Online Appendix: Sequential Monte Carlo with Model Tempering

Marko Mlikota and Frank Schorfheide

This Appendix consists of the following sections:

A. Computational Details
B. Illustration 2: VAR with Stochastic Volatility
C. Illustration 3: A Nonlinear DSGE Model

A Computational Details

The presentations of the mutation algorithm in Section A.1 and the BSPF in in Section A.2 are based on Herbst and Schorfheide (2015).

A.1 SMC Particle Mutation

Algorithm 2 (Particle Mutation).
In Step 2(c) in iteration n of Algorithm 1:

1. *Compute an importance sampling approximation $\tilde{\Sigma}_n$ of $\mathbb{V}_{\pi_n}[\theta]$ based on the particles $\{\theta^i_{n-1}, \tilde{W}^i_n\}_{i=1}^N$.*
2. *Compute the average empirical rejection rate $\hat{R}_{n-1}(\hat{\zeta}_{n-1})$, based on the Mutation step in iteration $n-1$. The average is computed across the N_{blocks} blocks.*
3. *Let $\hat{c}_1 = c^*$ and for $n > 2$ adjust the scaling factor according to*
$$\hat{c}_n = \hat{c}_{n-1} f\big(1 - \hat{R}_{n-1}(\hat{\zeta}_{n-1})\big),$$
where
$$f(x) = 0.95 + 0.10 \frac{e^{16(x-0.25)}}{1 + e^{16(x-0.25)}}.$$
4. *Define $\hat{\zeta}_n = \big[\hat{c}_n, vech(\tilde{\Sigma}_n)'\big]'$.*
5. *For each particle i, run N_{MH} steps of a Random Walk Metropolis-Hastings Algorithm using the proposal density*
$$\vartheta^{i,m}_n | \hat{\zeta}_n \sim N\left(\theta^{i,m-1}_n, \hat{c}^2_n \tilde{\Sigma}_n\right). \tag{A.1}$$

A.2 (Particle) Filtering

We use a bootstrap particle filter (BSPF) to approximate the likelihood function in the model with stochastic volatility. In the description of the filter we denote the latent state by s_t.

Algorithm 3 (Bootstrap Particle Filter).

1. **Initialization.** *Draw the initial particles from the distribution $s^j_0 \overset{iid}{\sim} p(s_0|\theta)$ and set $W^j_0 = 1$, $j = 1, \ldots, M$.*
2. **Recursion.** *For $t = 1, \ldots, T$:*
   (a) **Forecasting s_t.** *Draw $\tilde{s}^j_t$ from the state-transition density $p(\tilde{s}_t | s^j_{t-1}, \theta)$.*
   (b) **Forecasting y_t.** *Define the incremental weights*
$$\tilde{w}^j_t = p(y_t | \tilde{s}^j_t, Y_{1:t-1}, \theta) \tag{A.2}$$
   The predictive density $p(y_t|Y_{1:t-1}, \theta)$ can be approximated by
$$\hat{p}(y_t|Y_{1:t-1}, \theta) = \frac{1}{M} \sum_{j=1}^{M} \tilde{w}^j_t W^j_{t-1}. \tag{A.3}$$

(c) *Define the normalized weights*

$$\tilde{W}_t^j = \tilde{w}_t^j W_{t-1}^j \Big/ \frac{1}{M}\sum_{j=1}^{M} \tilde{w}_t^j W_{t-1}^j. \tag{A.4}$$

(d) **Selection.** *Resample the particles, for instance, via multinomial resampling. Let* $\{s_t^j\}_{j=1}^M$ *denote* M *iid draws from a multinomial distribution characterized by support points and weights* $\{\tilde{s}_t^j, \tilde{W}_t^j\}$ *and set* $W_t^j = 1$ *for* $j =, 1\ldots, M$*. An approximation of* $\mathbb{E}[h(s_t)|Y_{1:t},\theta]$ *is given by* $\bar{h}_{t,M} = \frac{1}{M}\sum_{j=1}^M h(s_t^j)W_t^j$*.*

3. **Likelihood Approximation.** *The approximation of the log-likelihood function is given by*

$$\ln \hat{p}(Y_{1:T}|\theta) = \sum_{t=1}^{T} \ln\left(\frac{1}{M}\sum_{j=1}^{M} \tilde{w}_t^j W_{t-1}^j\right). \tag{A.5}$$

