Notes on Linear Rational Expectation models

Alex Carrasco Martinez
(alex.carmar93@gmail.com)

First year master in economics - PUC Rio

December, 2017

Abstract

I document some widely used methods in the solution and estimation of linear rational expectation models. I start presenting different solution methods and explaining how we can use these algorithms to solve simple models. Then, I make a brief review of Bayesian statistics. After that, I explain in a general way the usefulness of the Kalman Filter and give some useful examples. Finally, I present a simpler sampler algorithm used in Bayesian estimation of linear rational expectation models and expose the way we implement it in a small open economy model using Peruvian data.

This document does not have any relation with the university beyond the fact that I study there. Errors are my own.

CONTENTS

1 Solution methods **4**

1.1 Undetermined coefficients method 4

1.2 Generalized Schur Decomposition 5

1.3 Sims' method 5

1.4 Examples 6

2 Bayesian Statistic: A brief review **10**

2.1 Likelihood function 10

2.2 Priors 12

2.3 Inference 12

2.3.1 Importance Sampling (IS) 13

2.3.2 Metropolis-Hastings (MH) algorithms: a first approximation 14

3 Kalman Filter **16**

3.1 Forecasting t based on $t-1$ 16

3.2 Updating the inference about $\mathbf{z}_t$ 17

3.3 Using the Kalman filter to evaluate the likelihood function 18

3.4 Smoothing 18

3.5 A simple example 19

4 Metropolis-Hastings algorithms for linear RE models **20**

4.1 Random Walk Metropolis Hastings (RWMH) algorithm 21

4.2 Block MH algorithm 21

4.3 Initializing the algorithm: the posterior mode 22

5 Application: An small open economy model **23**

5.1 Model 23

5.1.1 Households 23
5.1.2 Firms 23
5.1.3 Aggregation 24
5.2 Log-linearized version 24
5.3 Linear system representation 25
5.4 Bayesian estimation 29
5.4.1 Priors 30
5.4.2 Estimation and Results 30

1 Solution methods

In this section I will assume that we have already obtained the log-linearized version of the non-linear system of equations that determines an equilibrium.

1.1 Undetermined coefficients method

Following Uhlig (1995), the structural linear equations could be divided in three blocks:

$$0 = A\mathbf{x}_t + B\mathbf{x}_{t-1} + C\mathbf{y}_t + D\mathbf{z}_t \tag{1.1}$$

$$0 = \mathbb{E}_t[F\mathbf{x}_{t+1} + G\mathbf{x}_t + H\mathbf{x}_{t-1} + J\mathbf{y}_{t+1} + K\mathbf{y}_t + L\mathbf{z}_{t+1} + M\mathbf{z}_t] \tag{1.2}$$

$$\mathbf{z}_{t+1} = N\mathbf{z}_t + \mathbf{u}_{t+1}; \qquad \mathbb{E}_t[\mathbf{u}_{t+1}] = 0, \;\; \mathbb{E}_t[\mathbf{u}_{t+1}\mathbf{u}'_{t+1}] = \Sigma_u \tag{1.3}$$

where $\mathbf{x}_t$ is an $(m \times 1)$ endogenous state vector (k_t), $\mathbf{y}_t$ is a vector with the other endogenous variables $(n\times)$, $\mathbf{z}_t$ represents the k exogenous stochastic processes (λ_t), and $\mathbf{u}_t$ are the vector of structural shocks [1]. What one is looking for is the recursive equilibrium law of motion:

$$\mathbf{x}_t = P\mathbf{x}_{t-1} + P_z\mathbf{z}_t \tag{1.4}$$

$$\mathbf{y}_t = R\mathbf{x}_{t-1} + R_z\mathbf{z}_t \tag{1.5}$$

Uhlig (1995) shows that if there is recursive equilibrium law of motion solving the system given above, the coefficients matrices can be found as follows: let C^+ be the pseudo-inverse of C, and C^0 be an $(l-n)\times l$ matrix, whose rows form a basis of the null space of C',

1. P satisfies

$$\overbrace{\begin{bmatrix}\mathbf{0}_{(l-n)\times m}\\ F - JC^+A\end{bmatrix}}^{\Psi} P^2 + \overbrace{\begin{bmatrix}C^0A\\ JC^+B - G + KC^+A\end{bmatrix}}^{\Gamma} P + \overbrace{\begin{bmatrix}C^0B\\ KC^+B - H\end{bmatrix}}^{\Theta} = 0 \tag{1.6}$$

   Let $\Xi = \begin{bmatrix}\Gamma & \Theta\\ \mathbf{I} & \mathbf{0}\end{bmatrix}$ and $\Delta = \begin{bmatrix}\Psi & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}$, λ the generalized eigenvalue of Ξ respect to Δ with modulus less than 1, Ω the eigenvector associated to λ, then $P = \Omega diag(\lambda)\Omega^{-1}$. If $|\lambda_i| < 1$ for all $i = 1, \ldots, m$, then there exists a stable solution. Moreover, if the number of generalized eigenvalues with modulus less than one is equal to m, then that solution is locally unique.

2. R is given by

$$R = -C^+(AP + B) \tag{1.7}$$

3. Let

$$V = \begin{bmatrix}\mathbf{I}_k \otimes A & \mathbf{I}_k \otimes C\\ N' \otimes F + \mathbf{I}_k \otimes (FP + JR + G) & N' \otimes J + \mathbf{I}_k \otimes K\end{bmatrix}$$

   then

$$V\begin{bmatrix}vec(P_z)\\ vec(R_z)\end{bmatrix} = \begin{bmatrix}vec(D)\\ vec(LN + M)\end{bmatrix} \tag{1.8}$$

[1] It is assumed: a) (not so important) C is order $l\times n$ with $l \geq n$ and complete column rank, b) F is of size $(m+n+l)\times n$, and c) N has stable eigenvalues

1.2 Generalized Schur Decomposition

In this part of the document I will follow Klein (2000)'s work. Consider the system[2]

$$A\mathbb{E}_t\begin{bmatrix}\mathbf{x}_{t+1}\\ \mathbf{y}_{t+1}\end{bmatrix} = B\begin{bmatrix}\mathbf{x}_t\\ \mathbf{y}_t\end{bmatrix} + C\mathbf{z}_t \tag{1.9}$$

$$\mathbf{z}_{t+1} = \Phi\mathbf{z}_t + D\mathbf{u}_{t+1} \tag{1.10}$$

Again, we are looking for a recursive equilibrium law of motion described by matrices P, Q, R, S as above[3].

Teorema 1.1. *(**Complex Generalized Schur Decomposition**) Let $A, B \in \mathbb{C}^{(m+n)\times(m+n)}$, and $P(z) = Az - B$, then $\exists\ Q, Z \in \mathcal{C}^{(m+n)\times(m+n)}$ orthonormal matrices such that*

1. *$QAZ = S$ is upper triangular matrix.*
2. *$QBZ = T$ is upper triangular matrix.*
3. *$\mathbb{I}[s_{ii} = 0] + \mathbb{I}[t_{ii} = 0] < 2$ for all i.*
4. *The set of generalized eigenvalues of B respect to A is given by $\lambda(B, A) = \{z \in \mathbb{C} : |P(z)| = 0\} = \{t_{ii}/s_{ii} : s_{ii} \neq 0\}$.*
5. *The pairs (s_{ii}, t_{ii}), $i = 1, \ldots, m+n$ can be arranged in any order.*

Let m_s be the number of stable eigenvalues, i.e., $|s_{ii}| > |t_{ii}|$. Blanchard and Khan (1980) conditions can be restated as: a) If $m_s = m$, then there exist a locally unique stable solution; b) If $m_s < m$, then there doesn't exist a stable solution; and c) If $m_s > m$, then there exist multiple stable solutions. Arranging S, T such that the m_s stable generalized eigenvalues come first, then

$$Z = \begin{bmatrix}Z_{11} & Z_{12}\\ Z_{21} & Z_{22}\end{bmatrix}, \quad Q = \begin{bmatrix}Q_1\\ Q_2\end{bmatrix}$$

Let assume that Z_{11} is square and invertible (Case a), then the unique stable solution is given by

$$P = Z_{11}S_{11}^{-1}T_{11}Z_{11}^{-1} \tag{1.11}$$

$$P_z = -Z_{11}S_{11}^{-1}T_{11}Z_{11}^{-1}Z_{12}M + Z_{11}S_{11}^{-1}[T_{12}M - S_{12}M\Phi + Q_1C] + Z_{12}M\Phi \tag{1.12}$$

$$R = Z_{21}Z_{11}^{-1} \tag{1.13}$$

$$R_z = (Z_{22} - Z_{21}Z_{11}^{-1}Z_{12})M \tag{1.14}$$

$$vec(M) = [\Phi' \otimes S_{22} - \mathbf{I}_k \otimes T_{22}]^{-1}vec(Q_2C) \tag{1.15}$$

1.3 Sims' method

Sims (2002) suggested to cast the model in the form:

$$\Gamma_0(\theta)\mathbf{y}_t = \Gamma_1(\theta)\mathbf{y}_{t-1} + C + \Psi_z\mathbf{z}_t + \Pi\eta_t \tag{1.16}$$

for every t where C is a matrix of constants, $\mathbf{z}_t$ is a vector of exogenous process, and η_t is an expectational error satisfying $\mathbb{E}_t(\eta_{t+1}) = 0$ for all t. The solution of this system is

$$\mathbf{y}_t = \Theta_1(\theta)\mathbf{y}_{t-1} + \Theta_c(\theta) + \Theta_0(\theta)\mathbf{z}_t + \Theta_y(\theta)\sum_{s=1}^{\infty}\Theta_f^{s-1}(\theta)\Theta_z(\theta)\mathbb{E}_t[\mathbf{z}_{t+s}] \tag{1.17}$$

[2] Klein (2000) generalizes the concept of predetermined variables, $\mathbf{x}_{t+1} = \mathbf{x}_t + \xi_{t+1}$ where $\xi_{t+1}|\mathcal{F}_t$ is a difference martingale.

[3] However, this method considers solution in the form: $\mathbf{x}_t = R\mathbf{x}_{t-1} + S\mathbf{z}_t + \xi_t$.

If we specify the initial system making the $\mathbf{z}_t$ vector an iid process, then

$$\mathbf{y}_t = \Theta_1(\theta)\mathbf{y}_{t-1} + \Theta_c(\theta) + \Theta_0(\theta)\mathbf{z}_t \tag{1.18}$$

Moreover if there is no C, then

$$\mathbf{y}_t = \Theta_1(\theta)\mathbf{y}_{t-1} + \Theta_0(\theta)\mathbf{z}_t \tag{1.19}$$

where this specific form for the solution will be extremely useful to show Bayesian estimation methods.

1.4 Examples

Basic RBC model

CPOs and clearing market conditions:

$$\begin{aligned}
Y_t &= A_t K_t^{\alpha} N_t^{1-\alpha} \\
C_t^{-\sigma} &= \beta E_t[C_{t+1}^{-\sigma}(1 + r_{t+1} - \delta)] \\
BN_t^{\varphi} &= W_t C_t^{-\sigma} \\
K_{t+1} &= Y_t + (1-\delta)K_t - C_t \\
r_t &= \alpha \frac{Y_t}{K_t} \\
W_t &= (1-\alpha)\frac{Y_t}{N_t} \\
ln(A_t) &= \rho_a ln(A_{t-1}) + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma_a^2)
\end{aligned}$$

The non-stochastic steady state of the model is characterized by $x^{ss} = x_t^{ss} = x_{t+1}^{ss}$ for every variable and $\varepsilon_t = 0$ for all $t \geq 0$. It is easy to see that $A^{ss} = 1$ and $r^{ss} = \beta^{-1} - 1 + \delta$. Then

$$\begin{aligned}
&\frac{K^{ss}}{N^{ss}} = (\alpha / r^{ss})^{\frac{1}{1-\alpha}}, \qquad && \frac{Y^{ss}}{N^{ss}} = (\alpha / r^{ss})^{\frac{\alpha}{1-\alpha}} \\
&W^{ss} = (1-\alpha)Y^{ss}/N^{ss}, && \frac{C^{ss}}{N^{ss}} = \frac{Y^{ss}}{N^{ss}} - \delta\frac{K^{ss}}{N^{ss}} \\
&N^{ss} = \Big[\frac{W^{ss}}{B}\Big(\frac{C^{ss}}{N^{ss}}\Big)^{-\sigma}\Big]^{1/(\varphi+\sigma)}
\end{aligned}$$

Log-linearization

$$\begin{aligned}
0 &= y_t - \frac{C^{ss}}{Y^{ss}}c_t + \frac{K^{ss}}{Y^{ss}}[(1-\delta)k_t - k_{t+1}] \\
0 &= a_t + \alpha k_t + (1-\alpha)n_t - y_t \\
0 &= w_t - \sigma c_t - \varphi n_t \\
0 &= r^{ss}[y_t - k_t] - \tilde{r}_t \\
0 &= y_t - n_t - w_t \\
\\
0 &= \sigma[c_t - E_t c_{t+1}] + \beta E_t \tilde{r}_{t+1} \\
\\
0 &= \rho_a a_{t-1} + \varepsilon_t - a_t
\end{aligned}$$

where $x_t = \ln(X_t) - \ln(X^{ss})$ and $\tilde{r}_t = r_t - r^{ss}$. Following Uhlig (1995) the model could be written in matrix system form, $\mathbf{x}_t = [k_{t+1}]$, $\mathbf{y}_t = [y_t, c_t, n_t, w_t, \tilde{r}_t]'$ and $\mathbf{z}_t = [a_t]$, $F = G = H = L = M = 0$,

$N = \rho_a$, and

$$A = \begin{bmatrix} -\frac{K^{ss}}{Y^{ss}} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, B = \begin{bmatrix} \frac{K^{ss}}{Y^{ss}}(1-\delta) \\ \alpha \\ 0 \\ -r^{ss} \\ 0 \end{bmatrix}, C = \begin{bmatrix} 1 & -\frac{C^{ss}}{Y^{ss}} & 0 & 0 & 0 \\ -1 & 0 & (1-\alpha) & 0 & 0 \\ 0 & -\sigma & -\varphi & 1 & 0 \\ r^{ss} & 0 & 0 & 0 & -1 \\ 1 & 0 & -1 & -1 & 0 \end{bmatrix}, D = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{1.20}$$

