# A crash course of Bayesian DSGE estimation

# II. Some applications

Takeki Sunakawa
Hitotsubashi University
May 19, 2022 @ Keio University

# What will be covered today

- Random-walk MH algorithm and its application to New-Keynesian model (HS, ch. 4)
- Solving real business cycle model with Dynare (Miao, chs. 14, 15)

# A bird's-eye view: How it works?

- Taking a set of model parameters $\theta$ as given, put the REE solution and observation equation together to form a state-space representation.

$$y_t = A(\theta) + B(\theta)x_t + e_t, \qquad e_t \sim N(0, H(\theta))$$
$$x_t = P(\theta)x_{t-1} + Q(\theta)\epsilon_t, \qquad \epsilon_t \sim N(0, S_e(\theta))$$

- Then having the data $Y$, we calculate the likelihood function of the parameters, $L(Y|\theta)$, from the state-space representation using the Kalman filter.

- We conjecture some form of the prior distribution of the parameters, $p(\theta)$.

- Using the Bayes' theorem, we have the posterior distribution of the parameters.

$$p(\theta|Y) \propto p(\theta)L(Y|\theta)$$

- We use the Metropolis-Hasting algorithm, a Monte-Carlo sampling method, to approximate the shape of the posterior distribution.

- We do inferences based on the posterior distribution.

# RWMH algorithm

# Random-Walk MH algorithm

- The most widely used method to generate draws from posterior distributions of a DSGE model.
- The proposal distribution $q\left(\vartheta \middle| \theta^{i-1}\right)$ is expressed as the random walk

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

where $\eta$ is drawn from a distribution with mean zero.

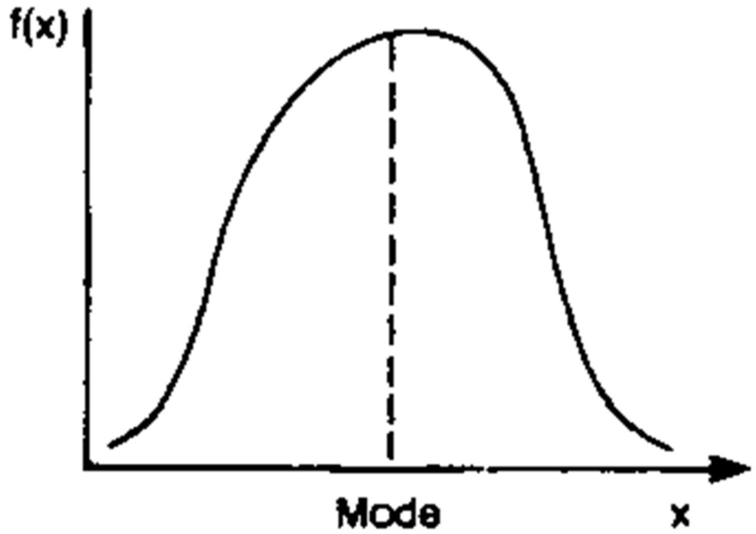
- With a prior distribution typically used, the posterior distribution (hopefully) has a well-behaved elliptical shape.

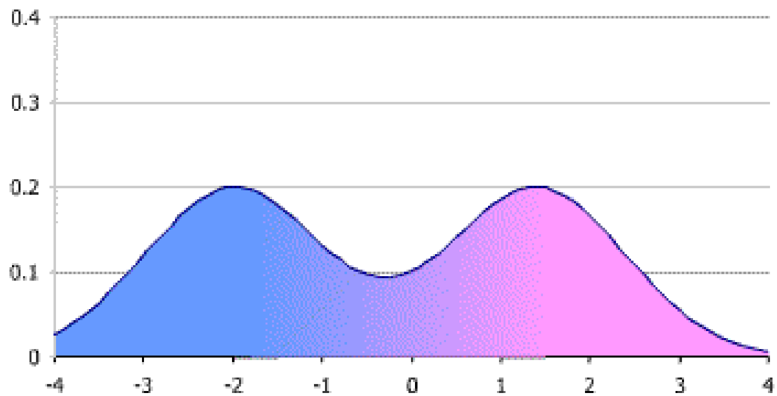
# Identification problems

- The posterior distributions could be very non-elliptical.
- This is due to identification problems
    - *Very flat* posteriors:

      Local identification problems (see the earlier example)
    - *Multimodal* posteriors:

      Inflation persistence can be generated by either persistent cost-push shocks or firm’s inability to re-optimize prices.



# Generic MH algorithm

(**Generic MH Algorithm**) For $i = 1, \dots, N$:

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

$$\alpha(\vartheta|\theta^{i-1}) = \min\left\{1, \frac{p(Y|\vartheta)p(\vartheta)/q(\vartheta|\theta^{i-1})}{p(Y|\theta^{i-1})p(\theta^{i-1})/q(\theta^{i-1}|\vartheta)}\right\}$$

and $\theta^i = \theta^{i-1}$ otherwise, where

$p(\theta)$ is the prior density and

$p(Y|\theta)$ is the likelihood of data $Y$ computed by Kalman filter for a given parameter set $\theta$.

- We replace the posterior densities by $p(\theta)p(Y|\theta)$, which does not require the marginal data density $p(Y)$. Note that $p(\theta|Y) = \frac{p(\theta)p(Y|\theta)}{p(Y)}$.

# RWMH algorithm

- The mean and variance of $q\left(\vartheta\middle|\theta^{i-1}\right)$ are $\theta^{i-1}$ and $c^2\hat{\Sigma}$, then the drawn parameter can be written as
$$\vartheta = \theta^{i-1} + \eta$$

where $\eta$ follows a mean zero with variance $c^2\hat{\Sigma}$.

- Given the symmetric nature of $q\left(\vartheta\middle|\theta^{i-1}\right)$, the acceptance probability becomes
$$\alpha = \min\left\{1, \frac{p(Y|\vartheta)p(\vartheta)}{p\left(Y\middle|\theta^{i-1}\right)p(\theta^{i-1})}\right\}$$

# Tuning the algorithm

- We use a multivariate normal proposal distribution for $\eta$,

$$\eta \sim N\left(0, c^2\hat{\Sigma}\right)$$

or equivalently,

$$\vartheta \sim q\left(\vartheta \middle| \theta^{i-1}\right) = N\left(\theta^{i-1}, c^2\hat{\Sigma}\right)$$

- We have two parameters for tuning the algorithm, $c$ and $\hat{\Sigma}$.

- $\hat{\Sigma}$ controls the relative variance and correlations in the proposal distribution. For example:
  - Suppose parameters in $\theta$ are highly correlated.
  - If $\hat{\Sigma}$ does not capture this correlation, e.g., using a diagonal matrix for $\hat{\Sigma}$, then the proposal $\vartheta$ is unlikely to reflect the correlation.
  - Draws $\theta$ are more likely rejected, and to raise the acceptance rate, $c$ is tuned to small values.
  - The chain will be very highly correlated and will have high variance, which is not desirable.