B Illustration 2: A VAR with Stochastic Volatility

B.1 Prior Specification

Prior for (Φ_1, Φ_2, Σ)**.** We use a Minnesota-type prior for the reduced-form VAR coefficients that appear in the homoskedastic version of the VAR in (21). The specification of the Minnesota prior follows Del Negro and Schorfheide (2012). The prior is indexed by hyperparameters λ_1, λ_2, and λ_3, and is implemented through dummy observations stacked into (Y^*, X^*). We use three sets of dummy observations, written as $Y_j^* = X_j^*\Phi + U_j$:

$$\begin{aligned}
\begin{bmatrix} \lambda_1 \underline{s}_1 & 0 \\ 0 & \lambda_1 \underline{s}_2 \end{bmatrix} &= \begin{bmatrix} \lambda_1 \underline{s}_1 & 0 & 0 \\ 0 & \lambda_1 \underline{s}_2 & 0 \end{bmatrix} \Phi + \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}, \\
\begin{bmatrix} \lambda_2 \underline{y}_1 & \lambda_2 \underline{y}_2 \end{bmatrix} &= \begin{bmatrix} \lambda_2 \underline{y}_1 & \lambda_2 \underline{y}_2 & \lambda_2 \end{bmatrix} \Phi + \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}, \\
\begin{bmatrix} \underline{s}_1 & 0 \\ 0 & \underline{s}_2 \end{bmatrix} &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Phi + \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix},
\end{aligned}$$

where $\underline{y}_i$ and $\underline{s}_i$ are the mean and standard deviation of y_i. The first set of dummy observations implies that the VAR coefficients are centered at univariate unit-root representations. The second set of dummy observations implies that if the lagged value y_{t-1} take the value $\underline{y}$, then the current value y_t will be close to $\underline{y}$. The third set of dummy observations induces a prior for the covariance matrix of u_t and is repeated λ_3 times. The dummy observations induce a conjugate MNIW prior for (Φ, Σ):

$$\Sigma \sim IW(\underline{S}, \underline{\nu}), \quad \Phi|\Sigma \sim MN(\underline{\mu}, \Sigma \otimes \underline{P}^{-1}),$$

with

$$\underline{\nu} = T^* - k, \quad \underline{S} = S^*, \quad \underline{\mu} = \Phi^*, \quad \underline{P} = X^{*\prime}X^*,$$

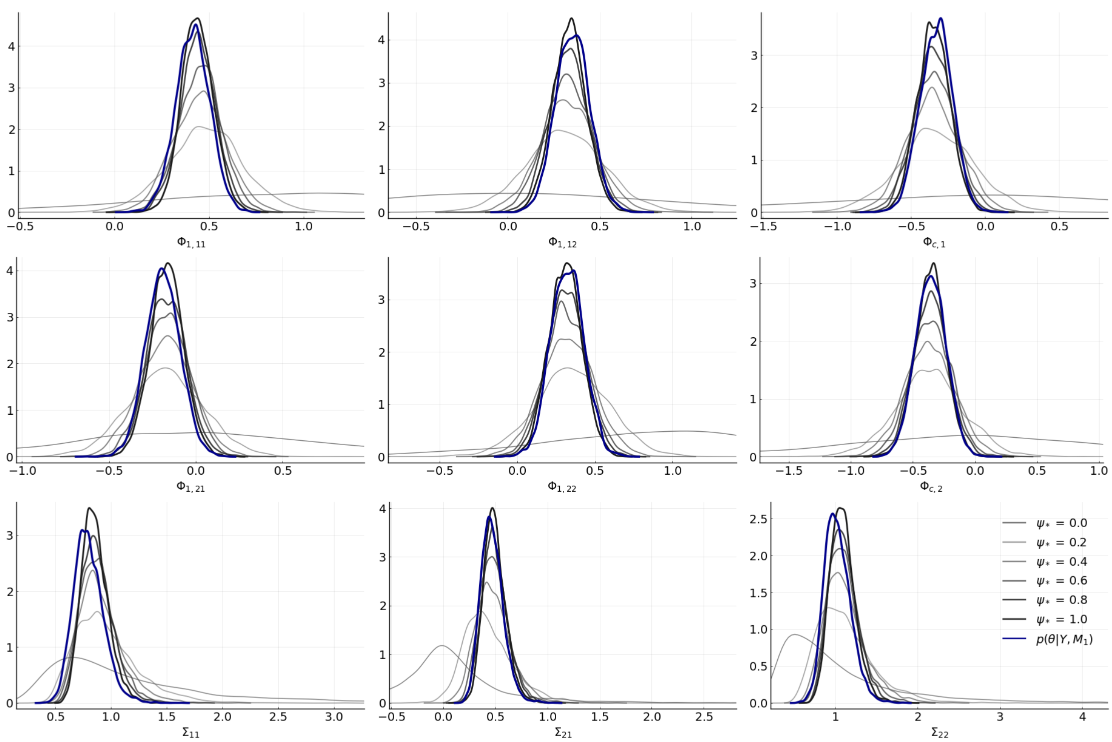
where $\Phi^* = (X^{*\prime}X^*)^{-1}X^{*\prime}Y^*$ and $S^* = (Y^* - X^*\Phi^*)'(Y^* - X^*\Phi^*)$. We set $\lambda_1 = 1$, $\lambda_2 = 1$, and $\lambda_3 = 3$.