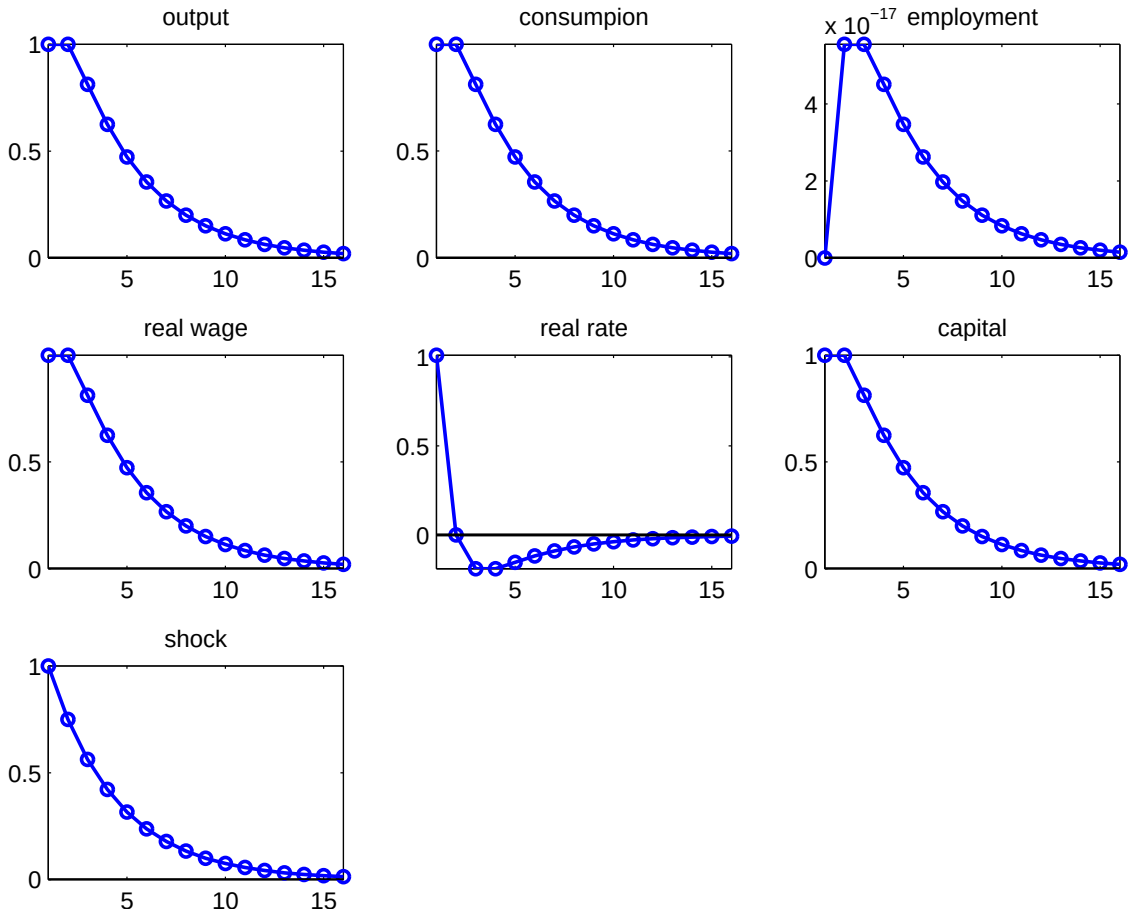
$$J = \begin{bmatrix} 0 & -\sigma & 0 & 0 & \beta \end{bmatrix}, \quad K = \begin{bmatrix} 0 & \sigma & 0 & 0 & 0 \end{bmatrix} \tag{1.21}$$

Following Klein (2000), $\Phi = \rho_a$, $D = \sigma_a$, and

$$A = \begin{bmatrix} \frac{K^{ss}}{Y^{ss}} & 0 & 0 & 0 & 0 & 0 \\ & & \mathbf{0}_{4\times 6} & & & \\ 0 & 0 & \sigma & 0 & 0 & -\beta \end{bmatrix}, B = \begin{bmatrix} \frac{K^{ss}}{Y^{ss}}(1-\delta) & 1 & -\frac{C^{ss}}{Y^{ss}} & 0 & 0 & 0 \\ \alpha & -1 & 0 & (1-\alpha) & 0 & 0 \\ 0 & 0 & -\sigma & -\varphi & 1 & 0 \\ -r^{ss} & r^{ss} & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 & -1 & 0 \\ 0 & 0 & \sigma & 0 & 0 & 0 \end{bmatrix}, D = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{1.22}$$

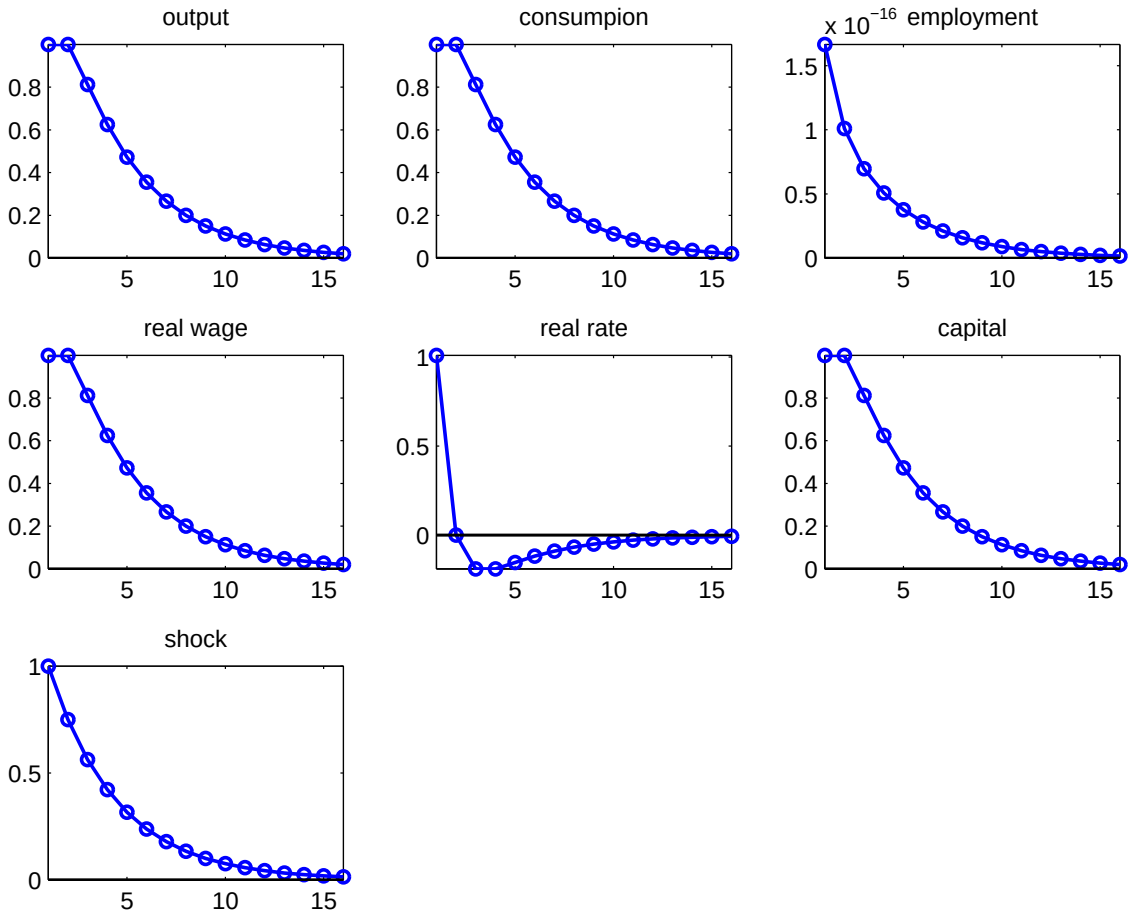
The script `main.m` gets the IRFs for this model using both methods. Results are shown in Fig. 1 and Fig. 2.

Figure 1. *Simple RBC model (solved by Uhlig's solution method)*



Note. This figure was created using REM_ucm.m and IRF.m functions in main_sol.m script.

FIGURE 2. *RBC model (IRFs using Klein's solution method)*



Note. This figure was created using REM_gschur.m and IRF.m functions in main_sol.m script.

BASIC NEW KEYNESIAN MODEL

In the basic New Keynesian model (as in Galí (2008)), there is not endogenous state variables, i.e., $m = 0$.

$$\begin{aligned}
\pi_t &= \beta \mathbb{E}_t \pi_{t+1} + \kappa \tilde{y}_t \\
\tilde{y}_t &= -\sigma^{-1}[i_t - \mathbb{E}_t \pi_{t+1} - r_t^n] + \mathbb{E}_t \tilde{y}_{t+1} \\
y_t &= \tilde{y}_t + y_t^n \\
i_t &= \rho + \phi_\pi \pi_t + \phi_y (y_t - y_t^{ss}) + v_t
\end{aligned}$$

$$\begin{aligned}
r_t^n &= \rho + \psi_{ya} \sigma \mathbb{E}_t \Delta a_{t+1} + (1 - \rho_z) z_t \\
y_t^n &= \psi_{ya} a_t + \psi_y
\end{aligned}$$

with $\kappa = \frac{(1-\theta)(1-\theta\beta)}{\theta}\zeta$, $\zeta = \frac{\varphi+\alpha+\sigma(1-\alpha)}{1-\alpha+\alpha\epsilon}$, $\psi_{ya} = \frac{1+\varphi}{(1-\alpha)\sigma+\varphi+\alpha}$, and $\psi_y = -\frac{(1-\alpha)(\mu-\ln(1-\alpha))}{\sigma(1-\alpha)+\varphi+\alpha}$.

STATIONARY TECHNOLOGY PROGRESS

Suppose $a_t = \rho_a a_{t-1} + \varepsilon_t^a$ where $\rho_a \in [0, 1)$. Turning off every shock, $a_t^{ss} = \lim_{k\to\infty} a_{t+k} = \lim_{k\to\infty} \rho_a^k a_{t-1} = 0$ and $y_t^{ss} = \psi_y$. Hence the NK model could be simplified

$$\beta \mathbb{E}_t \pi_{t+1} = \pi_t - \kappa \tilde{y}_t$$

$$\sigma^{-1}\mathbb{E}_t\pi_{t+1} + \mathbb{E}_t\tilde{y}_{t+1} = \sigma^{-1}\phi_\pi\pi_t + [1+\sigma^{-1}\phi_y]\tilde{y}_t + [\sigma^{-1}\phi_y\psi_{ya} + \psi_{ya}(1-\rho_a)]a_t - \sigma^{-1}(1-\rho_z)z_t + \sigma^{-1}v_t$$

Let $\mathbf{y}_t = [\pi_t, \tilde{y}_t]'$ and $\mathbf{z}_t = [a_t, z_t, v_t]'$, the system could be written in matrix form

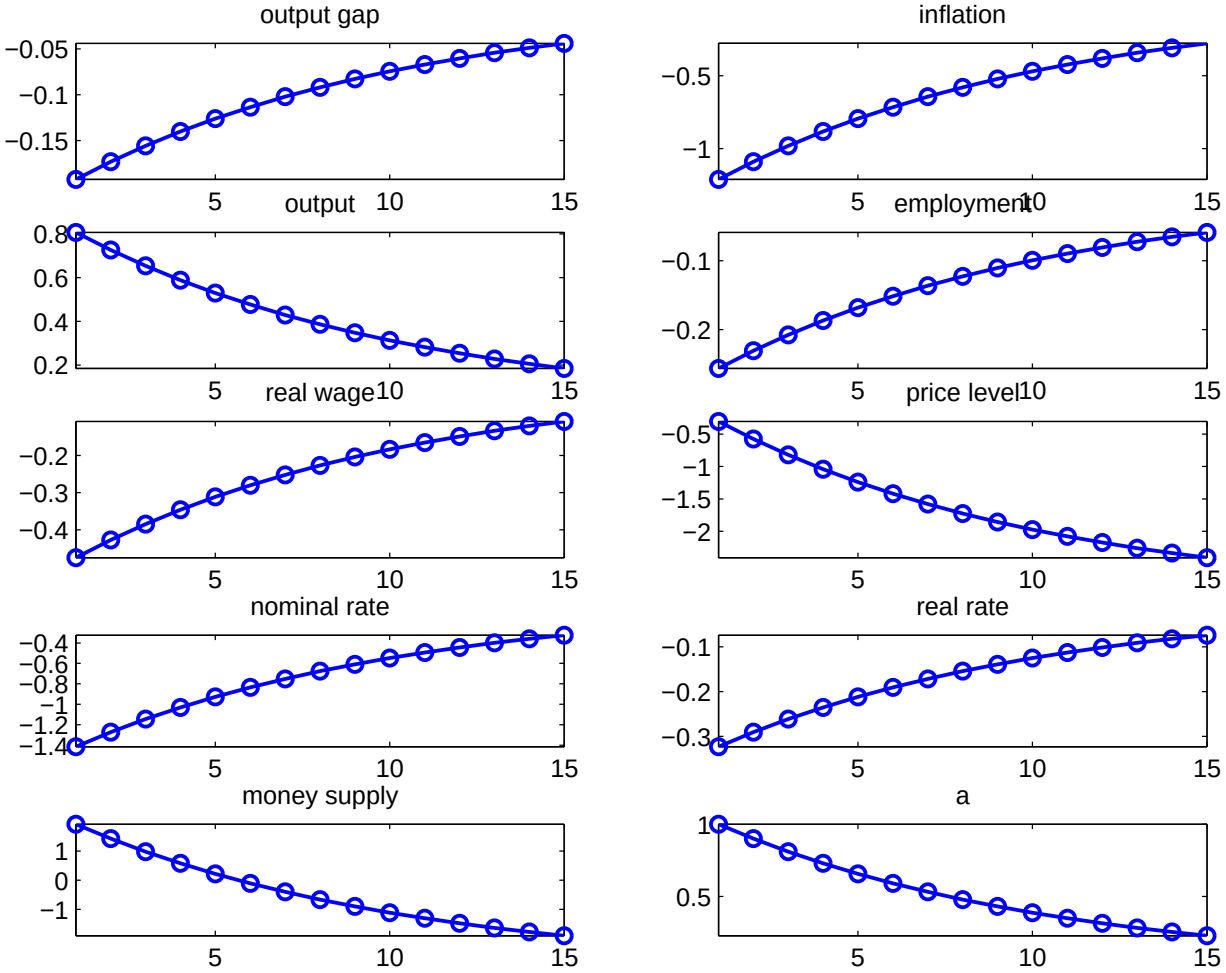
$$\begin{bmatrix} \beta & 0 \\ \frac{1}{\sigma} & 1 \end{bmatrix} \begin{bmatrix} \mathbb{E}_t\pi_{t+1} \\ \mathbb{E}_t\tilde{y}_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & -\kappa \\ \frac{\phi_\pi}{\sigma} & 1+\frac{\phi_y}{\sigma} \end{bmatrix} \begin{bmatrix} \pi_t \\ \tilde{y}_t \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ \psi_{ya}[(1-\rho_a)+\frac{\phi_y}{\sigma}] & -\frac{1-\rho_z}{\sigma} & \frac{1}{\sigma} \end{bmatrix} \begin{bmatrix} a_t \\ z_t \\ v_t \end{bmatrix} \tag{1.23}$$

with

$$\begin{bmatrix} a_{t+1} \\ z_{t+1} \\ v_{t+1} \end{bmatrix} = \begin{bmatrix} \rho_a & 0 & 0 \\ 0 & \rho_z & 0 \\ 0 & 0 & \rho_v \end{bmatrix} \begin{bmatrix} a_t \\ z_t \\ v_t \end{bmatrix} + \begin{bmatrix} \sigma_a & 0 & 0 \\ 0 & \sigma_z & 0 \\ 0 & 0 & \sigma_v \end{bmatrix} \begin{bmatrix} u^a_{t+1} \\ u^z_{t+1} \\ u^v_{t+1} \end{bmatrix}$$