- Obtaining a good $\hat{\Sigma}$ is difficult.
  - Usually, $\hat{\Sigma}$ is set to be the negative of the inverse Hessian of logged probability density function at the mode.
  - If the posterior distribution is multivariate normal, it is the covariance matrix.

$$\ln p(\theta) \propto -\frac{1}{2}(\theta-\bar{\theta})'\Sigma^{-1}(\theta-\bar{\theta}), \qquad \left[\left.\frac{\partial \ln p(\theta)}{\partial\theta\partial\theta'}\right|_{\theta=\bar{\theta}}\right]^{-1} = \Sigma$$

Thus, the inverse Hessian might be a good approximation to the covariance matrix.

  - Hessian can be obtained by running a numerical optimization routine before MCMC. However, The numerical approximation of the Hessian may be inaccurate.

- A partially adaptive approach may work well:
  - First, generate a set of posterior draws based on a reasonable initial choice for $\hat{\Sigma}$ (say, the covariance matrix for the prior distribution).
  - Compute the sample covariance matrix from the first draws and use it as $\hat{\Sigma}$ in a second run, and so on.

- $c$ is adjusted to ensure a “reasonable” acceptance rate.
    - If the sampler accepts too frequently, it may make very small movements. Then $c$ should be *larger*.
    - if the sampler rejects too frequently, it may be got stuck in one region of the parameter space. Then $c$ should be *smaller*.
- Most practitioners target an acceptance rate between 0.20 and 0.40.
    - Roberts et al. (1997) has derived a limit optimal rate of 0.234 for the special case of multivariate normal.

# Applications:
# A New Keynesian model

# A log-linearized New Keynesian model (Herbst and Schorfheide, 2015)

- Now we are ready to apply our techniques to the small-scale New Keynesian model.
- The equilibrium conditions are

$$\hat{c}_t = E_t \hat{c}_{t+1} - \tau^{-1}\left(\hat{R}_t - E_t \hat{\pi}_{t+1} - \rho_z \hat{z}_t\right)$$
$$\hat{\pi}_t = E_t \hat{\pi}_{t+1} + \kappa \hat{c}_t$$
$$\hat{R}_t = \rho_R \hat{R}_{t-1} + (1 - \rho_R)(\psi_1 \hat{\pi}_t + \psi_2 \hat{c}_t) + \epsilon_{R,t}$$
$$\hat{y}_t = \hat{c}_t + \hat{g}_t$$
$$\hat{g}_t = \rho_g \hat{g}_{t-1} + \epsilon_{g,t}$$
$$\hat{z}_t = \rho_z \hat{z}_{t-1} + \epsilon_{z,t}$$

# State equation

- The equilibrium conditions are summarized into

$$\mathcal{A}E_t\{x_{t+1}\} + \mathcal{B}x_t + \mathcal{C}x_{t-1} + \mathcal{E}\epsilon_t = 0$$

where

$$x_t = \left[\hat{c}_t, \hat{\pi}_t, \hat{R}_t, \hat{y}_t, \hat{g}_t, \hat{z}_t\ \right]'$$
$$\epsilon_t = \left[\epsilon_{z,t}, \epsilon_{g,t}, \epsilon_{R,t}\right]'$$

- The REE solution to the equilibrium condition

$$x_t = Ps_{t-1} + Q\epsilon_t, \qquad \epsilon_t \sim N(0, \mathrm{S}_e)$$

This is **the state equation**.

# Observation equation

- We have observed variables, which linked to model variables by

$$\Delta y_t^{obs} = \gamma^{(Q)} + (\hat{y}_t - \hat{y}_{t-1} + \hat{z}_t)$$
$$\pi_t^{obs} = \pi^{(A)} + 4\hat{\pi}_t$$
$$R_t^{obs} = \pi^{(A)} + r^{(A)} + 4\gamma^{(Q)} + 4\hat{R}_t$$

- These equations are summarized into

$$y_t = A + Bx_t + e_t, \qquad e_t \sim N(0, H)$$

This is **the observation equation**.

# State-space representation

- Then we have a state-space representation:

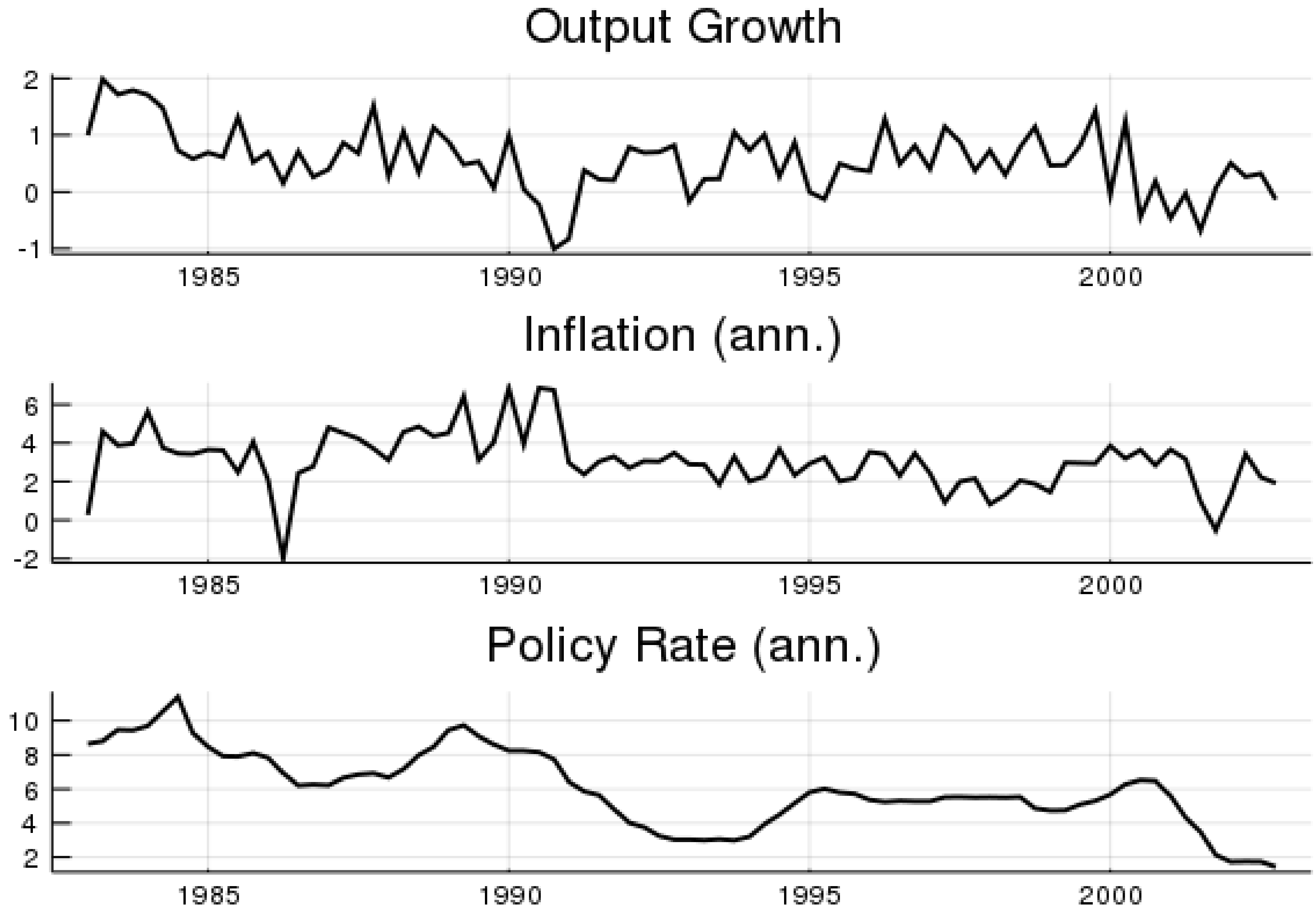
$$
\begin{aligned}
y_t &= A(\theta) + B(\theta)x_t + e_t, \quad && e_t \sim N(0, H(\theta)) \\
x_t &= P(\theta)x_{t-1} + Q(\theta)\epsilon_t, \quad && \epsilon_t \sim N(0, S_e(\theta))
\end{aligned}
$$