Prior for ρ_i**.** The prior for each ρ_i is Uniform on $[0, 1]$.

Prior for ξ_i**.** The prior of ξ_i is specified as an inverse Gamma distribution. It is parameterized as scaled inverse χ^2 distribution with density $p(\xi^2|s^2,\nu) \propto (\xi^2)^{-\nu/2-1}\exp[-\nu s^2/(2\xi^2)]$, where $\sqrt{s^2}$ is 0.3 and ν is 2.0. The density of ξ_i is obtained by the change of variables $\xi = \sqrt{\xi^2}$.

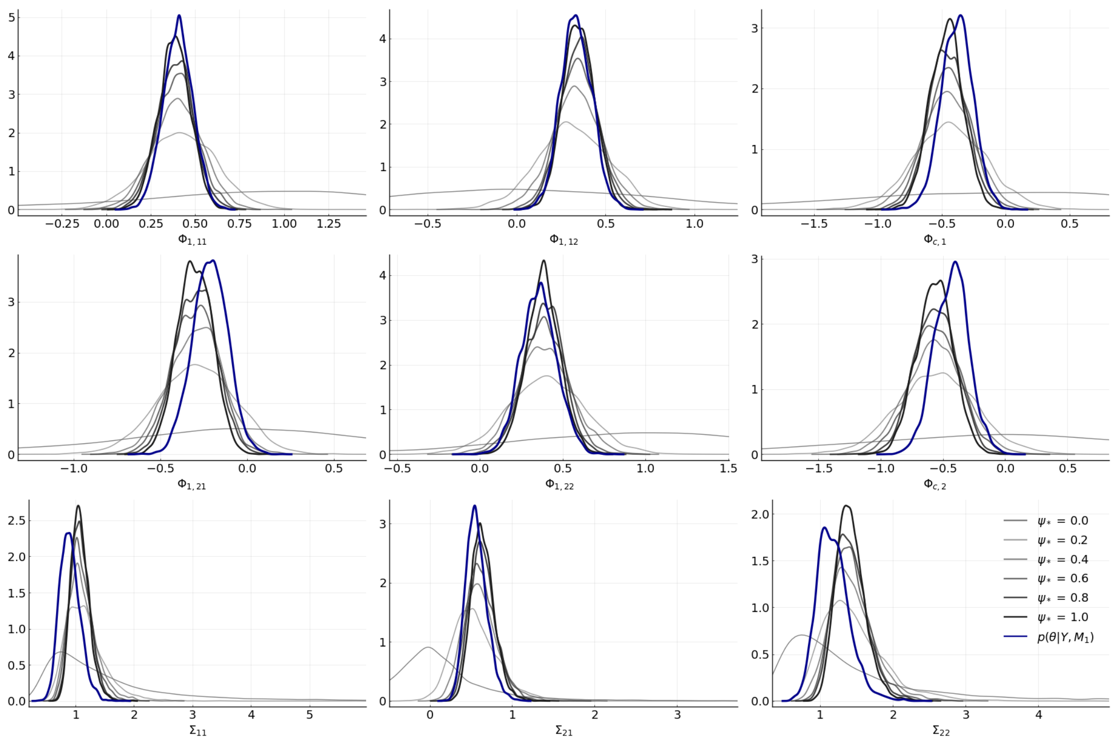
B.2 Further Results for the VAR-SV

Fig. A-1: VAR-SV: Target and Approximate Posterior Densities for DGP 1



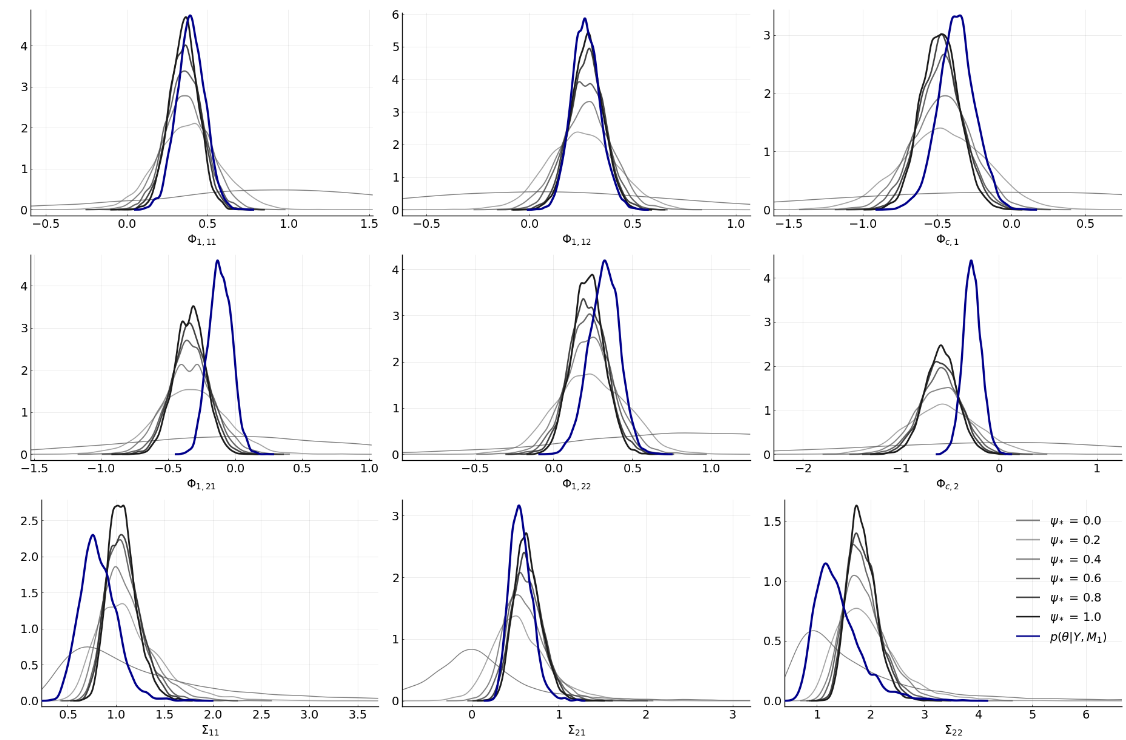
Notes: Each plot refers to a different parameter. The approximating posterior densities obtained from the tempered M_0 likelihood function for $\psi_* \in \{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$ are plotted in shades (the larger ψ_* the darker) of gray. The M_1 posterior is depicted in blue. The stochastic volatility parameters $\rho_i, \xi_i,\ i = 1, 2$ are not displayed because model M_0 is uninformative for them.

Fig. A-2: VAR-SV: Target and Approximate Posterior Densities for DGP 2



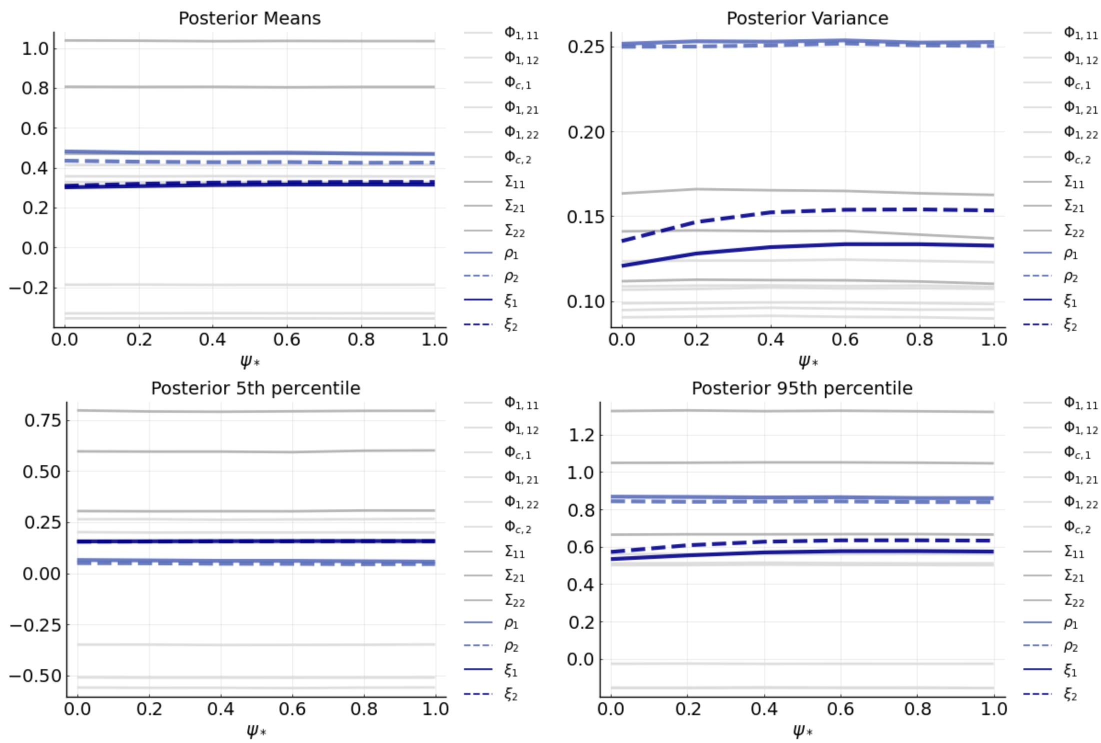
Notes: Each plot refers to a different parameter. The approximating posterior densities obtained from the tempered M_0 likelihood function for $\psi_* \in \{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$ are plotted in shades (the larger ψ_* the darker) of gray. The M_1 posterior is depicted in blue. The stochastic volatility parameters $\rho_i, \xi_i,\ i = 1, 2$ are not displayed because model M_0 is uninformative for them.