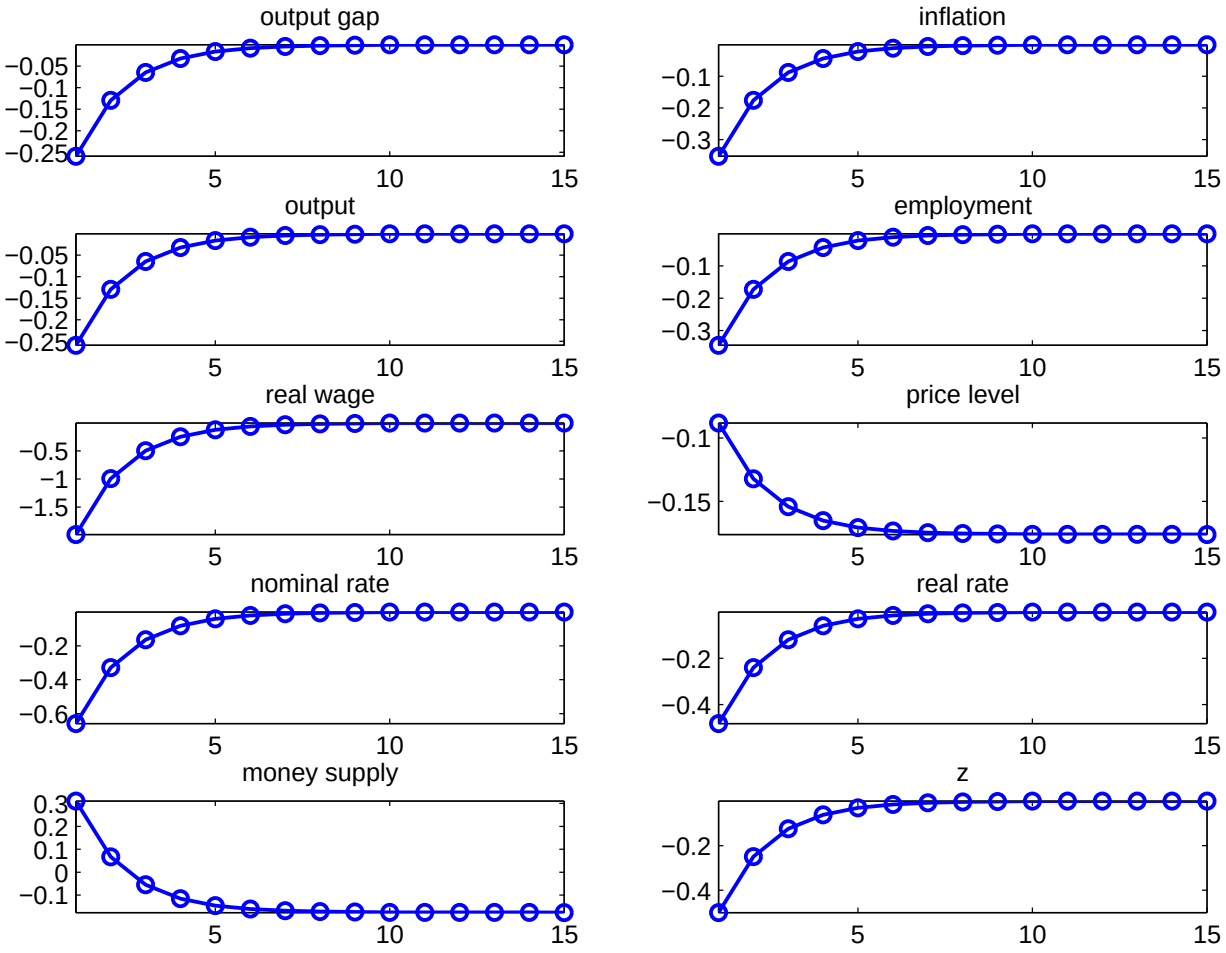
Figure 3, Fig. 4, and Fig. 5 show impulse-response functions of the basic new Keynesian model for technology, preference, and monetary shocks, respectively.

FIGURE 3. *Stationary technology progress, technology shock*



Note. This figure was created using REM_gschur.m and IRF.m functions in main_sol.m script.

FIGURE 4. *Stationary technology progress, preference shock*



Note. This figure was created using REM_gschur.m and IRF.m functions in main_sol.m script.

2 BAYESIAN STATISTIC: A BRIEF REVIEW

Based on Herbst and Schorfheide (2015). A Bayesian model consists of a joint distribution of data Y and parameters $\boldsymbol{\theta}$, $\mathbb{P}(Y, \boldsymbol{\theta})$. Specifically, we are interested in

$$\mathbb{P}(\boldsymbol{\theta}|Y) = \frac{\mathbb{P}(Y|\boldsymbol{\theta})\mathbb{P}(\boldsymbol{\theta})}{\mathbb{P}(Y)} \tag{2.1}$$

In order to turn linear RE models into Bayesian models we need to specify: a) probability distribution of the structural shocks, and b) prior probability distribution for $\boldsymbol{\theta}$. Once we solve the model conditional on $\boldsymbol{\theta}$, and get the *state-space representation* of the structural model, we can use an algorithm that evaluates the likelihood function for a given data set (this requires a filter due to the presence of unobservable variables).

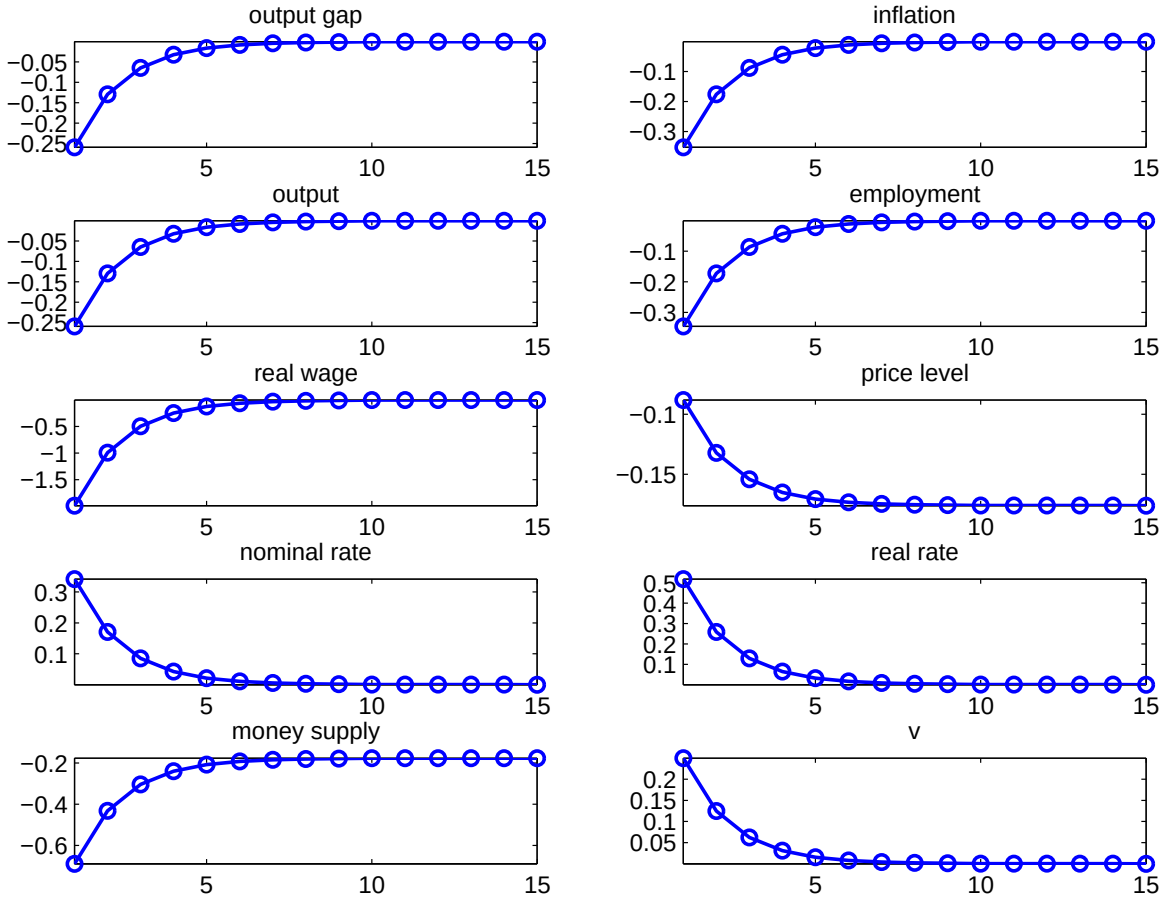
2.1 LIKELIHOOD FUNCTION

Consider the equilibrium law of motion of the structural model (see section 1)

$$\mathbf{z}_t = A_1(\boldsymbol{\theta})\mathbf{z}_{t-1} + A_\varepsilon(\theta)\boldsymbol{\varepsilon}_t \tag{2.2}$$

In order to construct a likelihood function, we have to relate the model variables $\mathbf{z}_t$ to a set of observable variables $\mathbf{y}_t$. Thus, the specification of the empirical model is completed by a set of *measurement*

FIGURE 5. *Stationary technology progress, monetary shock*



Note. This figure was created using REM_gschur.m and IRF.m functions in main_sol.m script.

equations:

$$\mathbf{y}_t = \Psi_0(\boldsymbol{\theta}) + \Psi_1(\boldsymbol{\theta})t + \Psi_2(\boldsymbol{\theta})\mathbf{z}_t + \boldsymbol{u}_t \tag{2.3}$$

where $\boldsymbol{u}_t$ is the vector of measurement errors. The state-space representation of the model provides a joint density for the observations and latent states given the parameters,

$$\mathbb{P}(Y_{1:T}, Z_{1:T}|\boldsymbol{\theta}) = \prod_{t=1}^{T} \mathbb{P}(\mathbf{y}_t, \mathbf{z}_t|Y_{1:t-1}, Z_{1:t-1}, \boldsymbol{\theta}) = \prod_{t=1}^{T} \mathbb{P}(\mathbf{y}_t|\mathbf{z}_t, \boldsymbol{\theta})\mathbb{P}(\mathbf{z}_t|\mathbf{z}_{t-1}, \boldsymbol{\theta})$$

But in order to compute the posterior distribution we need to compute $\mathbb{P}(Y_{1:T}|\boldsymbol{\theta}) = \prod_{t=1}^{T} \mathbb{P}(\mathbf{y}_t|Y_{1:t-1}, \boldsymbol{\theta})$, thus we must integrate the state variables, $\mathbf{z}_t$.

Generic filter. Let $\mathbb{P}(\mathbf{z}_0|Y_{1:0}, \boldsymbol{\theta}) = \mathbb{P}(\mathbf{z}_0|\boldsymbol{\theta})$. For $t = 1 : T$,

1. From iteration $t-1$ we have $\mathbb{P}(\mathbf{z}_{t-1}|Y_{1:t-1}, \boldsymbol{\theta})$.
2. Forecasting t given $t-1$:
   - *Transition equation*:

$$\mathbb{P}(\mathbf{z}_t|Y_{1:t-1}, \boldsymbol{\theta}) = \int \mathbb{P}(\mathbf{z}_t|\mathbf{z}_{t-1}, Y_{1:t-1}, \boldsymbol{\theta})\mathbb{P}(\mathbf{z}_{t-1}|Y_{1:t-1}, \boldsymbol{\theta})d\mathbf{z}_{t-1} \tag{2.4}$$

   - *Measurement equation*:

$$\mathbb{P}(\mathbf{y}_t|Y_{1:t-1}, \boldsymbol{\theta}) = \int \mathbb{P}(\mathbf{y}_t|\mathbf{z}_t, Y_{1:t-1}, \boldsymbol{\theta})\mathbb{P}(\mathbf{z}_t|Y_{1:t-1}, \boldsymbol{\theta})d\mathbf{z}_t \tag{2.5}$$

3. Updating with Bayes theorem. Once $\mathbf{y}_t$ becomes available,

$$\mathbb{P}(\mathbf{z}_t|Y_{1:t},\boldsymbol{\theta}) = \frac{\mathbb{P}(\mathbf{y}_t|\mathbf{z}_t,Y_{1:t-1},\boldsymbol{\theta})\mathbb{P}(\mathbf{z}_t|Y_{1:t-1},\boldsymbol{\theta})}{\mathbb{P}(\mathbf{y}_t|Y_{1:t-1},\boldsymbol{\theta})} \tag{2.6}$$

Generally we are interested in obtaining the whole distribution, however if the errors are Gaussian and given the log-linearized model, then Kalman filter can be used recursively to compute conditional means and covariances and thereby evaluate the likelihood of the distribution. Hence, it is common in linear DSGE modelling to specify Gaussian disturbances

$$\boldsymbol{\varepsilon}_t \sim \mathcal{IN}(0,\Sigma_\varepsilon) \tag{2.7}$$
$$\boldsymbol{u}_t \sim \mathcal{IN}(0,\Sigma_u) \tag{2.8}$$
$$\mathbf{z}_0 \sim \mathcal{IN}(\hat{\mathbf{z}}_{0|0},\mathbf{P}_{0|0}) \tag{2.9}$$

I will give further details in section 3.

2.2 Priors

As a first step in the elicitation of the prior distribution, it is useful to group the elements of $\boldsymbol{\theta}$ into three different categories. The first group, denoted $\boldsymbol{\theta}_{ss}$, contains parameters that affect the steady state of the DSGE model. The second group of parameters characterizes the law of motion of the exogenous shock process: $\boldsymbol{\theta}_{exo}$. Finally, the last group of parameters controls the endogenous propagation mechanisms without affecting the steady state of the model, $\boldsymbol{\theta}_{endo}$.