where $\theta = \left[\tau, \kappa, \psi_1, \psi_2, \rho_R, \rho_g, \rho_z, r^{(A)}, \pi^{(A)}, \gamma^{(Q)}, \sigma_R, \sigma_g, \sigma_z\right]'$.

- We can apply Kalman filter to obtain the likelihood function $p(Y|\theta)$ at $\theta$ given the data $Y$.

# Data

- We use three observables: $\{\Delta y_t^{obs}, \pi_t^{obs}, R_t^{obs}\}$
  - quarterly per capita GDP growth
  - quarterly inflation
  - the annualized federal funds rate
- The sample period is from 1983:I to 2002:IV (in the period of the great moderation), giving us a total of 80 observations.


Output Growth
2
1
0
-1
1985
1990
1995
2000
Inflation (ann.)
6
4
2
0
-2
1985
1990
1995
2000
Policy Rate (ann.)
10
8
6
4
2
1985
1990
1995
2000

# Bayesian inference

- Recall that we need the prior $p(\theta)$ and likelihood $p(Y|\theta)$ to obtain the posterior distribution of the parameters
$$p(\theta|Y) \propto p(\theta)p(Y|\theta)$$
- Priors are based on the information on parameters from micro data and other presumptions by the econometrician.
- Priors are also useful to avoid identification problems.

# Priors (somewhat diffused)

- Para(1) and Para(2) are:
  - the means and standard deviations for Beta, Gamma, and Normal distributions;
  - the upper and lower bound for Uniform distributions
  - $s$ and $\nu$ for Inverse Gamma distributions

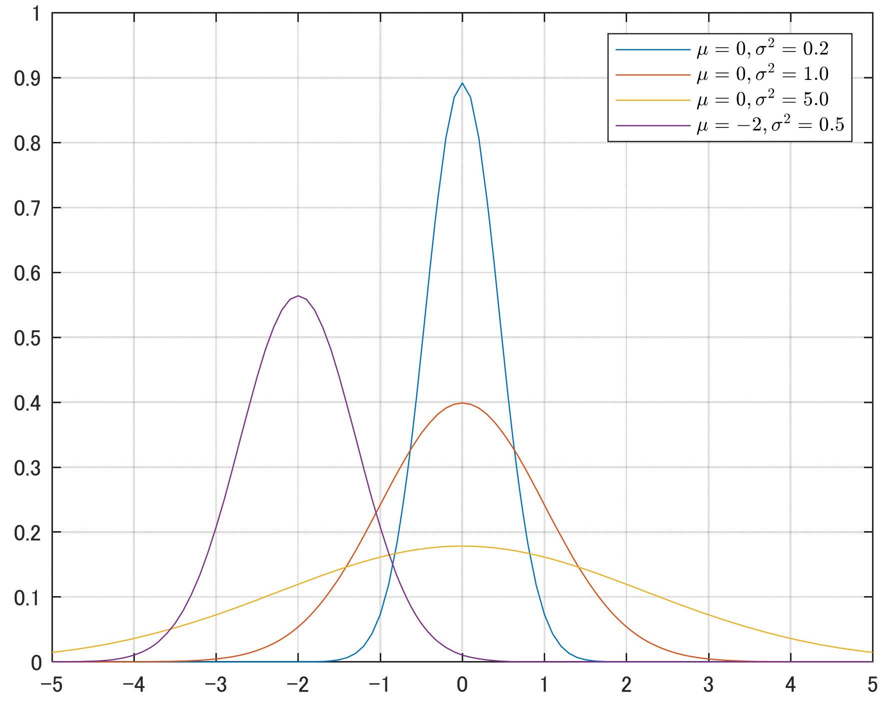
| Name | Domain | Prior | | |
|---|---|---|---|---|
| | | Density | Para (1) | Para (2) |
| $\tau$ | $\mathbb{R}^+$ | Gamma | 2.00 | 0.50 |
| $\kappa$ | $\mathbb{R}^+$ | Uniform | 0.00 | 1.00 |
| $\psi_1$ | $\mathbb{R}^+$ | Gamma | 1.50 | 0.25 |
| $\psi_2$ | $\mathbb{R}^+$ | Gamma | 0.50 | 0.25 |
| $\rho_R$ | $[0, 1)$ | Uniform | 0.00 | 1.00 |
| $\rho_G$ | $[0, 1)$ | Uniform | 0.00 | 1.00 |
| $\rho_Z$ | $[0, 1)$ | Uniform | 0.00 | 1.00 |
| $r^{(A)}$ | $\mathbb{R}^+$ | Gamma | 0.50 | 0.50 |
| $\pi^{(A)}$ | $\mathbb{R}^+$ | Gamma | 7.00 | 2.00 |
| $\gamma^{(Q)}$ | $\mathbb{R}$ | Normal | 0.40 | 0.20 |
| $100\sigma_R$ | $\mathbb{R}^+$ | InvGamma | 0.40 | 4.00 |
| $100\sigma_G$ | $\mathbb{R}^+$ | InvGamma | 1.00 | 4.00 |
| $100\sigma_Z$ | $\mathbb{R}^+$ | InvGamma | 0.50 | 4.00 |

# Normal distribution: $x \in (-\infty, \infty)$

- PDF is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

where $\mu$ is the mean and $\sigma$ is the standard deviation of the distribution.