Fig. A-3: VAR-SV: Target and Approximate Posterior Densities for DGP 3



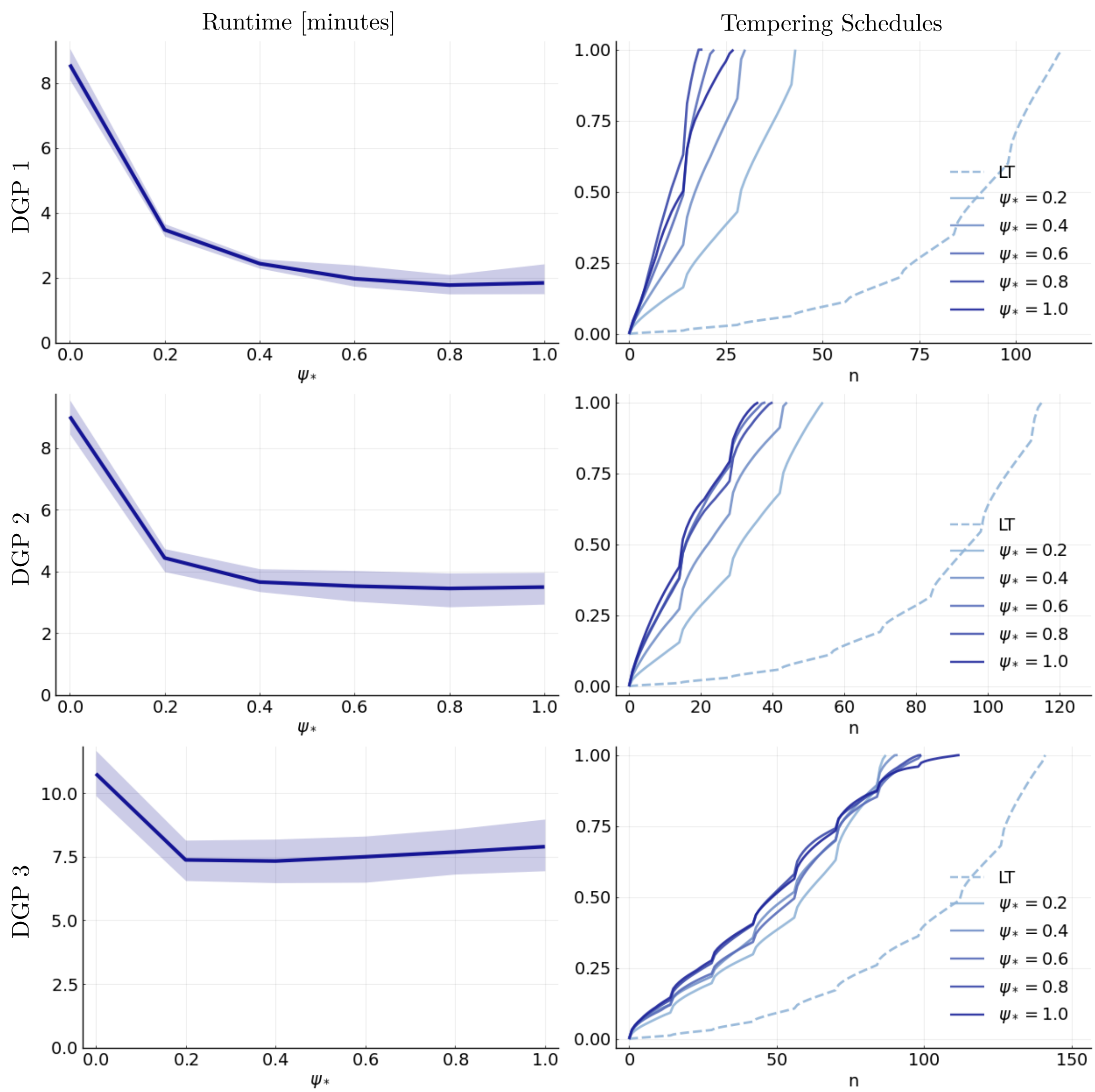
Notes: Each plot refers to a different parameter. The approximating posterior densities obtained from the tempered M_0 likelihood function for $\psi_* \in \{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$ are plotted in shades (the larger ψ_* the darker) of gray. The M_1 posterior is depicted in blue. The stochastic volatility parameters $\rho_i, \xi_i,\ i = 1, 2$ are not displayed because model M_0 is uninformative for them.