The tacit assumption underlying posterior inference with a prior that is constructed (at least in part) from a non-sample information is that $\mathbb{P}(Y|\mathcal{X}^0,\boldsymbol{\theta}) \approx \mathbb{P}(Y|\boldsymbol{\theta})$. This assumption is a reasonable one if the observations in $\mathcal{X}^0$ pre-date the observations in Y or if Y contains macro time series while $\mathcal{X}^0$ contains micro level data. Priors for $\boldsymbol{\theta}_{exo}$ are difficult to specify, however there is suggestion: i) specify a prior for $\boldsymbol{\theta}_{exo}$; ii) generate draws from this prior; iii) get sample moments based on simulated data; and iv) adjust the prior until the predictive distribution of the sample moments is consistent with prior beliefs. Sometimes, the linear DSGE model is re-parameterized in terms of its steady-state rather than $\boldsymbol{\theta}_{ss}$ to avoid generation of implausible steady-state.

The joint prior distribution is the product of the marginal densities. The domain of the prior is truncated to ensure that the linearized rational expectations model has a unique stable solution.

2.3 Inference

In a nutshell, Bayesian inference amounts to characterize properties of the posterior distribution $\mathbb{P}(\boldsymbol{\theta}|Y)$. Unfortunately, for many models it is not possible such a characterization analytically. In order to compute posterior quantiles and moments of functions $h(\boldsymbol{\theta})$ we have to rely on numerical techniques.

The posterior distribution, $\pi(\boldsymbol{\theta}) \equiv \mathbb{P}(\boldsymbol{\theta}|Y)$, summarizes the information about $\boldsymbol{\theta}$ after having observed data Y. Once we obtained $\pi(\boldsymbol{\theta})$, we can make a *direct sampling* to obtain an iid sample, $\{\boldsymbol{\theta}^i\}_i$ and compute a Monte Carlo approximation

$$\overline{h}_N = N^{-1}\sum_{i=1}^{N} h(\boldsymbol{\theta}^i) \tag{2.10}$$

Then, provided that $\mathbb{V}_\pi[h(\boldsymbol{\theta})] < \infty$ we have that

$$\overline{h}_N \xrightarrow{a.s.} \mathbb{E}_\pi[h] \tag{2.11}$$

$$\sqrt{N}(\overline{h}_N - \mathbb{E}_\pi[h]) \xrightarrow{d} \mathcal{N}(0, \mathbb{V}_\pi[h]) \tag{2.12}$$

From a Bayesian perspective it is optimal to make decisions that minimize the posterior expected loss of the decision maker . Let $\mathcal{L}(h(\boldsymbol{\theta}), \delta)$ denote the loss function under which the decision δ is evaluated, then posterior expected loss associated with a decision rule δ is given by

$$\rho(\delta|Y) = \int_\Theta \mathcal{L}(h(\boldsymbol{\theta}), \delta)\pi(\boldsymbol{\theta})d\boldsymbol{\theta}$$

Thus, a Bayes decision is a decision that minimizes the posterior expected loss:

$$\delta^* = arg\min_{\delta \in \mathcal{D}} \rho(\delta|Y) \tag{2.13}$$

to evaluate the posterior risk for each choice δ we can make a Monte Carlo approximation of the form

$$\overline{\rho}_N(\delta|Y) = N^{-1}\sum_{i=1}^{N} \mathcal{L}(h(\boldsymbol{\theta}^i), \delta)$$

$$\delta_N^* = arg\min_{\delta \in \mathcal{D}} \overline{\rho}(\delta|Y)$$

where the $\boldsymbol{\theta}^i$'s are draws from the posterior $\pi(\boldsymbol{\theta})$. For example, to point estimation: $\mathcal{L}(h(\boldsymbol{\theta}), \delta) = (h(\boldsymbol{\theta}) - \delta)^2$ or $\mathcal{L}(h(\boldsymbol{\theta}), \delta) = |h(\boldsymbol{\theta}) - \delta|$; to connected interval estimation: $\mathcal{L}(h(\boldsymbol{\theta}), \delta) = (\delta_u - \delta_l) + \max_{\lambda \in \mathbb{R}_-} \lambda[\mathbb{I}(\delta_l \leq h(\boldsymbol{\theta}) \leq \delta_u) - (1-\alpha)]$.

Identification problems is common in linear DSGE models an typically come in two variates: i) local identification problems in which the likelihood function is fairly flat in certain directions of the parameter space; and ii) global identification problems in which the likelihood function is multi-modal.

2.3.1 Importance Sampling (IS)

One way to obtained a sample from the posterior distribution is to direct sampling. Instead of this, we could approximate $\pi(.)$ by using a different, tractable density $g(\boldsymbol{\theta})$ that is easy to sample from and then reweigh the draws. Because in many applications the posterior density can only be evaluated up to a constant of proportionality, we write

$$\pi(\boldsymbol{\theta}) = \frac{f(\boldsymbol{\theta})}{c} \tag{2.14}$$

The IS algorithm is based on the following identity:

$$\mathbb{E}_\pi[h(\boldsymbol{\theta})] = \int h(\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta} = \frac{1}{c}\int h(\boldsymbol{\theta})\frac{f(\boldsymbol{\theta})}{g(\boldsymbol{\theta})}g(\boldsymbol{\theta})d\boldsymbol{\theta} \tag{2.15}$$

we can define importance weights as

$$w(\boldsymbol{\theta}) = \frac{f(\boldsymbol{\theta})}{g(\boldsymbol{\theta})}, \qquad \text{[unnormalized]} \tag{2.16}$$

$$v(\boldsymbol{\theta}) = \frac{w(\boldsymbol{\theta})}{\int w(\boldsymbol{\theta})g(\boldsymbol{\theta})d\boldsymbol{\theta}}, \quad \text{[normalized]} \tag{2.17}$$

Algorithm 1 (Importance sampling)

- *For $i = 1$ to N, draw an iid sample $\boldsymbol{\theta}^i \sim g(\boldsymbol{\theta})$ and compute the unnormalized importance weights*

$$w^i = \frac{f(\boldsymbol{\theta}^i)}{g(\boldsymbol{\theta}^i)} \tag{2.18}$$

- *Compute the normalized importance weights*

$$W^i = \frac{w^i}{N^{-1}\sum_{i=1}^N w^i} \tag{2.19}$$

An approximation of $\mathbb{E}_\pi[h(\boldsymbol{\theta})]$ is given by

$$\overline{h}_N = N^{-1}\sum_{i=1}^N W^i h(\boldsymbol{\theta}^i) \tag{2.20}$$

If $\mathbb{E}_g[hf/g] < \infty$ and $\mathbb{E}_g[f/g] < \infty$, then the Monte Carlo estimate $\overline{h}_N \xrightarrow{a.s.} \mathbb{E}_\pi[h(\boldsymbol{\theta})]$. Moreover, it can be deduced that

$$\sqrt{N}(\overline{h}_N - \mathbb{E}_\pi[h]) \xrightarrow{d} \mathcal{N}(0, \Omega(h)) \tag{2.21}$$

where $\Omega(h) = \mathbb{V}_g[(\pi/g)(h - \mathbb{E}_\pi[h])]$. We can now define the following inefficiency factor (relative to direct sampling),

$$InEff_\infty = \frac{\Omega(h)}{\mathbb{V}_\pi[h]} \tag{2.22}$$

This factor is typically greater than one, and using a crude approximation one can factorize $\Omega(h) \approx \mathbb{V}_\pi[h](\mathbb{V}_g[\pi/g] + 1)$, then

$$\Omega(h) \approx 1 + \mathbb{V}_g[\pi/g]$$

Then the approximation of the asymptotic inefficiency factor is independent of the function $h(.)$ and highlights that the larger the variance of the importance weights, the less accurate the Monte Carlo approximation using IS algorithm relative to the accuracy that could be achieved with a direct iid sampling.

2.3.2 Metropolis-Hastings (MH) algorithms: a first approximation

The basic idea is to constructs a Markov Chain such that the stationary distribution associated with this Markov chain is unique and equals the posterior distribution of interest, thus, instead of generate a sequence of independent draws the MH algorithm generates a sequence of serially correlated draws.

Algorithm 2 (Generic MH algorithm). *For $i = 1$ to N:*

- *Draw $\boldsymbol{\vartheta}$ from a pre-specified density $q(\boldsymbol{\vartheta}|\boldsymbol{\theta}^{i-1})$.*
- *Set $\boldsymbol{\theta}^i = \boldsymbol{\vartheta}$ with probability*

$$\alpha(\boldsymbol{\vartheta}|\boldsymbol{\theta}^{i-1}) = \min\left\{1, \frac{\mathbb{P}(Y|\boldsymbol{\vartheta})\mathbb{P}(\boldsymbol{\vartheta})/q(\boldsymbol{\vartheta}|\boldsymbol{\theta}^{i-1})}{\mathbb{P}(Y|\boldsymbol{\theta}^{i-1})\mathbb{P}(\boldsymbol{\theta}^{i-1})/q(\boldsymbol{\theta}^{i-1}|\boldsymbol{\vartheta})}\right\} \tag{2.23}$$

and $\boldsymbol{\theta}^i = \boldsymbol{\theta}^{i-1}$ otherwise.

Thus, implicitly, the MH algorithm characterizes a Markov transition kernel $K(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}})$. For algorithm 2 to generate a sequence of draws from the posterior distribution a necessary condition is that the posterior distribution is an invariant distribution under the transition kernel $K(.|.)$, i.e.,

$$\pi(\boldsymbol{\theta}) = \int K(\boldsymbol{\theta}|\tilde{\boldsymbol{\theta}})\mathbb{P}(\tilde{\boldsymbol{\theta}}|Y)d\tilde{\boldsymbol{\theta}} \tag{2.24}$$

Verifying the invariance property is relative straightforward. But this property is not sufficient to guarantee that the Monte Carlo average of draws $h(\boldsymbol{\theta})$ converge to the posterior expectation $\mathbb{E}_\pi[h(\boldsymbol{\theta})]$. In particular, one needs to ensure that the transition kernel has a unique invariant distribution, that repeated application of the transition kernel leads to convergence to the unique invariant distribution regardless of the chain's initialization, and that the autocorrelation of the draws $\boldsymbol{\theta}^i$ generated by the Markov chain decays sufficiently fast such that sample averages converge to population means.

Similar we can estimate the asymptotic variance for $\overline{h}_N(\boldsymbol{\theta})$ using HAC estimators (since the sample is not iid), e.g., a New-West consistent estimator

$$HAC[\overline{h}] = N^{-1}\Big(\hat{\gamma}_0 + 2\sum_{l=1}^{L}[1 - l/(L+1)]\hat{\gamma}_l\Big) \tag{2.25}$$

While the reduction in MC algorithm is desirable, it is unrealistic in practice because to generate a negatively correlated chain we must know the posterior distribution.

3 KALMAN FILTER

Based on Hamilton (1994). The Kalman filter is an algorithm for sequentially updating a linear projection for the system. Among other benefits, it provides a way to calculate exact finite-sample forecasts and the exact likelihood function when it is assumed Gaussian distributions. The state-space representation of the model of interest is

$$\begin{aligned}\mathbf{z}_t &= A_1(\boldsymbol{\theta})\mathbf{z}_{t-1} + A_\varepsilon(\boldsymbol{\theta})\varepsilon_t, \qquad \text{[Transition equation]}\\ \mathbf{y}_t &= \Psi_3(\boldsymbol{\theta})\mathbf{x}_t + \Psi_2(\boldsymbol{\theta})\mathbf{z}_t + \boldsymbol{u}_t, \quad \text{[Measurement equation]}\end{aligned}$$

where $\mathbf{x}_t = [1, t]$ and in general would represent a vector of exogenous or predetermined variables and additionally I will assume

$$\mathbb{E}[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_\tau'] = \begin{cases}\Sigma_\varepsilon, & for\ \ t=\tau\\ 0, & otherwise\end{cases}$$

$$\mathbb{E}[\boldsymbol{u}_t\boldsymbol{u}_\tau'] = \begin{cases}\Sigma_u, & for\ \ t=\tau\\ 0, & otherwise\end{cases}$$

$$\mathbb{E}[\boldsymbol{\varepsilon}_t\boldsymbol{u}_t'] = 0$$

i.e., I have not yet assumed normality of the errors, so I will derive first the linear projections of state variables using Kalman filter. The observable data is $\mathcal{F}_t \equiv (\mathbf{y}_1', \mathbf{y}_2', \dots, \mathbf{y}_T', \mathbf{x}_1', \mathbf{x}_2', \dots, \mathbf{x}_T')'$, and the objective is to obtain estimations of the state vector at t given the information at t

$$\begin{aligned}\hat{\mathbf{z}}_{t|t} &\equiv \hat{\mathbb{E}}(\mathbf{z}_t|\mathcal{F}_t)\\ \mathbf{P}_{t|t} &\equiv \mathbb{E}[(\mathbf{z}_t - \hat{\mathbf{z}}_{t|t})(\mathbf{z}_t - \hat{\mathbf{z}}_{t|t})']\end{aligned}$$

The recursion begins with $\hat{\mathbf{z}}_{0|0}$ which denotes a forecast of $\mathbf{z}_0$ based on no observations and this is just the unconditional mean $\mathbb{E}[\mathbf{z}_0]$ with associated MSE $\mathbf{P}_{0|0} = \mathbb{E}[\mathbf{z}_0 - \mathbb{E}[\mathbf{z}_0])(\mathbf{z}_0 - \mathbb{E}[\mathbf{z}_0])']$. If the model is completely stationary this matrices take the following for

$$\begin{aligned}\mathbb{E}[\mathbf{z}_0] &= 0\\ vec(\mathbf{P}_{0|0}) &= [\mathbf{I} - (A_1 \otimes A_1)]^{-1} vec(A_\varepsilon \Sigma_\varepsilon A_\varepsilon')\end{aligned}$$

where for simplicity of exposition I have removed the dependence of the state-space representation on deep parameters. However, if the system is not stationary then the initialization is based on the analyst's best guess and the confidence on it.