# Gamma distribution: $x \in (0, \infty)$

- PDF is

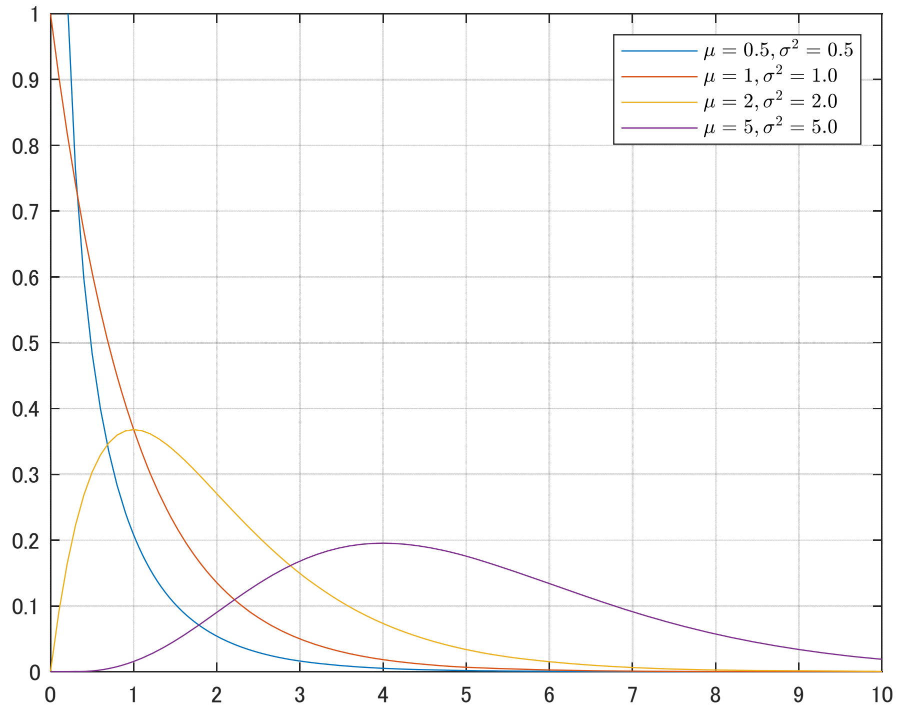
$$f(x) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}}$$

where $k > 0, \theta > 0$, and $\Gamma(k)$ is the Gamma function.

- Mean and variance are

$$E[x] = k\theta,$$
$$Var[x] = k\theta^2$$

- When $\theta = 1$, it is called the standard Gamma distr.



# Beta distribution: $x \in (0,1)$

- PDF is

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$$

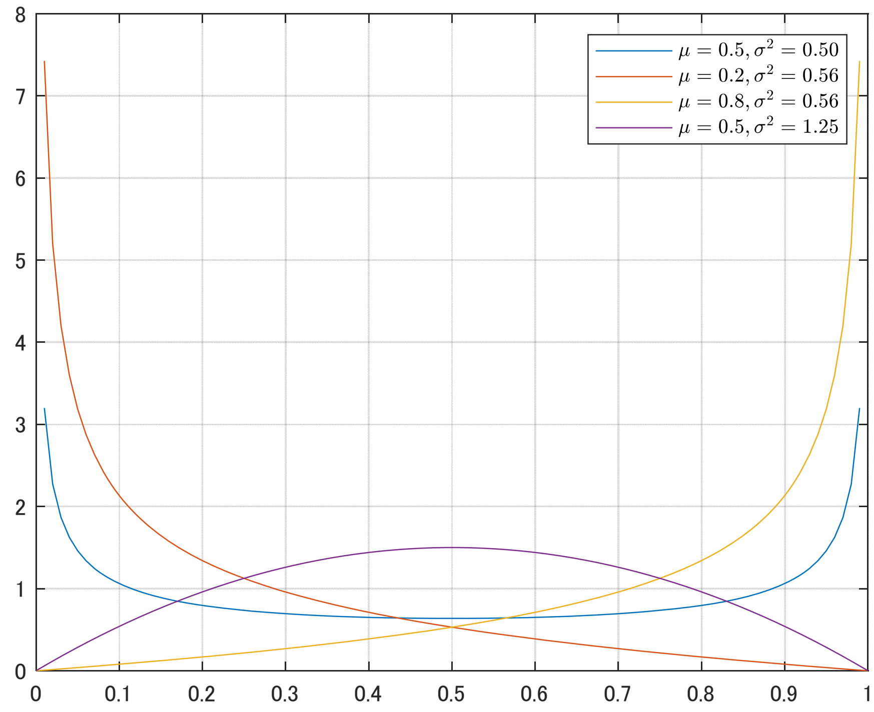
where $\alpha > 0, \beta > 0$, and

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\alpha)}{\Gamma(\alpha,\beta)}$$

- Mean and variance are

$$E[x] = \frac{\alpha}{\alpha+\beta},$$

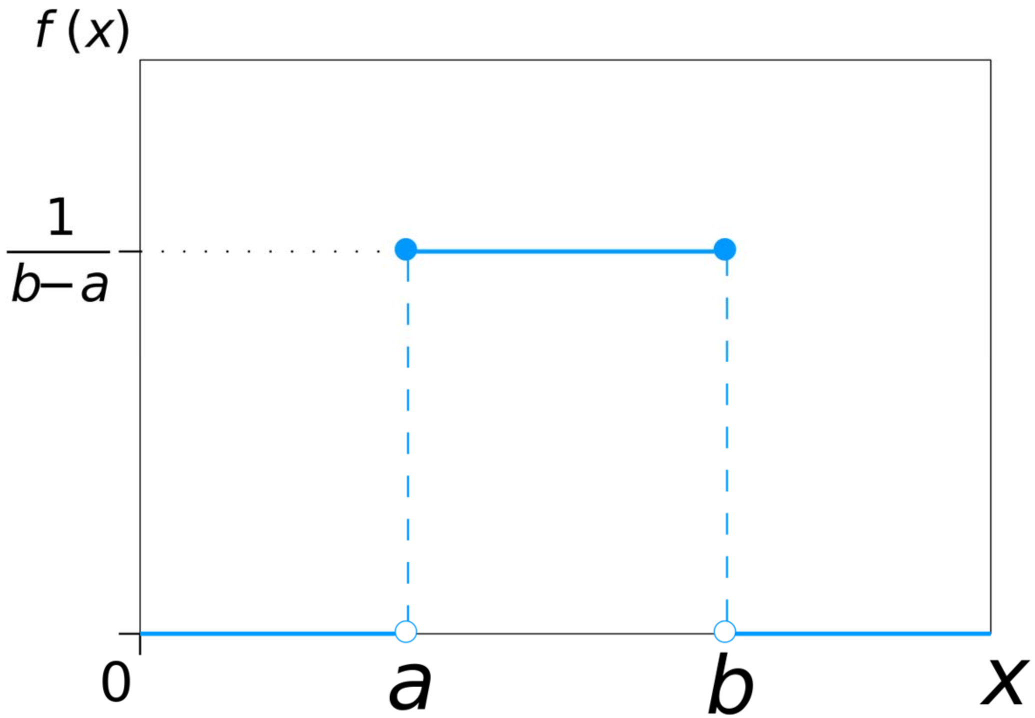
$$Var[x] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$



# Uniform distribution: $x \in [a, b]$

- PDF is

$f(x) = \frac{1}{b-a}$ for $x \in [a, b]$ and 0 otherwise.