Fig. A-4: VAR-SV: Monte Carlo Approximations of Posterior Statistics for DGP 1



Notes: Each panel shows the Monte Carlo approximation of the respective posterior statistic as a function of the tempering parameter ψ_* for the approximating model. Depicted are means across $N_{run} = 200$ runs.

Fig. A-5: VAR-SV: Runtime and Tempering Schedule



Notes: The left panel shows the mean runtime and 90% confidence interval across $N_{run} = 200$ runs. The right panel illustrates the evolution of the tempering schedule by plotting the median value of the tempering parameter at each stage n.

C Illustration 3: A Nonlinear DSGE Model

C.1 Equilibrium Conditions, Steady State, and Log-linearization

We write the social planner's problem stated in the main text as

$$\begin{aligned} V(K,S) &= \max_{C,L,K'} u(B,C,L) + \beta \mathbb{E}_{S'|S}[V(K',S')] \\ &\text{s.t.} \quad C + I + K\Phi(K'/K) = Y, && \text{(A.6)} \\ &\qquad Y = f(Z,K,L), && \text{(A.7)} \\ &\qquad I = K' - (1-\delta)K\ . && \text{(A.8)} \end{aligned}$$

We use the following functional forms:

$$\begin{aligned} u(B,C,L) &= \frac{C^{1-\tau}-1}{1-\tau} - B\frac{L^{1+1/\nu}}{1+1/\nu}, \\ f(Z,K,L) &= ZK^{\alpha}L^{1-\alpha}, \\ \Phi(K'/K) &= \phi_1\left(\frac{\exp(-\phi_2(K'/K-1)) + \phi_2(K'/K-1) - 1}{\phi_2^2}\right). \end{aligned}$$

The exogenous processes evolve according to:

$$\begin{aligned} Z &= Z_* e^{\hat{z}}, \quad \hat{z}' = \rho_z \hat{z} + \sigma_z \varepsilon_z', \\ B &= B_* e^{\hat{b}}, \quad \hat{b}' = \rho_b \hat{b} + \sigma_b \varepsilon_b'\ . \end{aligned}$$

Throughout this section we use $f_i(\cdot)$ to denote the derivative of a function $f(\cdot)$ with respect to its i'th argument.

C.1.1 First-Order Conditions (FOCs)

Substitute (A.7) and (A.8) into (A.6) and then take FOCs with respect to L and K'. The FOC for L takes the form

$$u_2(B,C,L)f_3(Z,K,L) + u_3(B,C,L) = 0.$$

Using the functional forms, this leads to

$$(1-\alpha)\frac{Y}{L} = BC^{\tau}L^{1/\nu}. \tag{A.9}$$

Now write

$$C = ZK^{\alpha}L^{1-\alpha} - K' + (1-\delta)K - K\Phi(K'/K).$$

The FOC for K' takes the form:

$$-u_2(B,C,L)\big[1+\Phi_1(K'/K)\big] + \beta\mathbb{E}\left[V_1(K',S')\right] = 0.$$

Plugging in the expressions for $u_2(\cdot)$ and $V_1(\cdot)$ we obtain

$$\begin{aligned} &C^{-\tau}\left[1+\Phi_1(K'/K)\right] && \text{(A.10)} \\ &\quad = \beta\mathbb{E}\left[C'^{-\tau}\left(\alpha\frac{Y'}{K'} + 1 - \delta - \Phi(K''/K') + \Phi_1(K''/K')\frac{K''}{K'}\right)\right], \end{aligned}$$

where

$$\Phi_1(x) = \frac{\phi_1}{\phi_2}\left[1 - \exp\{-\phi_2(x-1)\}\right]. \tag{A.11}$$

C.1.2 Steady State

Rather than taking (Z_*, B_*) as given and solving for (Y_*, L_*) and the remaining steady states, we proceed in the other direction and solve for (Z_*, B_*) as a function of (Y_*, L_*). Notice that the adjustment costs are zero in steady state because $\Phi(1) = 0$. Moreover, $\Phi_1(1) = 0$. We deduce from (A.10) that

$$\frac{1}{\beta} = \alpha \frac{Y_*}{K_*} + (1-\delta),$$

which implies that

$$K_* = \frac{\alpha}{1/\beta - (1-\delta)} Y_*. \tag{A.12}$$

The capital accumulation equation implies that

$$I_* = \delta K_* = \frac{\alpha\delta}{1/\beta - (1-\delta)} Y_*. \tag{A.13}$$

The aggregate resource constraint implies that

$$C_* = Y_* - I_* = \left(1 - \frac{\alpha\delta}{1/\beta - (1-\delta)}\right) Y_*. \tag{A.14}$$

The production function can be solved for Z_*:

$$Z_* = \frac{Y_*}{K_*^{\alpha} L_*^{1-\alpha}} = \left(\frac{1/\beta - (1-\delta)}{\alpha}\right)^{\alpha} \left(\frac{Y_*}{L_*}\right)^{1-\alpha}. \tag{A.15}$$

Finally, we solve (A.9) for B to obtain B_*:

$$B_* = (1-\alpha)\frac{Y_*}{L_*} C_*^{-\tau} L_*^{-1/\nu}.$$

In the numerical illustration we set $Y_* = L_* = 1$.