3.1 FORECASTING t BASED ON $t-1$

The *transition (or state) equation* is used to forecast $\mathbf{z}_t$,

$$\begin{aligned}\hat{\mathbf{z}}_{t|t-1} &= A_1\hat{\mathbb{E}}_{t-1}[\mathbf{z}_{t-1}] + \hat{\mathbb{E}}_{t-1}[\boldsymbol{\varepsilon}_t]\\ &= A_1\hat{\mathbf{z}}_{t-1|t-1}\end{aligned}$$

where $\hat{\mathbf{z}}_{t|t-1} \equiv \hat{\mathbb{E}}[\mathbf{z}_t|\mathcal{F}_{t-1}]$ and the MSE

$$\begin{aligned}\mathbf{P}_{t|t-1} &= \mathbb{E}[(\mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1})(\mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1})']\\ &= \mathbb{E}[(A_1\mathbf{z}_{t-1} + A_\varepsilon\boldsymbol{\varepsilon}_t - \hat{\mathbf{z}}_{t|t-1})(A_1\mathbf{z}_{t-1} + A_\varepsilon\boldsymbol{\varepsilon}_t - \hat{\mathbf{z}}_{t|t-1})']\\ &= A_1\mathbb{E}[(\mathbf{z}_{t-1} - \mathbf{z}_{t-1|t-1})(\mathbf{z}_{t-1} - \mathbf{z}_{t-1|t-1})']A_1' + A_\varepsilon\mathbb{E}[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t']A_\varepsilon'\\ &= A_1\mathbf{P}_{t-1|t-1}A_1' + A_\varepsilon\Sigma_\varepsilon A_\varepsilon'\end{aligned}$$

Moreover, we can forecast $\mathbf{y}_t$ based on $\mathcal{F}_{t-1}$ using the *measurement equation* and the fact that $\mathbf{x}_t$ is exogenous, i.e.,

$$\hat{\mathbb{E}}[\mathbf{z}_t|\mathbf{x}_t, \mathcal{F}_{t-1}] = \hat{\mathbb{E}}[\mathbf{z}_t|\mathcal{F}_{t-1}] = \hat{\mathbf{z}}_{t|t-1}$$

then

$$\begin{aligned}\hat{\mathbf{y}}_{t|t-1} &\equiv \hat{\mathbb{E}}[\mathbf{y}_t|\mathbf{x}_t, \mathcal{F}_{t-1}] \\ &= \Psi_3\mathbf{x}_t + \Psi_2\hat{\mathbf{z}}_{t|t-1}\end{aligned}$$

and this prediction is associated with the MSE

$$\begin{aligned}\mathbb{E}[(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})'] &= \Psi_2\mathbb{E}[(\mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1})(\mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1})']\Psi_2' + \mathbb{E}[\boldsymbol{u}_t\boldsymbol{u}_t'] \\ &= \Psi_2\mathbf{P}_{t|t-1}\Psi_2' + \Sigma_u\end{aligned}$$

3.2 Updating the inference about $\mathbf{z}_t$

To update a linear projection I will use the general form of any linear projection:

$$\begin{aligned}\hat{\mathbb{E}}[w_t|q_t] &= \mathbb{E}[w_t] + (\mathbb{C}[w_t q_t])(\mathbb{V}[q_t])^{-1} \times [q_t - \mathbb{E}[q_t]] \\ \hat{\mathbb{V}}[w_t|q_t] &= \mathbb{V}[w_t] - \mathbb{C}[w_t, q_t]\{\mathbb{V}[q_t]\}^{-1}\mathbb{C}[w_t, q_t]'\end{aligned}$$

for any w_t and q_t variables. If we make an update conditional on a specific information set, e.g., $\{\mathbf{x}_t, \mathcal{F}_{t-1}\}$, the linear projection becomes:

$$\hat{\mathbf{z}}_{t|t} = \hat{\mathbf{z}}_{t|t-1} + \Big\{\mathbb{E}[(\mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1})(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})']\Big\}\Big\{\mathbb{E}[(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})']\Big\}^{-1}[\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1}]$$

where $\hat{\mathbf{z}}_{t|t} \equiv \hat{\mathbb{E}}[\mathbf{z}_t|\mathbf{y}_t, \mathbf{z}_t, \mathcal{F}_{t-1}] = \hat{\mathbb{E}}[\mathbf{z}_t|\mathcal{F}_t]$. Note that

$$\begin{aligned}\mathbb{E}[(\mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1})(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})'] &= \mathbb{E}[(\mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1})(\Psi_2(\mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1}) + \boldsymbol{u}_t)'] \\ &= \mathbf{P}_{t|t-1}\Psi_2'\end{aligned}$$

In the same way, the MSE associated with this updated projection, denoted as $\mathbf{P}_{t|t}$, is

$$\begin{aligned}\mathbf{P}_{t|t} &= \mathbb{E}[(\mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1})(\mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1})'] \\ &\quad - \Big\{\mathbb{E}[(\mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1})(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})']\Big\}\Big\{\mathbb{E}[(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})']\Big\}^{-1}\Big\{\mathbb{E}[(\mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1})(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})']\Big\}' \\ &= \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\Psi_2'[\Psi_2\mathbf{P}_{t|t-1}\Psi_2' + \Sigma_u]^{-1}\Psi_2\mathbf{P}_{t|t-1}\end{aligned}$$

Hence, I have shown how to obtained recursive equations for the Kalman filter. Start with $\hat{\mathbf{z}}_{0|0}$ and $\mathbf{P}_{0|0}$, and then compute

$$\hat{\mathbf{z}}_{t|t-1} = A_1\hat{\mathbf{z}}_{t-1|t-1} \tag{3.1}$$
$$\mathbf{P}_{t|t-1} = A_1\mathbf{P}_{t-1|t-1}A_1' + A_\varepsilon\Sigma_\varepsilon A_\varepsilon' \tag{3.2}$$
$$\hat{\mathbf{y}}_{t|t-1} = \Psi_3\mathbf{x}_t + \Psi_2\hat{\mathbf{z}}_{t|t-1} \tag{3.3}$$
$$\mathbf{F}_{t|t-1} = \Psi_2\mathbf{P}_{t|t-1}\Psi_2' + \Sigma_u \tag{3.4}$$
$$\hat{\mathbf{z}}_{t|t} = \hat{\mathbf{z}}_{t|t-1} + \mathbf{P}_{t|t-1}\Psi_2'\mathbf{F}_{t|t-1}^{-1}[\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1}] \tag{3.5}$$
$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\Psi_2'\mathbf{F}_{t|t-1}^{-1}\Psi_2\mathbf{P}_{t|t-1} \tag{3.6}$$

for $t = 1, \ldots, T$ and where $\mathbf{F}_{t|t-1} = \mathbb{E}[(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})'$.

Observation: *Sometimes the interest is in obtaining forecasts of state and observable vector*[4]*, then we can start by specifying* $\hat{\mathbf{z}}_{1|0}$ *and* $\mathbf{P}_{1|0}$ *instead of* $\hat{\mathbf{z}}_{0|0}$ *and* $\mathbf{P}_{0|0}$ *and continue obtaining directly*

$$\hat{\mathbf{z}}_{t+1|t} = A_1\hat{\mathbf{z}}_{t|t-1} + \mathbf{K}_t[\mathbf{y}_t - \Psi_3\mathbf{x}_t - \Psi_2\hat{\mathbf{z}}_{t|t-1}] \tag{3.7}$$

$$\hat{\mathbf{y}}_{t+1|t} = \Psi_3\mathbf{x}_{t+1} + \Psi_2\hat{\mathbf{z}}_{t+1|t} \tag{3.8}$$

$$\mathbf{P}_{t+1|t} = [A_1 - \mathbf{K}_t\Psi_2]\mathbf{P}_{t|t-1}[A_1 - \mathbf{K}_t\Psi_2]' + \mathbf{K}_t\Sigma_u\mathbf{K}_t' + A_\varepsilon\Sigma_\varepsilon A_\varepsilon' \tag{3.9}$$

where $\mathbf{K}_t = A_1\mathbf{P}_{t|t-1}\Psi_2'[\Psi_2\mathbf{P}_{\mathbf{t|t-1}}\Psi_2' + \Sigma_u]$ *is the* ***gain matrix*** *and* $\mathbb{E}[(\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t})(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})'] = \Psi_2\mathbf{P}_{t+1|t}\Psi_2' + \Sigma_u$.

3.3 Using the Kalman filter to evaluate the likelihood function

When I obtained the recursive equations of the Kalman filter I have assumed that the matrices $A_1, A_\varepsilon, \Sigma_\varepsilon, \Psi_3, \Psi_2, \Sigma_u$ are known which is an assumption that is not always satisfied, then the natural question is concern in getting estimations for those matrices (or the deep parameter, $\boldsymbol{\theta}$). Now I will suppose a distribution for $\{\boldsymbol{\varepsilon}_t, \boldsymbol{u}_t\}_{t=1}^T$ and $\mathbf{z}_0$ and, finally, evaluate the likelihood function. The evaluation of the likelihood function serves for objectives: i) to estimate the unknown structural parameters, $\boldsymbol{\theta}$, by ML estimation; and ii) to obtained the posterior distribution for $\boldsymbol{\theta}$ in a Bayesian estimation context. Then in the rest of this document I will assume:

$$\boldsymbol{\varepsilon}_t \sim \mathcal{IN}(0, \Sigma_\varepsilon)$$
$$\boldsymbol{u}_t \sim \mathcal{IN}(0, \Sigma_u)$$
$$\mathbf{z}_0 \sim \mathcal{IN}(\hat{\mathbf{z}}_{0|0}, \mathbf{P}_{0|0})$$

Thus, it is easy to obtained the following distributions:

$$\mathbf{z}_t|\mathbf{x}_t, \mathcal{F}_{t-1}, \boldsymbol{\theta} \sim \mathcal{N}(\hat{\mathbf{z}}_{t-1|t-1}, \mathbf{P}_{t-1|t-1}) \tag{3.10}$$

$$\mathbf{y}_t|\mathbf{x}_t, \mathcal{F}_{t-1}, \boldsymbol{\theta} \sim \mathcal{N}(\hat{\mathbf{y}}_{t|t-1}, \mathbf{F}_{t|t-1}) \tag{3.11}$$

$$\mathbf{z}_t|\mathcal{F}_t, \boldsymbol{\theta} \sim \mathcal{N}(\hat{\mathbf{z}}_{t|t}, \mathbf{P}_{t|t}) \tag{3.12}$$

Hence,

$$f(\mathbf{y}_t|\mathbf{x}_t, \mathcal{F}_{t-1}, \boldsymbol{\theta}) = (2\pi)^{-k/2}|\mathbf{F}_{t|t-1}|^{-1/2}\exp\Big\{-\frac{1}{2}(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})'\mathbf{F}_{t|t-1}^{-1}(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})\Big\} \tag{3.13}$$

where k is the dimension of the vector of observable variables, $\boldsymbol{y}_t$. The sample likelihood function is given by

$$\mathbb{P}(Y_{1:T}|\boldsymbol{\theta}) = \prod_{T=1}^{T} f(\mathbf{y}_t|\mathbf{x}_t, \mathcal{F}_{t-1}, \boldsymbol{\theta}) \tag{3.14}$$

Note that in the initial transition equation $\mathbf{x}_t = [1, t]'$, then $Y_{1:t-1} = [\mathbf{x}_t, \mathbf{F}_{t-1}]$ then we have already get the likelihood of the observed data!

3.4 Smoothing

Commonly, the state vector is of interest for its own, then we would like to form an inference about the value of $\mathbf{z}_t$ based on the full set of data collected, i.e., $\hat{\mathbf{z}}_{t|T} \equiv \hat{\mathbb{E}}(\mathbf{z}|\mathcal{F}_T)$ which is called the *smoothed* estimate of $\mathbf{z}_t$. First of all, obtained

$$\mathbb{E}[(\mathbf{z}_t - \hat{\mathbf{z}}_{t|t})(\mathbf{z}_{t+1} - \hat{\mathbf{z}}_{t+1|t})'] = \mathbb{E}[(\mathbf{z}_t - \hat{\mathbf{z}}_{t|t})(A_1\mathbf{z}_t + \boldsymbol{\varepsilon}_{t+1} - A_1\hat{\mathbf{z}}_{t|t})']$$

[4] Actually, Hamilton (1994) uses this approach.

$$= \mathbf{P}_{t|t}A_1'$$

Thus,

$$\hat{\mathbb{E}}[\mathbf{z}_t|\mathbf{z}_{t+1}, \mathcal{F}_t] = \hat{\mathbf{z}}_{t|t} + \mathbf{J}_t[\mathbf{z}_{t+1} - \hat{\mathbf{z}}_{t+1|t}]$$

where $\mathbf{J}_t \equiv \mathbf{P}_{t|t}A_1'\mathbf{P}_{t+1|t}^{-1}$. Take a time to realize that

$$\hat{\mathbb{E}}[\mathbf{z}_t|\mathbf{z}_{t+1}, \mathcal{F}_T] = \hat{\mathbf{z}}_{t|t} + \mathbf{J}_t[\mathbf{z}_{t+1} - \hat{\mathbf{z}}_{t+1|t}]$$

and, finally, we get the recursive equation

$$\hat{\mathbf{z}}_{t|T} = \hat{\mathbf{z}}_{t|t} + \mathbf{J}_t[\hat{\mathbf{z}}_{t+1|T} - \hat{\mathbf{z}}_{t+1|t}] \tag{3.15}$$

i.e., to obtain $\{\hat{\mathbf{z}}_{t|T}\}_{t=1}^{T}$ we must proceed by backward induction. The MSE associated is

$$\mathbf{P}_{t|T} = \mathbf{P}_{t|t} + \mathbf{J}_t[\mathbf{P}_{t+1|T} - \mathbf{P}_{t+1|t}]\mathbf{J}_t' \tag{3.16}$$

3.5 A simple example

I have created the function `KalmanFilter.m` which makes all the computations presented here. I will use two different filters to obtained the HP cycle component of Peruvian GDP ($y_t \equiv 100\ln(Y_t)$). The first one, *state-space I*, follows directly the HP filter formulas and its state-space representation:

$$\begin{aligned} y_t &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \tau_t \\ \tau_{t-1} \end{bmatrix} + c_t \\ \begin{bmatrix} \tau_t \\ \tau_{t-1} \end{bmatrix} &= \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \tau_{t-1} \\ \tau_{t-2} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \varepsilon_t \end{aligned}$$

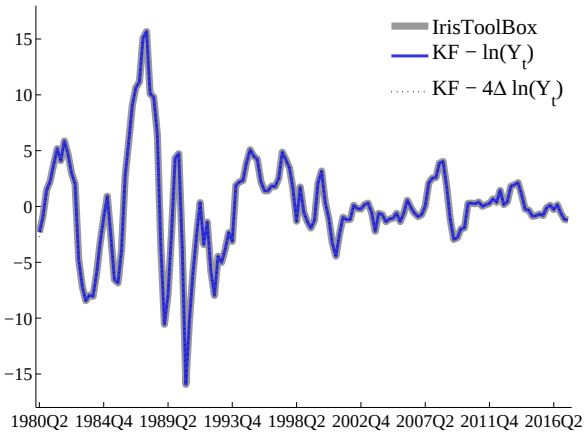
This is the traditional HP filter: $\lambda = \mathbb{V}(c_t)/\mathbb{V}(\varepsilon_t)$. The second one, *state-space II*, takes as observable variable the annualized percentage change of the output, i.e., $\Delta^a y_t \equiv 4\Delta y_t$, and its state-space representation is

$$\begin{aligned} \Delta^a y_t &= \begin{bmatrix} 1 & 4 & -4 \end{bmatrix} \begin{bmatrix} \tau_t \\ c_t \\ c_{t-1} \end{bmatrix} \\ \begin{bmatrix} \tau_t \\ c_t \\ c_{t-1} \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \tau_{t-1} \\ c_{t-1} \\ c_{t-2} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \boldsymbol{\varepsilon}_t \end{aligned}$$

Since these are non-stationary models I initialized the filter with a very large matrix $\mathbf{P}_{0|0}$. Both state-space models will lead to the same cycle component, as it is shown in Fig. 6, when there is an adequate calibration for the variance errors. One of the advantages of using Kalman filter is the we obtain a measure of the uncertainty about the filter or smoothed series, see Fig. 7 also.