https://en.wikipedia.org/wiki/Continuous_uniform_distribution

# Inverse Gamma: $x \in (0, \infty)$

- Compared to Gamma, the distribution is fat-tailed.
- PDF is (Dynare uses a bit non-standard one; see Ingvar Strid's note)

$$p(\sigma|a,b) = 2C_g^{-1}\left(\frac{b}{2},\frac{2}{ba^2}\right)\sigma^{-b-1}e^{-\frac{ba^2}{2\sigma^2}}$$

$$\propto \sigma^{-\nu-1}e^{-\frac{ba^2}{2\sigma^2}}$$

where $C_g\left(\frac{b}{2},\frac{2}{ba^2}\right) = \Gamma\left(\frac{b}{2}\right)\left(\frac{2}{ba^2}\right)^{\frac{b}{2}}$. (see p. 27 in the textbook with $b = \nu$ and $a = s$)

The log density is

$$\ln p(\sigma|a,b) = \ln(2) - \ln\Gamma\left(\frac{b}{2}\right) - \frac{b}{2}\ln\left(\frac{2}{ba^2}\right) - (b+1)\ln(\sigma) - \frac{ba^2}{2\sigma^2}$$

- Mean and variance of $x = \sigma$ is

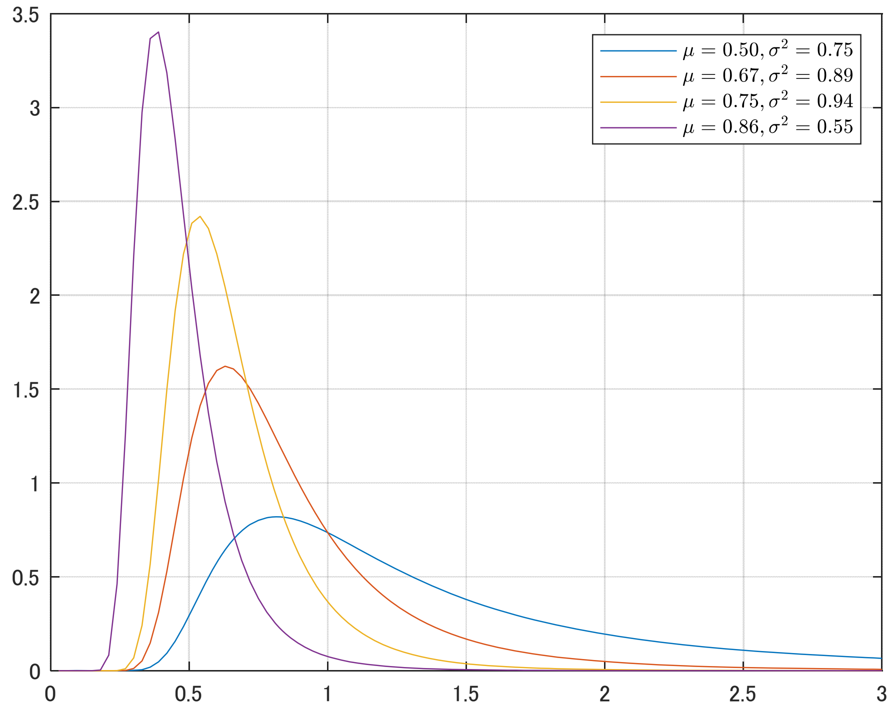
$$E[x] = \sqrt{\frac{ba^2}{2}} \frac{\Gamma\left(\frac{b-1}{2}\right)}{\Gamma\left(\frac{b}{2}\right)}, b > 1$$

and

$$Var[x] = \frac{ba^2}{b-2} - E(x)^2, b > 2$$

- **Note**: Dynare requires the mean and standard deviation for inputs.

For example, if $a = 0.4$ and $b = 4$, $E[x] = 0.501$ and $Var[x] = 0.262^2$.



# Dynare code: as_est.mod

```
var notr yy dp nomr gshk zshk rshk dy_obs pi_obs
r_obs;
varexo errgshk errzshk errrshk;
parameters tau kappa psi1 psi2 rho_R rho_g
rho_z rA piA gammaQ;

varobs dy_obs pi_obs r_obs;
```

- First, we define variables and set parameters.
  - varobs: the list of observed variables, which are also a part of the endogenous variables.
- Now, parameters are to be estimated.

```
model;
#bet  = 1/(1+rA/400);

yy + (1/tau)*nomr - (1-rho_g)*gshk - rho_z/tau*zshk - yy(+1)
- (1/tau)*dp(+1) = 0;

dp = kappa*(yy-gshk) + bet*dp(+1);

nomr = rho_R*nomr(-1) + (1-rho_R)*psi1*dp + (1-
rho_R)*psi2*(yy-gshk) + rshk;

gshk = rho_g*gshk(-1) + errgshk;
zshk = rho_z*zshk(-1) + errzshk;
rshk = errrshk;
```

- Then we input the log-linearized equations.

```
dy_obs = 1*(yy - yy(-1) + zshk) + gammaQ;
pi_obs = 4*dp + piA;
r_obs  = 4*nomr + (piA+rA+4*gammaQ);

end;
```

- We also have observation equations.

```
shocks;
var dy_obs;
stderr 0.1159;
var pi_obs;
stderr 0.2941;
var r_obs;
stderr 0.4475;
end;
```

- When estimating the model, a measurement error for each observed variable is between shocks; and end;
  - We multiply 0.2 by the standard deviation of each observed variable.

```
estimated_params;
tau,   gamma_pdf, 2.0, 0.5;
kappa, uniform_pdf, , , 0.0, 1.0;
psi1,  gamma_pdf, 1.50, 0.25;
psi2,  gamma_pdf, 0.50, 0.25;
rho_R, uniform_pdf, , , 0.0, 1.0;
rho_g, uniform_pdf, , , 0.0, 1.0;
rho_z, uniform_pdf, , , 0.0, 1.0;
rA,    gamma_pdf, 0.5, 0.5;
piA,   gamma_pdf, 7.0, 2.0;
gammaQ, normal_pdf, 0.4, 0.2;
stderr errrshk, inv_gamma_pdf, 0.5013, 0.2620;
stderr errgshk, inv_gamma_pdf, 1.2533, 0.6551;
stderr errzshk, inv_gamma_pdf, 0.6266, 0.3275;
end;
```

- Prior distributions are specified between estimated_params; and end;
- Dynare requires the mean and standard deviation for inputs of Inverse Gamma distribution.

$$E[x] = \sqrt{\frac{ba^2}{2}}\frac{\Gamma\left(\frac{b-1}{2}\right)}{\Gamma\left(\frac{b}{2}\right)}, b > 1$$

and

$$Var[x] = \frac{ba^2}{b-2} - E(x)^2, b > 2$$

```
estimated_params_init;
tau      , 2.09;
kappa    , 0.98;
psi1     , 2.25;
psi2     , 0.65;
rho_R    , 0.81;
rho_g    , 0.98;
rho_z    , 0.93;
rA       , 0.34;
piA      , 3.16;
gammaQ   , 0.51;
stderr errrshk , 0.19;
stderr errgshk , 0.65;
stderr errzshk , 0.24;
end;
```