C.1.3 Log-Linearization

Log-linearizing Equations (A.6), (A.7), (A.8), and (A.9) yields:

$$\hat{y} = \frac{C_*}{Y_*}\hat{c} + \frac{I_*}{Y_*}\hat{i} \tag{A.16}$$

$$\hat{y} = \hat{z} + \alpha\hat{k} + (1-\alpha)\hat{l} \tag{A.17}$$

$$\delta\hat{i} = \hat{k}' - (1-\delta)\hat{k} \tag{A.18}$$

$$(1 + 1/\nu)\hat{l} = \hat{y} - \hat{b} - \tau\hat{c}. \tag{A.19}$$

We proceed with the log-linearization of $\Phi_1(x)$ in (A.11). Differentiating with respect to the argument yields

$$\Phi_{11}(x) = \phi_1 \exp\{-\phi_2(x-1)\}.$$

Log-linearizing around $x = \exp(z) = 1$ leads to the approximation:

$$\Phi_1\big(\exp(z)\big) \approx \Phi_1(1) + \Phi_{11}(1)\cdot 1 \cdot (z - 0).$$

In turn, we can write

$$\Phi_1(K'/K) \approx \phi_1(\hat{k}' - \hat{k}),$$

which shows that the linex adjustment cost function is equivalent, up to second order, to a quadratic adjustment cost function

$$\Phi(K'/K) \approx \frac{\phi_1}{2}\big(K'/K - 1\big)^2.$$

We now turn to the log-linearization of (A.10) using the observation that $\Phi_1(1) = 0$:

$$\begin{aligned} & -\tau C_*^{-\tau}\hat{c} + C_*^{-\tau}\phi_1(\hat{k}' - \hat{k}) \\ = \quad & -\tau\beta C_*^{-\tau}(\alpha Y_*/K_* + 1 - \delta)\mathbb{E}[\hat{c}'] + \alpha\beta C_*^{-\tau}\frac{Y_*}{K_*}\mathbb{E}[\hat{y}' - \hat{k}'] + \phi_1\beta C_*^{-\tau}\mathbb{E}[\hat{k}'' - \hat{k}']. \end{aligned}$$

Multiplying by C_*^{τ}, using (A.12), and noting that $\hat{k}'$ is in the information for the conditional expectation $\mathbb{E}[\cdot]$ yields the simplified equation:

$$-\tau\hat{c} + \phi_1(\hat{k}' - \hat{k}) = -\tau\mathbb{E}[\hat{c}'] + \big(1 - \beta(1-\delta)\big)\big(\mathbb{E}[\hat{y}'] - \hat{k}'\big) + \phi_1\beta\big(\mathbb{E}[\hat{k}''] - \hat{k}'\big). \tag{A.20}$$

Equations (A.16) to (A.20) and the laws of motion for $\hat{z}$ and $\hat{b}$ form a linear rational expectations system that determines the dynamics of the model.

After setting $Y_* = L_* = 1$, the measurement equations in (25) can be written as

$$\ln Y^o = \hat{y} + \eta_Y, \quad \ln I^o = \ln\left(\frac{\alpha\delta}{1/\beta - (1-\delta)}\right) + \hat{i} + \eta_I, \quad \ln L^o = \hat{l} + \eta_l. \tag{A.21}$$

C.2 Model Solution, and Computational Details

While the approximate model M_0 refers to a first-order linearization around the steady state and is described in Section C.1 above, we obtain M_1 as a second-order linearization around the steady state, computed following Schmitt-Grohé and Uribe (2004). To implement it in Julia, we use the package *SolveDSGE*, developed by Richard Dennis and available at *https://github.com/RJDennis*.