Figure 6. *Kalmam filter: cycle component*

(a) Cycle comparisons



(b) State-space II

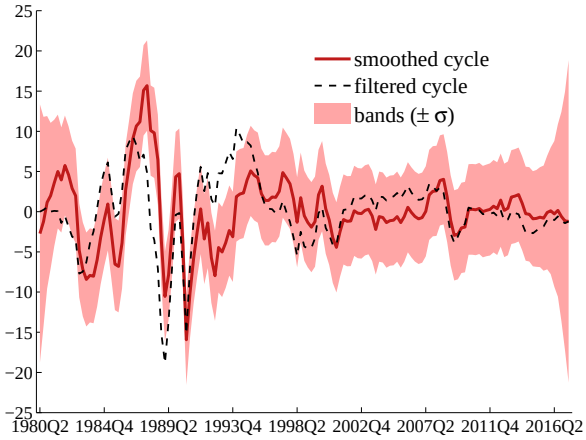
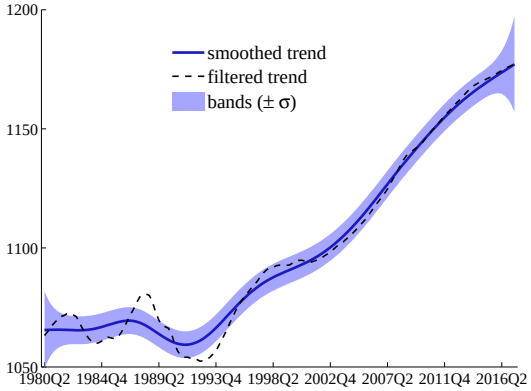
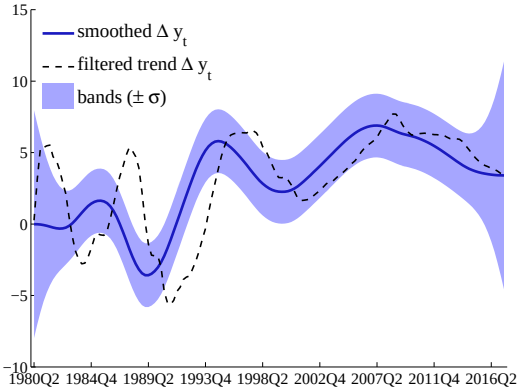


Figure 7. *Kalmam filter: permament component comparisons*

(a) State-space I: permanent component

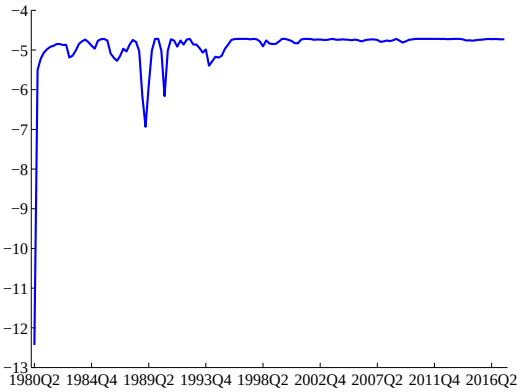


(b) State-space II: permanent component



(c) State-space I: likelihood

(d) State-space II: likelihood



4 Metropolis-Hastings algorithms for linear RE models

Now we will explore different alternatives to specify $q(.|.)$ in the MH algorithms. In general, we consider distributions of the form

$$q(.|\theta^{i-1}) = \mathbb{P}_t(.|\mu(\boldsymbol{\theta}^{i-1}), \Sigma(\boldsymbol{\theta}^{i-1}), \upsilon) \tag{4.1}$$

where $\mathbb{P}_t(.)$ refers to the density of a t-student distribution.

4.1 Random Walk Metropolis Hastings (RWMH) algorithm

The name of the algorithm comes from the random walk form of the proposal

$$\boldsymbol{\vartheta} = \boldsymbol{\theta}^{i-1} + \boldsymbol{\eta} \tag{4.2}$$

where $\boldsymbol{\eta}$ is mean zero with variance $\tau^2\hat{\Sigma}$. Given the symmetric nature of the proposal distribution: $q(\boldsymbol{\vartheta}|\boldsymbol{\theta}^{i-1}) = q(\boldsymbol{\theta}^{i-1}|\boldsymbol{\vartheta})$, then

$$\alpha = \min\left\{\frac{\mathbb{P}(\boldsymbol{\vartheta}|Y)}{\mathbb{P}(\boldsymbol{\theta}^{i-1}|Y)}, 1\right\} \tag{4.3}$$

For simplicity we can set $\upsilon = \infty$, i.e., we use a multivariate normal proposal distribution. Noting that the sampler can work very poorly if $q(.|.)$ is strongly at odds with the target distribution, then the variance of the proposal distribution $\hat{\Sigma}$ must incorporate information from the posterior, to potentially capture correlations between parameter posterior. We have two alternatives: i) following Schorfheide (2000), set $\hat{\Sigma}$ to be the negative of the inverse Hessian at the mode of the log posterior obtaining by running a numerical optimization; and ii) adaptive approach: first, generate a set of posterior draws based on a reasonable initial choice for $\hat{\Sigma}$, and second, compute the sample covariance matrix from the first sequence of posterior draws and use it as $\hat{\Sigma}$ in a second run of the RWMH algorithm.

The final parameter of the algorithm is the scaling factor τ and is typically adjusted to ensure a reasonable acceptance rate, $0.2 - 0.4$.

This algorithm works well when the target distribution is regular, i.e., elliptical form. There are two main sources of problems with this algorithm: i) local identification problems, i.e., posterior is flat in the direction of at least one of the parameters, and ii) global identification problems, i.e., bimodal posterior distributions due to potential observational equivalence (misspecification).

4.2 Block MH algorithm

Despite the careful choice of the proposal distribution $q(.|\boldsymbol{\theta}^{i-1})$ it is natural that the efficiency of the MH algorithm decreases as dimension of the parameter vector increases. One way to alleviate this problem is to break the parameter vector into blocks. Suppose the dimension of the parameter vector $\boldsymbol{\theta}$ is d, then let make a partition of the parameter space, Θ, in a N_b blocks.

Algorithm 3 (Block MH algorithm) ***Draw*** $\boldsymbol{\theta}^0 \in \Theta$ ***and then for*** $i = 1$ ***to*** N*:*

- *Create a partition of the parameter vector into N_b via some rule (perhaps probabilistic), unrelated to the current state of the Markov chain.*
- *For $b = 1, \ldots, N_b$:*
  1. *Draw $\boldsymbol{\vartheta}_b \sim q(.|[\boldsymbol{\theta}^i_{<b}, \boldsymbol{\theta}^{i-1}_b, \boldsymbol{\theta}^{i-1}_{>b}])$.*
  2. *With probability,*

$$\alpha = \max\left\{\frac{\mathbb{P}([\boldsymbol{\theta}^i_{<b}, \boldsymbol{\vartheta}_b, \boldsymbol{\theta}^{i-1}_{>b}]|Y) q(\boldsymbol{\theta}^i_b|[\boldsymbol{\theta}^i_{<b}, \boldsymbol{\vartheta}_b, \boldsymbol{\theta}^{i-1}_{>b}])}{\mathbb{P}([\boldsymbol{\theta}^i_{<b}, \boldsymbol{\theta}^{i-1}_b, \boldsymbol{\theta}^{i-1}_{>b}]|Y) q(\boldsymbol{\vartheta}_b|[\boldsymbol{\theta}^i_{<b}, \boldsymbol{\theta}^{i-1}_b, \boldsymbol{\theta}^{i-1}_{>b}])}, 1\right\} \tag{4.4}$$

  set $\boldsymbol{\theta}^i_b = \boldsymbol{\vartheta}_b$, otherwise set $\boldsymbol{\theta}^i_b = \boldsymbol{\theta}^{i-1}_b$.

A good rule of thumb is that we want the parameters within a block to be as correlated as possible while we want the parameters between blocks to be "as independent as possible".

Algorithm 4 (Random-Block MH algorithm)

- *Create a sequence of random partitions* $\{\Theta^i\}_i$ *of the parameter vector into* N_b *equally sized blocks as follow:*
  1. *assign an iid* $U[0,1]$ *draw to each element of* $\boldsymbol{\theta}$*;*
  2. *sort the parameters according to the assigned random number;*
  3. *let b'th block consist of parameters* $(b-1)N_b, \ldots, bN_b$*.*
- *Execute algorithm 3.*

4.3 Initializing the algorithm: the posterior mode

In order to initializing any algorithm presented here, is common to compute the mode of the posterior distribution. For sake of this objective, we must use a numeric optimization algorithm. The scripts I have coded uses `csminwel.m` function to obtain the prior mode which at the same time is implement by Dynare specifying the option `mode_compute = 4` in the `estimation` command. Then, the Hessian at the mode of the log posterior is taken as $\hat{\Sigma}$.

5 Application: An small open economy model

5.1 Model

The following model was proposed to work with for an empirical homework in the subject of *Macroeconomics II*.

5.1.1 Households

The households face the problem:

$$\max_{\{C_t,L_t,B_{t+1}\}_t} \sum_{t=0}^{\infty}\sum_{s^t}\beta^t\Big\{\Gamma_t(s^t)\Big(\frac{[C_t(s^t)-h\overline{C}_{t-1}(s^{t-1})]^{1-\sigma}}{1-\sigma}-\frac{L_t(s^t)^{1+\varphi}}{1+\varphi}\Big)\Big]\mathbb{P}(s^t)$$
$$+lm_t(s^t)\Big[B_t(s^t)+W_t(s^t)L_t(s^t)+\int_0^1\Pi_t^L(s^t)(i)di-P_t(s^t)C_t(s^t)$$
$$-\sum_{s_{t+1}}Q_{t,t+1}(s_{t+1}|s^t)B_{t+1}(s_{t+1},s^t)\mathbb{P}(s_{t+1},s^t|s^t)\Big]\Big\}$$

where $lm_t(s^t)$ is the lagrange multiplier for each restriction. The first order conditions for this problem are

$$\begin{aligned}C_t(s^t):&\quad \Gamma_t(s^t)[C_t(s^t)-h\overline{C}_{t-1}(s^{t-1})]^{-\sigma}\mathbb{P}(s^t)-lm_t(s^t)P_t(s^t)=0\\ B_{t+1}(s_{t+1},s^t):&\quad -lm_t(s^t)Q_{t,t+1}(s_{t+1}|s^t)\mathbb{P}(s_{t+1},s^t|s^t)+\beta lm_{t+1}(s_{t+1},s^t)=0\\ L_t(s^t):&\quad -\Gamma(s^t)L_t(s^t)^{\varphi}\mathbb{P}(s^t)+lm_t(s^t)W_t(s^t)=0\end{aligned}$$

In terms of the traditional notation (and more useful here), the equilibrium conditions are:

$$\frac{W_t}{P_t}=L_t^{\varphi}[C_t-hC_{t-1}]^{\sigma} \tag{5.1}$$
$$\frac{1}{1+i_t}=\beta\mathbb{E}\Big[\Big(\frac{C_t-hC_{t-1}}{C_{t+1}-hC_t}\Big)^{\sigma}\frac{P_t}{P_{t+1}}\omega_t\Big] \tag{5.2}$$

where $\omega_t=\frac{\Gamma_{t+1}}{\Gamma_t}$ and it was imposed the consistence condition $C_t=\overline{C}_t$. Moreover, the demand for the ith is given by:

$$C_t(i)=\Big[\frac{P_t(i)}{P_t}\Big]^{-\theta_t}C_t \tag{5.3}$$
$$C_t=\Big(\int_0^1 C_t(i)^{(\theta_t-1)/\theta_t}di\Big)^{\frac{\theta_t}{\theta_t-1}} \tag{5.4}$$
$$P_t=\Big(\int_0^1 P_t(i)^{1-\theta_t}di\Big)^{\frac{1}{1-\theta_t}} \tag{5.5}$$

5.1.2 Firms

Cost minimization of firms lead to the following equation for real marginal costs:

$$MC_t=\frac{1}{A_t}\Big[\frac{W_t/P_t}{1-\delta}\Big]^{1-\delta}\Big[\frac{RER_t}{\delta}\Big]^{\delta} \tag{5.6}$$

Price setting is as follows: a fraction λ of firms set prices at t while the other fraction only readjust its previous price with the following rule

$$P_t(i)=P_{t-1}(i)\Big(\frac{P_{t-1}}{P_{t-2}}\Big)^{\gamma}$$

Let $\Pi_{t-1,k} \equiv P_{t+k-1}/P_{t-1}$ and equal to $\Pi_{t-1,1}) \equiv \Pi_t$. Hence, profit maximization problem is:

$$\max_{P_t(i)} \sum_{k=0}^{\infty} (1-\lambda)^k \mathbb{E}_t \Big[Q_{t,t+k} \Pi^L_{t+k}(P_t(i), P_{t+k}, C_{t+k}) \Big]$$

where $\Pi^L_{t+k}(p, P, C) = [\tilde{p}_k - \mathcal{C}'_{t+k}] D_{t+k}(i, C)$ where $D_{t+k}(i, C) = \left[\frac{\tilde{p}_k}{P} \right]^{-\theta_{t+k}} C$, and $\tilde{p}_k = p \Pi^{\gamma}_{t-1,k}$. Then the first order condition for the firm is

$$\sum_{k=0}^{\infty} (1-\lambda)^k \mathbb{E}_t \Big[Q_{t,t+k} \frac{\partial}{\partial P_t(i)} \Pi^L_{t+k}(P^*_t(i), P_{t+k}, C_{t+k}) \Big] = 0 \tag{5.7}$$

where

$$\begin{aligned} \frac{\partial}{\partial P_t(i)} \Pi^L_{t+k}(P^*_t(i), P_{t+k}, C_{t+k}) =& (1-\theta_{t+k}) \Big(\frac{P_t}{P^*_t(i)} \Big)^{1+\theta_{t+k}} \\ & \times \Pi^{1+\theta_{t+k}}_{t,k} \Pi^{-\theta_{t+k}\gamma}_{t-1,k} C_{t+k} \Big[\frac{P^*_t(i)}{P_t} \frac{\Pi^{\gamma}_{t-1,k}}{\Pi_{t,k}} - \frac{\theta_{t+k}}{\theta_{t+k}-1} MC_{t+k} \Big] \end{aligned} \tag{5.8}$$