- The mode of parameters are obtained by maximizing the prior and likelihood.
- Initial values for optimization are between estimated_params_init; and end;

```
estimation(
datafile = us,
mode_compute = 6, // set to 6 for diffused prior
//mode_compute = 0,
//mode_file = as_est_mode,
mode_check,
mh_replic = 20000,
mh_drop = 0.25,
mh_jscale = 0.6
);
```

- Estimation command has many options:

  datafile: the name of the data file

  mode_compute: the algorithm used to compute the mode. For example,

    4: (default) Chris sims' csminwel

    6: A Monte-Carlo based optimization (**very powerful**)

    7: fminsearch based on simplex method

  mode_file: the mode previously computed (with mode_compute = 0, it is used without optimization routine). The filename is usually [mod file name]_mode.mat

  mode_check: check the curvature of the objective function around the mode

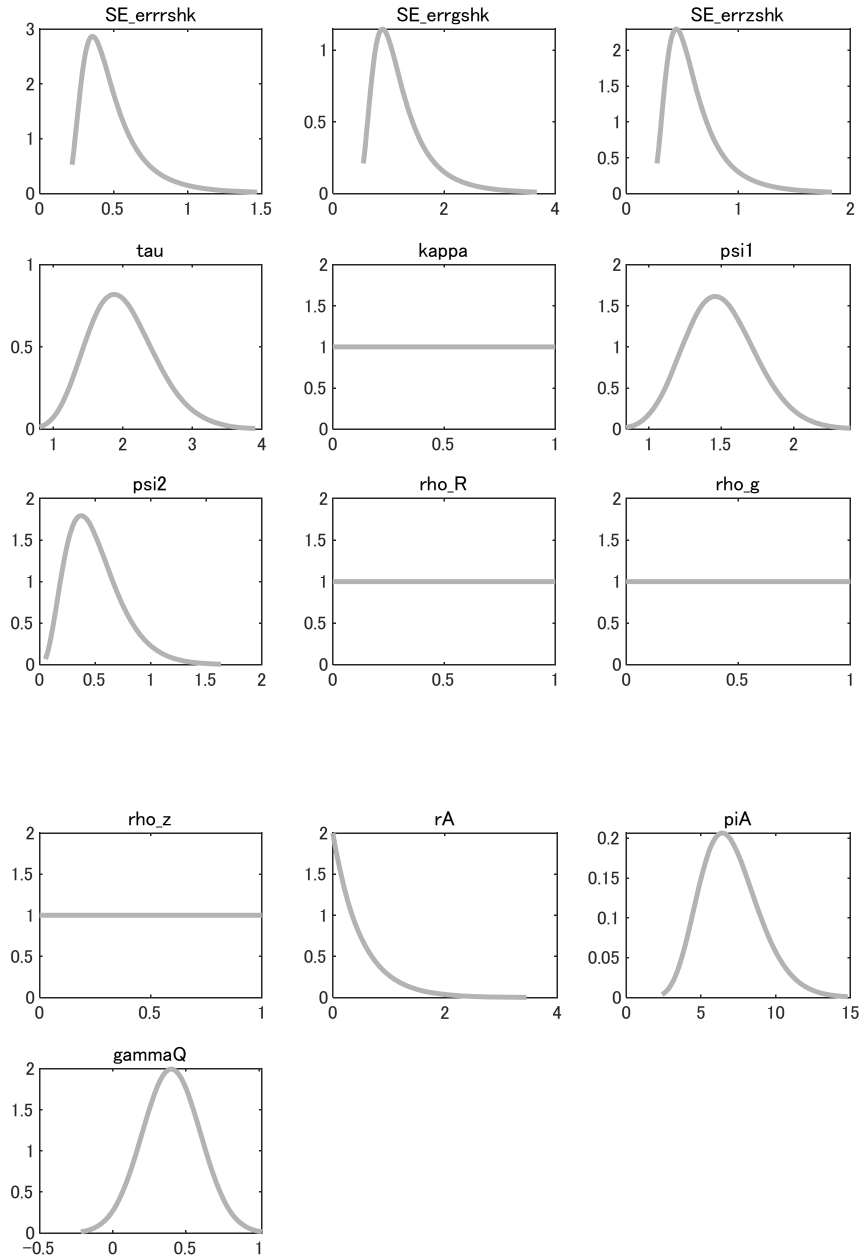
  mh_replic: the length of the MCMC chain (i.e., sampling period)

  mh_drop: the ratio of the burn-in period in the MCMC chain.