C.3 Further Results for the RBC Model

Fig. A-6: RBC Model: Simulated Data

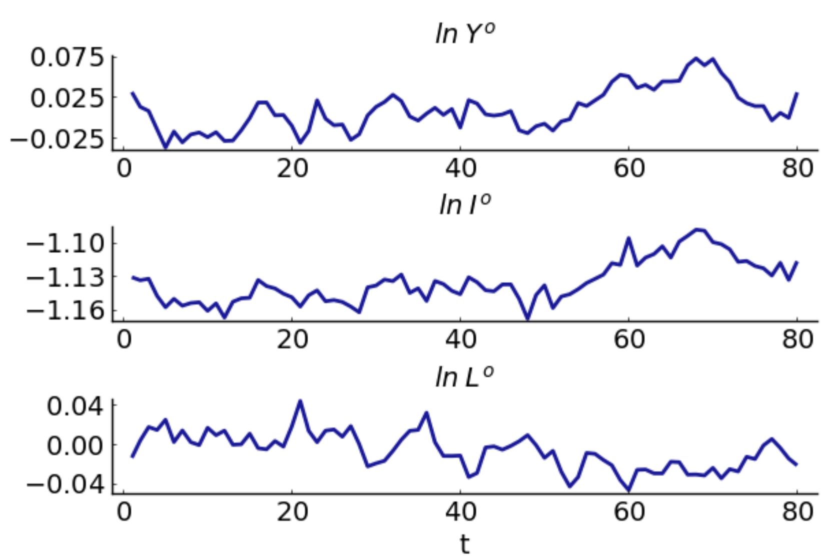
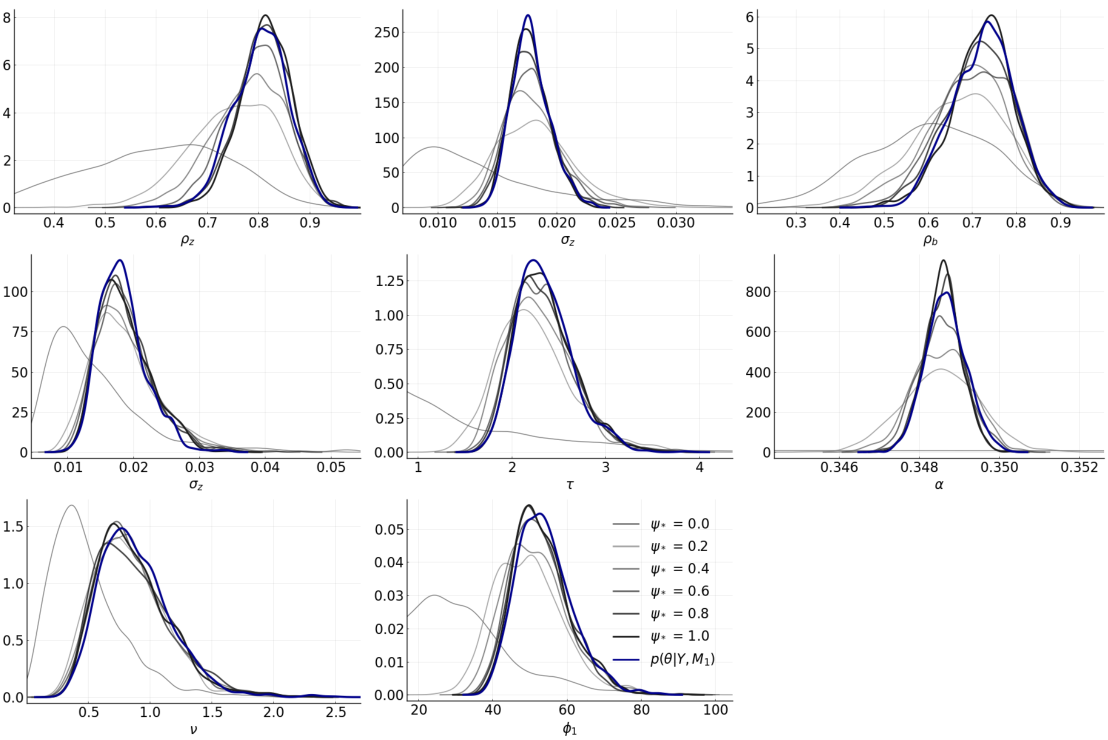
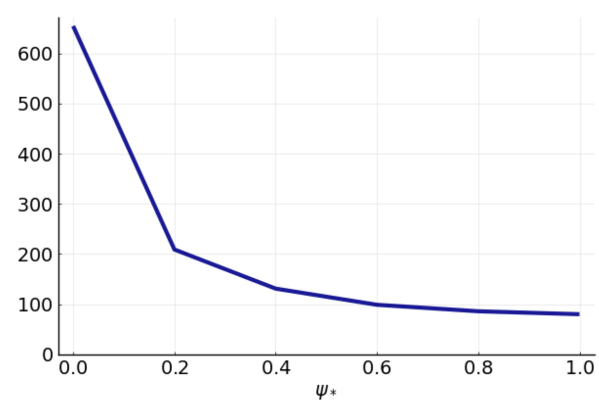


Fig. A-7: RBC Model: Target and Approximate Posterior Densities



Notes: Each plot refers to a different parameter. The approximating posterior densities obtained from the tempered M_0 likelihood function for $\psi_* \in \{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$ are plotted in shades (the larger ψ_* the darker) of gray. The M_1 posterior is depicted in blue.

Fig. A-8: RBC Model: Absolute Runtimes



Notes: Absolute runtime as a function of ψ_* based on a single run ($N_{run} = 1$).