5.1.3 Aggregation

Aggregate price index is given by the following recursive equation:

$$P^{1-\theta_t}_t = \lambda P^{*1-\theta_t}_t + (1-\lambda) \Pi^{(1-\theta_t)\gamma}_{t-2,1} P^{1-\theta_t}_{t-1} \tag{5.9}$$

If we integrate production function and use market clearing condition $Y_t(i) = C_t(i)$:

$$\Delta_t Y_t = A_t L^{1-\delta}_t Z^{\delta}_t \tag{5.10}$$

where

$$\Delta_t = \int_0^1 \Big[\frac{P_t(i)}{P_t} \Big]^{-\theta_t} di \approx 1 + \int_0^1 [-\theta(\hat{p}_t(i) - \hat{p}_t) - \ln(\overline{P}_t(i)/\overline{P}_t)\hat{\theta}_t] di = 1$$

Moreover, in order to express marginal cost in terms of real exchange rate and output, I derive aggregate factor demands:

$$Z_t \equiv \int_0^1 Z_t(i) di = \delta \frac{MC_t}{RER_t} \Delta_t Y_t \tag{5.11}$$

$$L_t \equiv \int_0^1 L_t(i) di = (1-\delta) \frac{MC_t}{W_t/P_t} \Delta_t Y_t \tag{5.12}$$

5.2 Log-linearized version

After some tedious algebra we can get the linear version of the model described above.

$$(1+h)\hat{y}_t = \mathbb{E}_t \hat{y}_{t+1} + h\hat{y}_{t-1} - (1-h)\sigma^{-1}[\hat{i}_t - \mathbb{E}_t \hat{\pi}_{t+1} + \hat{\omega}_t] \tag{5.13}$$

$$\hat{mc}_t = \zeta[\hat{y}_t - \rho_y \hat{y}_{t-1}] - \frac{1+\varphi}{1+\varphi\delta}\hat{a}_t + \frac{(1+\varphi)\delta}{1+\varphi\delta} r\hat{e}r_t \tag{5.14}$$

$$\hat{\pi}_t = \Lambda \hat{mc}_t + \beta \mathbb{E}_t[\hat{\pi}_{t+1} - \gamma\hat{\pi}_t] - \Theta\hat{\theta}_t + \gamma\hat{\pi}_{t-1} \tag{5.15}$$

$$\hat{i}_t = \hat{i}^f_t + \mathbb{E}_t \Delta\hat{s}_{t+1} + \hat{\phi}_t - \chi r\hat{e}r_t \tag{5.16}$$

$$\begin{aligned} \hat{i}_t = \rho_i \hat{i}_{t-1} + (1-\rho_i) \Big[& \phi_\pi \mathbb{E}_t \Big\{ \hat{\pi}_t + \hat{\pi}_{t+1} + \hat{\pi}_{t+2} \Big\} \\ & + \phi_c \hat{y}_t + \phi_{\Delta c}(\hat{y}_t - \hat{y}_{t-4}) + \phi_s(\Delta\hat{s}_t + \Delta\hat{s}_{t-1}) \Big] + \hat{v}_t \end{aligned} \tag{5.17}$$

$$\hat{rer}_t = \hat{rer}_{t-1} + \Delta\hat{s}_t + \hat{\pi}^f_t - \hat{\pi}_t \tag{5.18}$$
$$\hat{\pi}^f_t = \rho_\pi \hat{\pi}^f_{t-1} + \sigma_\pi \varepsilon^\pi_t \tag{5.19}$$
$$\hat{i}^f_t = \rho_{if} \hat{i}^f_{t-1} + \sigma_{if} \varepsilon^{if}_t \tag{5.20}$$
$$\hat{\omega}_t = \rho_\omega \hat{\omega}_{t-1} + \sigma_\omega \varepsilon^\omega_t \tag{5.21}$$
$$\hat{\theta}_t = \rho_\theta \hat{\theta}_{t-1} + \sigma_\theta \varepsilon^\theta_t \tag{5.22}$$
$$\hat{\phi}_t = \rho_\phi \hat{\phi}_{t-1} + \sigma_\phi \varepsilon^\phi_t \tag{5.23}$$
$$\hat{a}_t = \rho_a \hat{a}_{t-1} + \sigma_a \varepsilon^a_t \tag{5.24}$$
$$\hat{v}_t = \rho_v \hat{v}_{t-1} + \sigma_v \varepsilon^v_t \tag{5.25}$$

where $\zeta = \frac{(\varphi(1-h)+\sigma)(1-\delta)}{(1-h)(1+\varphi\delta)}$, $\rho_y = \frac{\sigma h}{\varphi(1-h)+\sigma}$, $\Lambda = \frac{\lambda[1-(1-\lambda)\beta]}{1-\lambda}$, and $\Theta = \frac{\Lambda}{\theta-1}$.

5.3 Linear system representation

In **??** I expose how to use the created own codes. One of the first steps is to represent the model as was suggested by Sims (2002):

$$\Gamma_0(\boldsymbol{\theta})\mathbf{z}_{t+1} = \Gamma_1(\boldsymbol{\theta})\mathbf{z}_t + \Psi_\varepsilon(\boldsymbol{\theta})\boldsymbol{\varepsilon}_{t+1} + \Pi\boldsymbol{\eta}_{t+1} \tag{5.26}$$

First, it will be useful to make some arrangements to the log-linearized system:

$$\begin{aligned}
&-(1+h)\hat{y}_{t+1} - \frac{(1-h)}{\sigma}\hat{i}_{t+1} + \mathbb{E}_{t+1}\hat{y}_{t+2} + \frac{(1-h)}{\sigma}\mathbb{E}_{t+1}\hat{\pi}_{t+2} - \frac{(1-h)}{\sigma}\hat{\omega}_{t+1} = -h\hat{y}_t \\
&-\zeta\hat{y}_{t+1} + \hat{mc}_{t+1} + \frac{1+\varphi}{1+\varphi\delta}\hat{a}_{t+1} - \frac{(1+\varphi)\delta}{1+\varphi\delta}\hat{rer}_{t+1} = -\zeta\rho_y\hat{y}_t \\
&(1+\beta\gamma)\hat{\pi}_{t+1} - \Lambda\hat{mc}_{t+1} - \beta\mathbb{E}_{t+1}[\hat{\pi}_{t+2}] + \Theta\hat{\theta}_{t+1} = \gamma\hat{\pi}_t \\
&\Delta\hat{s}_{t+1} = \hat{i}_t - \hat{i}^f_t + \chi\hat{rer}_t - \hat{\phi}_t + \eta^s_{t+1} \\
&(1-\rho_i)\phi_\pi\mathbb{E}_{t+1}\hat{\pi}_{t+2} = \hat{i}_t - \rho_i\hat{i}_{t-1} - (1-\rho_i)\Big[\phi_\pi\mathbb{E}_t[\hat{\pi}_t + \hat{\pi}_{t+1}] + \phi_c\hat{y}_t + \phi_{\Delta c}(\hat{y}_t - \hat{y}_{t-4}) \\
&\qquad\qquad + \phi_s(\Delta\hat{s}_t + \Delta\hat{s}_{t-1})\Big] - \hat{v}_t + (1-\rho_i)\phi_\pi\eta^{\mathbb{E}\pi}_{t+1} \\
&\hat{rer}_{t+1} - \Delta\hat{s}_{t+1} - \hat{\pi}^f_{t+1} + \hat{\pi}_{t+1} = \hat{rer}_t \\
&\hat{\pi}^f_{t+1} = \rho_\pi\hat{\pi}^f_t + \sigma_\pi\varepsilon^\pi_{t+1} \\
&\hat{i}^f_{t+1} = \rho_{if}\hat{i}^f_t + \sigma_{if}\varepsilon^{if}_{t+1} \\
&\hat{\omega}_{t+1} = \rho_\omega\hat{\omega}_t + \sigma_\omega\varepsilon^\omega_{t+1} \\
&\hat{\theta}_{t+1} = \rho_\theta\hat{\theta}_t + \sigma_\theta\varepsilon^\theta_{t+1} \\
&\hat{\phi}_{t+1} = \rho_\phi\hat{\phi}_t + \sigma_\phi\varepsilon^\phi_{t+1} \\
&\hat{a}_{t+1} = \rho_a\hat{a}_t + \sigma_a\varepsilon^a_{t+1} \\
&\hat{v}_{t+1} = \rho_v\hat{v}_t + \sigma_v\varepsilon^v_{t+1} \\
&\hat{y}_{t+1} = \mathbb{E}_t\hat{y}_{t+1} + \eta^y_{t+1} \\
&\hat{\pi}_{t+1} = \mathbb{E}_t\hat{\pi}_{t+1} + \eta^\pi_{t+1} \\
&\Delta\hat{s}_{t+1} = \mathbb{E}_t\Delta\hat{s}_{t+1} + \eta^s_{t+1} \\
&\mathbb{E}_{t+1}\hat{\pi}_{t+2} = \mathbb{E}_t\hat{\pi}_{t+2} + \eta^{\mathbb{E}\pi}_{t+1}
\end{aligned}$$

Let

$$\mathbf{z}_t = (\hat{y}_t, \hat{\pi}_t, \hat{i}_t, \hat{mc}_t, \hat{rer}_t, \Delta\hat{s}_t, \mathbb{E}_t[\hat{y}_{t+1}], \mathbb{E}_t[\hat{\pi}_{t+1}], \hat{y}_{t-1}, \hat{y}_{t-2}, \hat{y}_{t-3}, \hat{y}_{t-4}, \hat{i}_{t-1}, \Delta\hat{s}_{t-1}, \hat{\pi}^f_t, \hat{i}^f_t, \hat{\omega}_t, \hat{\theta}_t, \hat{\phi}_t, \hat{a}_t, \hat{v}_t)',$$

$\boldsymbol{\varepsilon}_t = (\varepsilon_t^{\pi}, \varepsilon_t^{i^f}, \varepsilon_t^{\omega}, \varepsilon_t^{\theta}, \varepsilon_t^{\phi}, \varepsilon_t^{a}, \varepsilon_t^{v})'$, and $\boldsymbol{\eta}_t = (\eta_t^{y}, \eta_t^{\pi}, \eta_t^{s}.\eta_t^{\mathbb{E}\pi})'$, then

$$\Psi_{\varepsilon}(\boldsymbol{\theta}) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \sigma_{\pi} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{if} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{\omega} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{\theta} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{\phi} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{a} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{v} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\Pi = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & (1-\rho_i)\phi_{\pi} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}'$$

$$
\Gamma_0(\boldsymbol{\theta}) = \begin{bmatrix}
-(1+h) & 0 & -\frac{1-h}{\sigma} & 0 & 0 & 0 & 1 & \frac{1-h}{\sigma} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1-h}{\sigma} & 0 & 0 & 0 & 0 \\
-\zeta & 0 & 0 & 1 & -\frac{(1+\varphi)\delta}{1+\varphi\delta} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1+\varphi}{1+\varphi\delta} & 0 \\
0 & (1+\beta\gamma) & 0 & -\Lambda & 0 & 0 & 0 & -\beta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \Theta & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & (1-\rho_i)\phi_\pi & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix} \tag{5.27}
$$

$$
\Gamma_1(\boldsymbol{\theta}) = \begin{bmatrix}
-h & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-\zeta\rho_y & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \gamma & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & \chi & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 \\
-\frac{(1-\rho_i)}{(\phi_c+\phi_{\Delta c})^{-1}} & -\frac{(1-\rho_i)}{\phi_\pi^{-1}} & 1 & 0 & 0 & -\frac{(1-\rho_i)}{\phi_s^{-1}} & 0 & -\frac{(1-\rho_i)}{\phi_\pi^{-1}} & 0 & 0 & 0 & \frac{(1-\rho_i)}{\phi_{\Delta c}^{-1}} & -\rho_i & -\frac{(1-\rho_i)}{\phi_s^{-1}} & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho_\pi & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho_{if} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho_\omega & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho_\theta & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho_\phi & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho_a & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho_v \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix} \tag{5.28}
$$

The function `sims_basic_model.m` takes as argument the parameter vector (those which will be estimated) and constructs this matrices.

