_{t=1}^T\}_{j=1}^{n+1}\}$: Same as Step 2 of Model V, except that there are now $n+1$ processes, and in Step 2a, σ^2_{n+1} does not follow the hierarchical prior.

11. $\{m_{\ln(\nu^o-2)}, v_{\ln(\nu^o-2)}, m_{\ln\kappa^2}, v_{\ln\kappa^2}, \{\nu^o_j, \kappa_j, \{S^o_{j,t}\}_{t=1}^T\}_{j=1}^{n+1}\}$: Same as Step 3 of Model IV, except that there are now $n+1$ processes, and in Step 3a, κ^2_{n+1} does not follow the hierarchical prior.

B.3 Bayes Factors

The Bayes factors were obtained by using the bridge sampling approach of Meng and Wong (1996): For two models A and B with the same parameter space Θ, priors π_A and π_B and likelihood $f_A(y|\theta)$ and $f_B(y|\theta)$, respectively, define $\mathrm{LR}_{A/B}(\theta) = \frac{f_A(y|\theta)\pi_A(\theta)}{f_B(y|\theta)\pi_B(\theta)}$ and $\mathrm{LR}_{B/A}(\theta) = \frac{f_B(y|\theta)\pi_B(\theta)}{f_A(y|\theta)\pi_A(\theta)}$. Then the Bayes factor BF can be written as

$$\mathrm{BF} = \frac{\int f_A(y|\theta)\pi_A(\theta)d\theta}{\int f_B(y|\theta)\pi_B(\theta)d\theta} = \frac{\mathbb{E}_B[\frac{\mathrm{LR}_{A/B}(\theta)}{\mathrm{LR}_{A/B}(\theta)+\mathrm{BF}}]}{\mathbb{E}_A[\frac{\mathrm{LR}_{B/A}(\theta)}{1+\mathrm{BF}\cdot\mathrm{LR}_{B/A}(\theta)}]}$$

where $E_A[\cdot]$ and $E_B[\cdot]$ are expectations with respect to the posterior of θ under model A and B, respectively. The Bayes factors can then be obtained by replacing the posterior expectations by averages from MCMC output, and by iterating the above to convergence.

It is useful to "bridge" the gap between the baseline model A_0 and a fairly distinct alternative model A_k by intermediate models via the identity $\mathrm{BF}_{A_0/A_k} = \prod_{i=1}^k \mathrm{BF}_{A_{i-1}/A_i}$ in obvious notation, as this improves numerical stability of the estimate. In our application, we set $k = 2$, so there is one intermediate model.

References

Chan, J., and I. Jeliazkov (2009): "Efficient Simulation and Integrated Likelihood Estimation in State Space Models," *International Journal of Mathematical Modeling and Numerical Optimization*, 1, 101–120.

Durbin, J., and S. J. Koopman (2002): "A Simple and Effcient Simulation Smoother for State Space Time Series Analysis," *Biometrika*, 89(3), 603–615.

Meng, X. L., and W. H. Wong (1996): "Simulating Ratios of Normalizing Constants via a Simple Identity: a Theoretical Exploration," *Stastistica Sinica*, 6, 831–860.

Müller, U. K., and M. W. Watson (2020): "Low-Frequency Analysis of Economic Time Series," *Chapter prepared for the Handbook of Econometrics.*

Stock, J. H., and M. W. Watson (2016): "Core Inflation and Trend Inflation," *Review of Economics and Statistics*, 98(4), 770–784.