  mh_jscale: scaling parameter for the proposal distribution

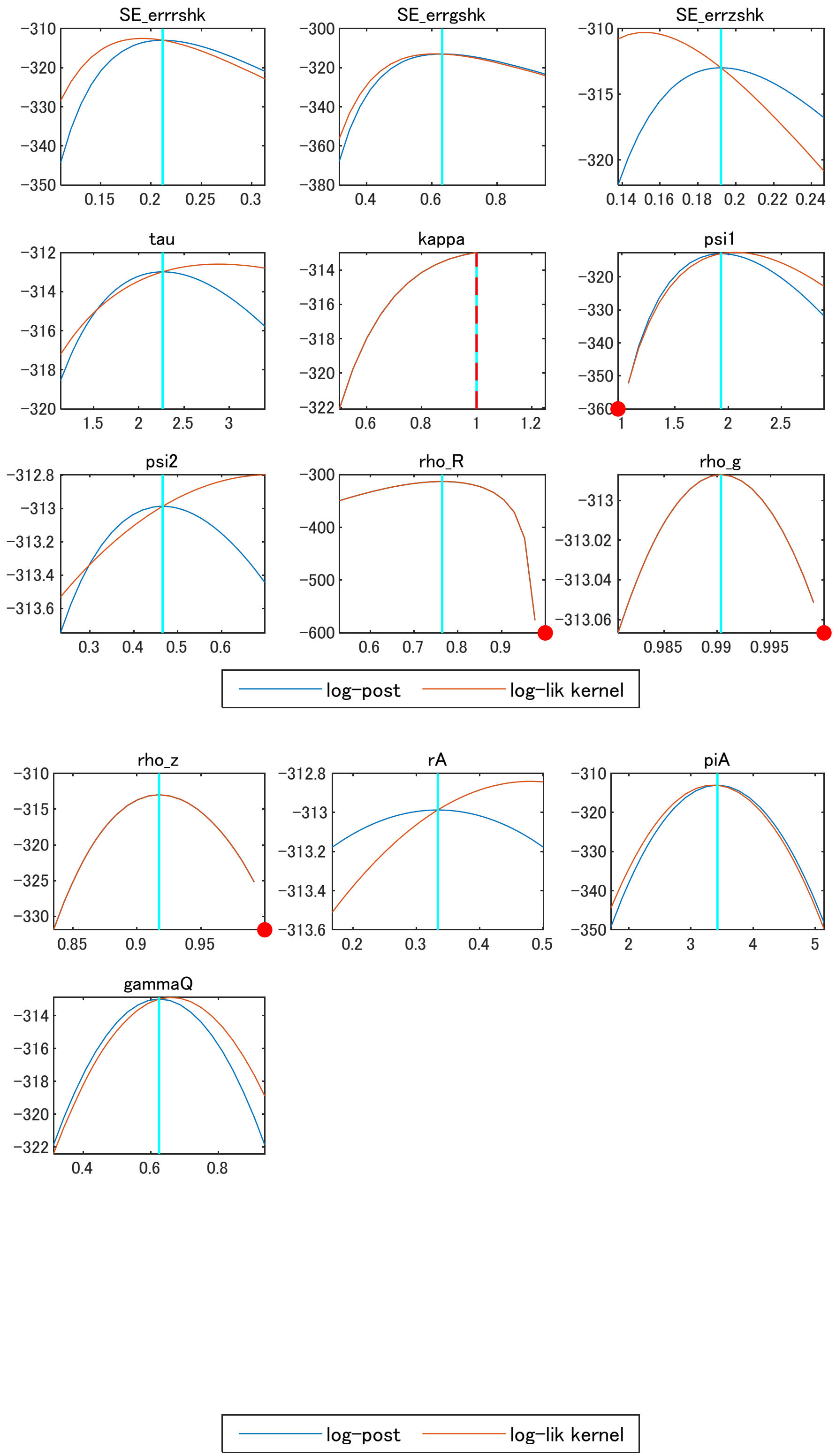
# Priors

```
estimated_params;
tau,   gamma_pdf, 2.0, 0.5;
kappa, uniform_pdf, , , 0.0, 1.0;
psi1,  gamma_pdf, 1.50, 0.25;
psi2,  gamma_pdf, 0.50, 0.25;
rho_R, uniform_pdf, , , 0.0, 1.0;
rho_g, uniform_pdf, , , 0.0, 1.0;
rho_z, uniform_pdf, , , 0.0, 1.0;
rA,    gamma_pdf, 0.5, 0.5;
piA,   gamma_pdf, 7.0, 2.0;
gammaQ, normal_pdf, 0.4, 0.2;
stderr errrshk, inv_gamma_pdf, 0.5013, 0.2620;
stderr errgshk, inv_gamma_pdf, 1.2533, 0.6551;
stderr errzshk, inv_gamma_pdf, 0.6266, 0.3275;
end;
```



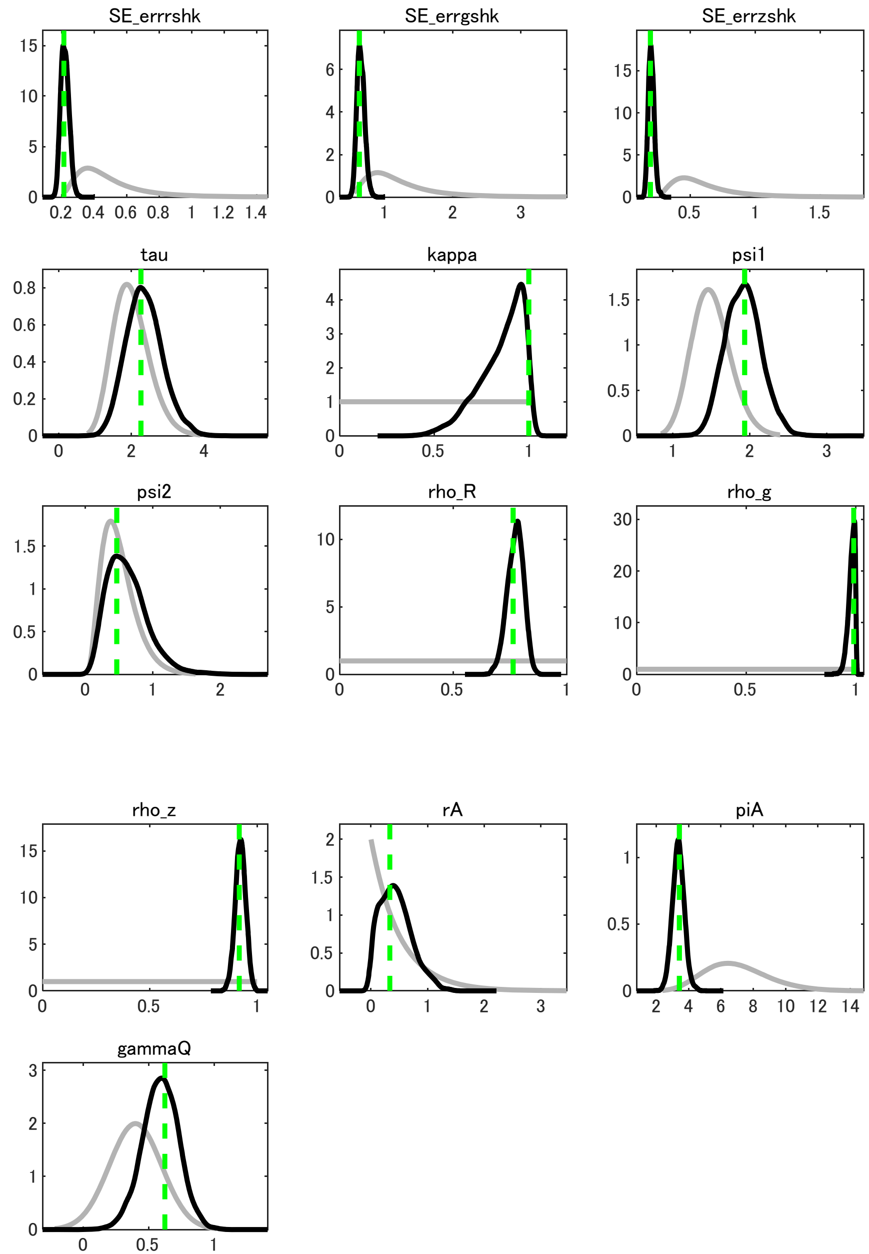
# Posterior mode

- With mode_check, check if the log posterior is maximized at the mode



# Posterior

- Check if the posterior are prior are different so that the data has some information on the parameters
$$p(\theta|Y) \propto p(\theta)p(Y|\theta)$$



# Posterior mode

RESULTS FROM POSTERIOR ESTIMATION

parameters

| | prior mean | mode | s.d. | prior | pstdev |
|---|---|---|---|---|---|
| tau | 2.0000 | 2.2640 | 0.5489 | gamm | 0.5000 |
| kappa | 0.5000 | 1.0000 | 0.1183 | unif | 0.2887 |
| psi1 | 1.5000 | 1.9323 | 0.2304 | gamm | 0.2500 |
| psi2 | 0.5000 | 0.4658 | 0.3025 | gamm | 0.2500 |
| rho_R | 0.5000 | 0.7645 | 0.0358 | unif | 0.2887 |
| rho_g | 0.5000 | 0.9903 | 0.0162 | unif | 0.2887 |
| rho_z | 0.5000 | 0.9173 | 0.0236 | unif | 0.2887 |
| rA | 0.5000 | 0.3335 | 0.2785 | gamm | 0.5000 |
| piA | 7.0000 | 3.4278 | 0.3917 | gamm | 2.0000 |
| gammaQ | 0.4000 | 0.6247 | 0.1404 | norm | 0.2000 |

standard deviation of shocks

| | prior mean | mode | s.d. | prior | pstdev |
|---|---|---|---|---|---|
| errrshk | 0.5013 | 0.2116 | 0.0274 | invg | 0.2621 |
| errgshk | 1.2533 | 0.6322 | 0.0543 | invg | 0.6551 |
| errzshk | 0.6267 | 0.1922 | 0.0227 | invg | 0.3276 |

- First, we compute the posterior mode, which maximizes the marginal density.