5.4 Bayesian estimation

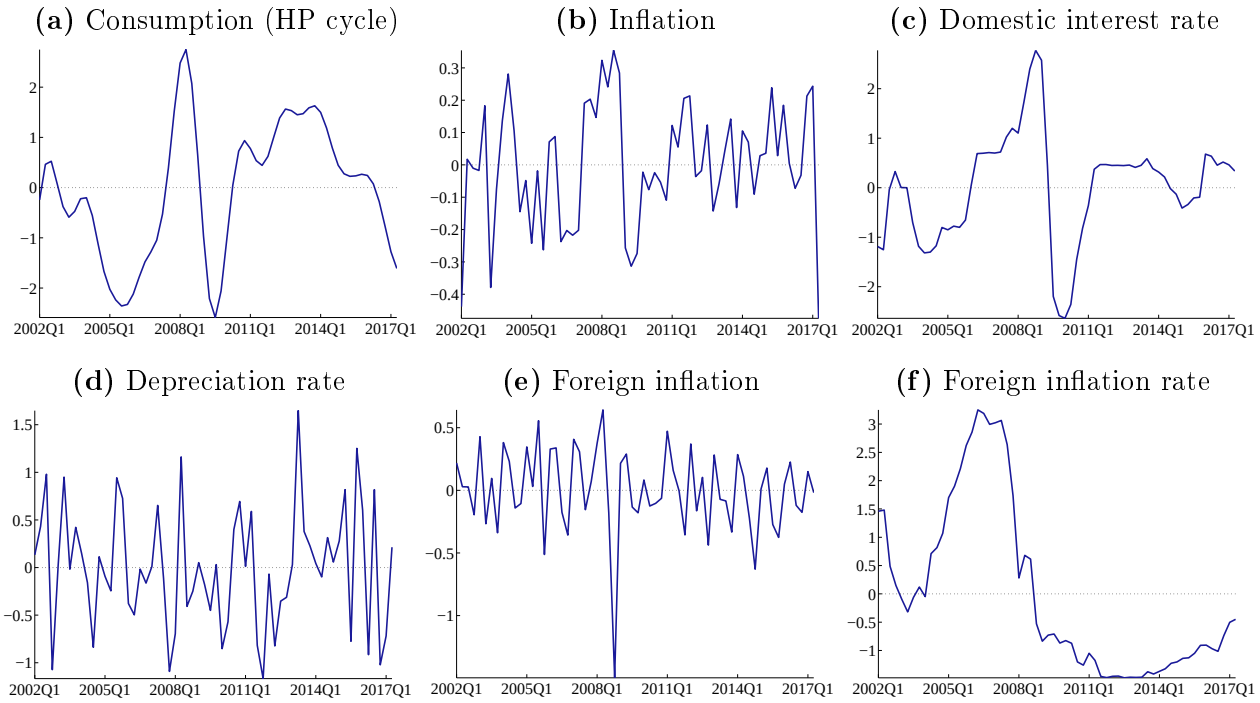
Bayesian estimation is perform using Random Walk Metropolis-Hastings algorithm. Both Dynare toolbox and own codes are used in estimation. **??** gives some details about the code employed for purpose. The observable variables are:

Table 1. *Observable variables and notation*

Variable	Model	Model notation
Quarterly Private Consumption	Consumption	$\hat{c}_t^{obs}$
Quarterly CPI inflation	Inflation	$\hat{\pi}_t^{obs}$
Annualized Interbank interest rate	Interest rate	$\hat{i}_t^{obs}$
Quarterly depreciation rate	Depreciation rate	$\Delta\hat{s}_t^{obs}$
Business partners inflation rate	Foreign inflation rate	$\hat{\pi}_t^{f,obs}$
Annualized yield two-years treasury bonds	Foreign interest rate	$\hat{i}_t^{f,obs}$

Figure 8 plots the observable series. It is evident that foreign interest rate will cause a difficulty because after international financial crisis this series seems to be at a lower bound. Later I will try to solve this problem based on Canova (2008).

Figure 8. *Data*



Source: Central Bank of Peru.

5.4.1 Priors

Prior distribution are specified following Castillo et al. (2006) and Herbst and Schorfheide (2015). There is only two parameter I maintain calibrated: $\beta = 0.99$ and $\chi = 1e-3$. The first one is commonly calibrated and guarantee an annualized interest rate steady-state of four percent, while the later is not estimated because draws over any distribution could problematic for this parameter. Table 2 show prior distributions for the rest of the variables.

Table 2. *Prior distribution*

P.	Description	Distribution	Mean	S.D	*Based on*
h	*habit formation*	BETA	0.7	0.1	Castillo et al. (2006)
σ	*Risk aversion coef.*	GAMMA	2	0.5	Herbst and Schorfheide (2015)
φ	*Inv. Frish elasticity*	InvGAMMA	1	0.3	Castillo et al. (2006)
δ	*Imported input participation*	UNIF*	0	1	No information
γ	*Degree of inflation indexation*	BETA	0.5	0.1	Castillo et al. (2006)
θ	*s.s. elasticity of substitution*	NORMAL	6	1	Castillo et al. (2006)
λ	*Degree of price flexibility*	BETA	0.66	0.1	Castillo et al. (2006)
ρ_i	*Taylor rule: smoothing*	BETA	0.5	0.2	Castillo et al. (2006)
ϕ_π	*Taylor rule: inflation*	NORMAL	1.5	0.3	Castillo et al. (2006)
ϕ_s	*Taylor rule: depreciation*	NORMAL	0.5	0.1	Castillo et al. (2006)
ϕ_c	*Taylor rule: output*	BETA	0.5	0.2	Castillo et al. (2006)
$\phi_{\Delta c}$	*Taylor rule: growth*	BETA	0.5	0.2	Castillo et al. (2006)
ρ_x	*X's persistence*	BETA	0.5	0.2	-
σ_x	*X's volatility*	InvGAMMA	0.4	0.3	-

Note. P.: parameter. * : Specify by bounds. *S.D.* : standard deviation. *s.s.* : Steady state.

5.4.2 Estimation and Results

The first exercise does not use measurement equations because there are seven stochastic shocks and six observable variable, so there is not a problem of stochastic singularity:

$$\hat{c}_t^{obs} = \hat{c}_t \tag{5.29}$$
$$\hat{\pi}_t^{obs} = \hat{\pi}_t \tag{5.30}$$
$$\hat{i}_t^{obs} = 4\hat{i}_t \tag{5.31}$$
$$\Delta\hat{s}_t^{obs} = \Delta\hat{s}_t \tag{5.32}$$
$$\hat{\pi}_t^{f,obs} = \hat{\pi}_t^f \tag{5.33}$$
$$\hat{i}_t^{f,obs} = 4\hat{i}_t^f \tag{5.34}$$

Since this is small scale model, some matrix representation are more or less manageable, then I estimate this model using own codes: `RMWH_dsge.m` and `KalmanFilter.m` functions are important in estimation and further details about this functions can be found in **??**. Additionally, there are three complement model dependent functions: `sims_basic_model.m`, `measure_eq_basic.m`, and `priors.m`, and they permit both state-space representation and posterior computation. For both estimations 100000 draws are generated using Random Walk Metropolis Hastings algorithm (RWMH). Sampling are initiated at the posterior mode computed by `csminwel.m` [implemented by Dynare with `mode_compute=4` option] function and for both exercises the acceptance ratio oscillates around a reasonable interval: $[23\% - 33\%]$.

Results are shown in Tab. 3. We can note that most of the estimations for mean, mode, and ninety percent confidence interval are very similar. There are many reasons why estimations are not equal: a)

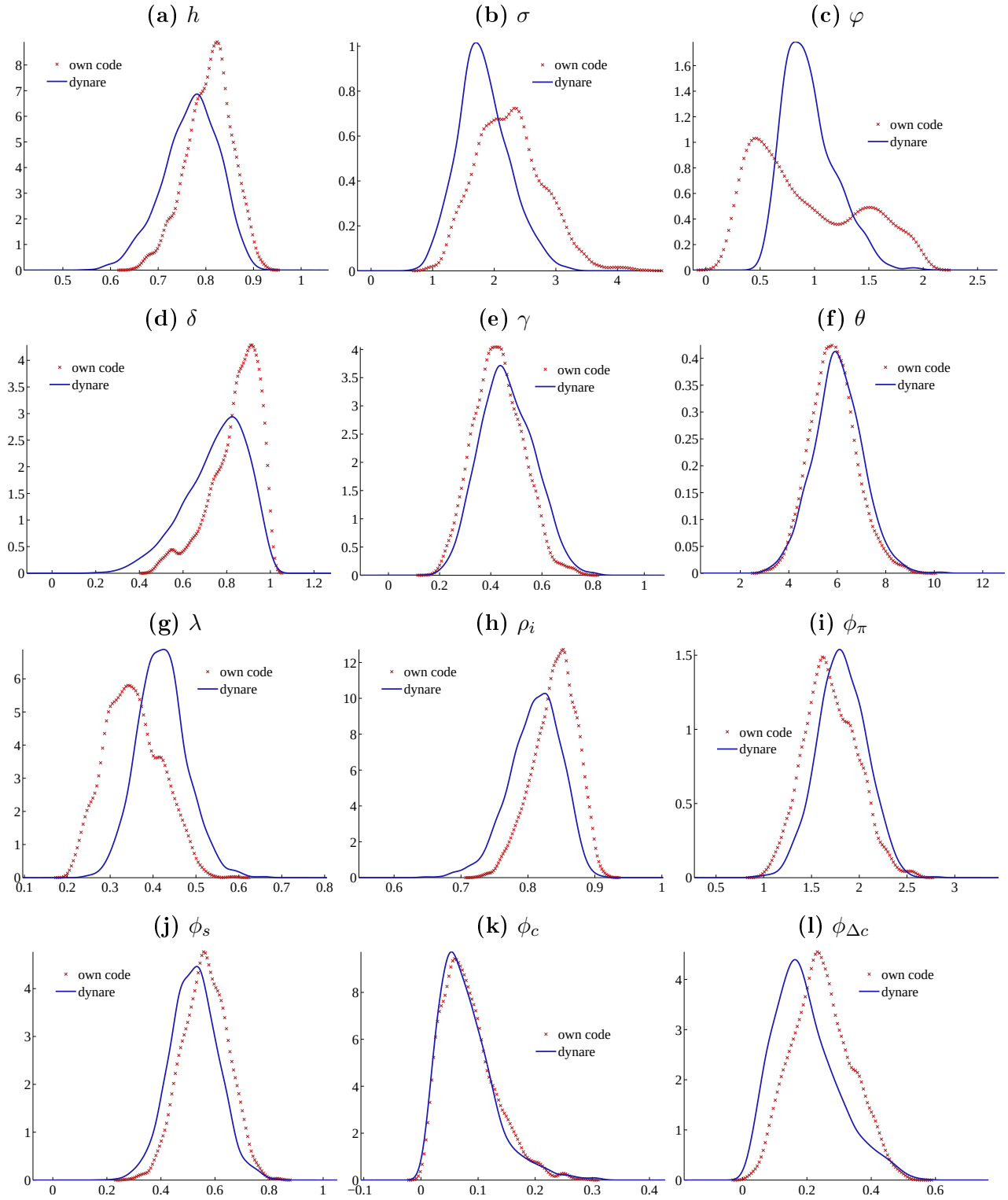
Table 3. *Posterior moments estimation*

Parameter	Posterior: Dynare Toolbox				Posterior: Own code			
	Mode	*Mean*	*5%*	*95%*	*Mode*	*Mean*	*5%*	*95%*
h	0.774	0.768	0.681	0.872	0.797	0.807	0.720	0.882
σ	1.655	1.825	1.107	2.488	1.920	2.256	1.417	3.184
φ	0.831	0.954	0.572	1.317	0.444	0.923	0.302	1.816
δ	0.743	0.750	0.520	0.963	0.837	0.841	0.613	0.977
γ	0.449	0.464	0.297	0.635	0.431	0.429	0.276	0.591
θ	6.002	6.019	4.389	7.684	6.003	5.835	4.314	7.527
λ	0.436	0.420	0.324	0.513	0.406	0.354	0.249	0.467
ρ_i	0.810	0.810	0.749	0.870	0.833	0.840	0.780	0.889
ϕ_π	1.779	1.815	1.411	2.250	1.711	1.702	1.253	2.168
ϕ_s	0.501	0.521	0.385	0.666	0.522	0.563	0.426	0.701
ϕ_c	0.052	0.081	0.009	0.144	0.050	0.084	0.023	0.173
$\phi_{\Delta c}$	0.159	0.196	0.029	0.342	0.217	0.247	0.100	0.409
ρ_π	0.067	0.087	0.015	0.156	0.069	0.086	0.025	0.173
ρ_{if}	0.866	0.856	0.805	0.907	0.881	0.878	0.831	0.918
ρ_ω	0.807	0.790	0.695	0.886	0.792	0.768	0.663	0.865
ρ_θ	0.497	0.496	0.193	0.855	0.497	0.509	0.165	0.847
ρ_ϕ	0.766	0.735	0.631	0.836	0.779	0.734	0.623	0.844
ρ_a	0.926	0.905	0.829	0.990	0.903	0.828	0.634	0.961
ρ_v	0.320	0.339	0.190	0.508	0.280	0.320	0.160	0.478
σ_π	0.336	0.348	0.298	0.403	0.336	0.344	0.294	0.400
σ_{if}	0.114	0.117	0.099	0.134	0.098	0.102	0.088	0.120
σ_ω	0.439	0.533	0.273	0.821	0.609	0.919	0.407	1.662
σ_θ	0.251	0.363	0.146	0.575	0.192	0.399	0.116	0.934
σ_ϕ	0.223	0.249	0.184	0.315	0.222	0.268	0.192	0.349
σ_a	0.777	0.829	0.634	1.012	0.878	1.097	0.739	1.655
σ_v	0.173	0.183	0.137	0.225	0.149	0.159	0.123	0.204
Log-Likelihood	-173.1370				-169.8048			
Log-Posterior	-183.0793				-180.6766			

Numerical algorithms implemented by Dynare works with higher precision than mines, and b) there can be other improvements in the basic RWMH algorithm implemented by Dynare. In any case, I consider that RMWH_dsge.m works reasonable well. To given a complete evaluation of my own code performance, Fig. 9 plots the posterior distributions.

An interesting result is that estimation for δ seems to be "too high". I consider that it is a consequence of model misspecification: Peru is characterized by the presence of partial dollarization (see, Castillo et al. (2009)): a) price dollarization, in which case prices are indexed to changes in the exchange rate, b) transaction dollarization, also known as currency substitution, in which dollars are accepted as the medium of payment, and c) financial dollarization, in which case dollars are preferred as a store of value. Hence, the effects of movements in depreciation rate over inflation are all captured by the marginal cost channel and then attributed to the specific production function.

Figure 9. *Posterior distributions*



References

Blanchard, O. and Khan, C. M. (1980). The solution of linear difference models under Rational expectations. *Econometrica*, pages 1305–1312.

Canova, F. (2008). Bridging DSGE models and the raw data. Economics Working Papers 1320, Department of Economics and Business, Universitat Pompeu Fabra.

Castillo, P., Montoro, C., and Tuesta, V. (2006). An Estimated Stochastic General Equilibrium Model with Partial Dollarization: A Bayesian Approach. Working Papers Central Bank of Chile 381, Central Bank of Chile.

Castillo, P., Montoro, C., and Tuesta, V. (2009). A Dynamic Stochastic General Equilibrium Model with Dollarization for the Peruvian Economy. Working Papers 2009-003, Banco Central de Reserva del Perú.

Galí, J. (2008). *Monetary policy, inflation and the Business cycle: an introduction to the New Keynesian framework*. Princeton University Press.

Hamilton, J. (1994). *Time Series Analysis*. Princeton University Press.

Herbst, E. and Schorfheide, F. (2015). *Bayesian Estimation of DSGE Models*. The Econometric and Tinbergen Institutes Lectures. Princeton University Press.

Klein, P. (2000). Using the generalized Schur form to solve a multivariate linear rational expectations model. *Journal of economic dynamics and control*, pages 1405–1423.

Schorfheide, F. (2000). Loss function-based evaluation of DSGE models. *Journal of Applied Econometrics*, 15(6):645–670.

Sims, C. (2002). Solving linear rational expectations models. *Computational Economics*, 20(1-2):1–20.

Uhlig, H. (1995). A toolkit for analyzing nonlinear dynamic stochastic models easily. Discussion Paper / Institute for Empirical Macroeconomics 101, Federal Reserve Bank of Minneapolis.