# Posterior

ESTIMATION RESULTS

Log data density is -334.721071.

parameters

| | prior mean | post. mean | 90% HPD interval | | prior | pstdev |
|---|---|---|---|---|---|---|
| tau | 2.000 | 2.3524 | 1.5341 | 3.1374 | gamm | 0.5000 |
| kappa | 0.500 | 0.8542 | 0.6801 | 0.9998 | unif | 0.2887 |
| psi1 | 1.500 | 1.9108 | 1.5280 | 2.2749 | gamm | 0.2500 |
| psi2 | 0.500 | 0.6180 | 0.1444 | 1.0347 | gamm | 0.2500 |
| rho_R | 0.500 | 0.7741 | 0.7183 | 0.8331 | unif | 0.2887 |
| rho_g | 0.500 | 0.9797 | 0.9583 | 1.0000 | unif | 0.2887 |
| rho_z | 0.500 | 0.9226 | 0.8833 | 0.9629 | unif | 0.2887 |
| rA | 0.500 | 0.4429 | 0.0028 | 0.8189 | gamm | 0.5000 |
| piA | 7.000 | 3.3612 | 2.7668 | 3.9488 | gamm | 2.0000 |
| gammaQ | 0.400 | 0.5889 | 0.3581 | 0.8063 | norm | 0.2000 |

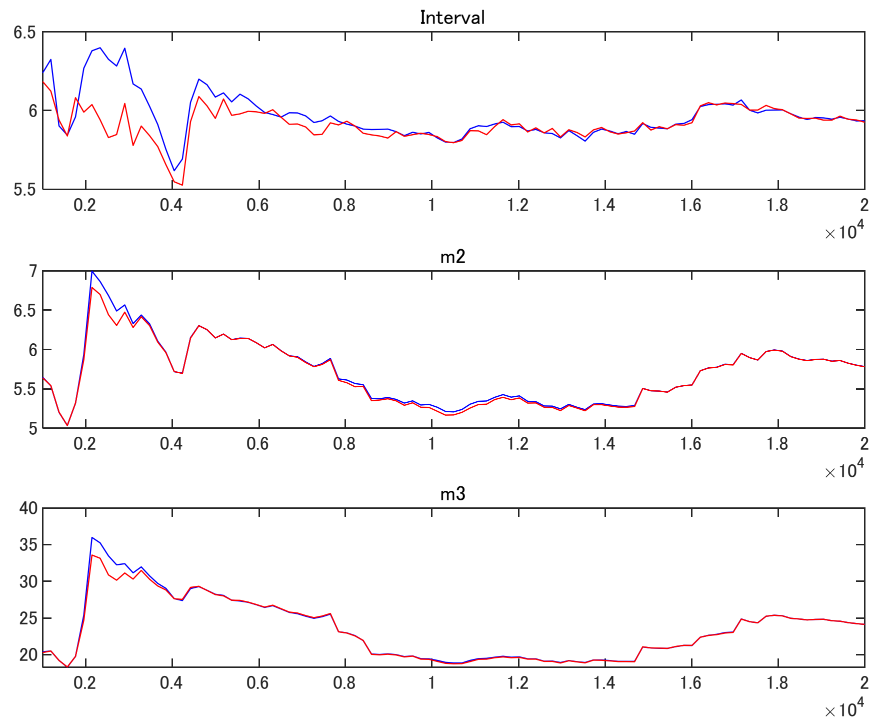
standard deviation of shocks

| | prior mean | post. mean | 90% HPD interval | | prior | pstdev |
|---|---|---|---|---|---|---|
| errrshk | 0.501 | 0.2180 | 0.1781 | 0.2587 | invg | 0.2621 |
| errgshk | 1.253 | 0.6533 | 0.5585 | 0.7370 | invg | 0.6551 |
| errzshk | 0.627 | 0.1990 | 0.1626 | 0.2344 | invg | 0.3276 |

- Then we do MCMC to infer the posterior distribution: Posterior mean and 90% credible intervals

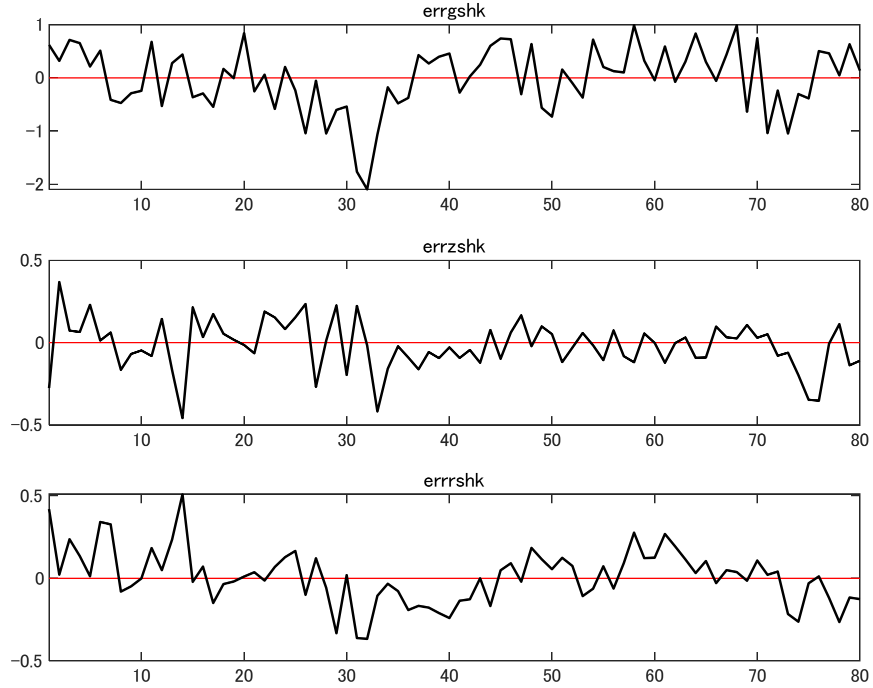
# Convergence

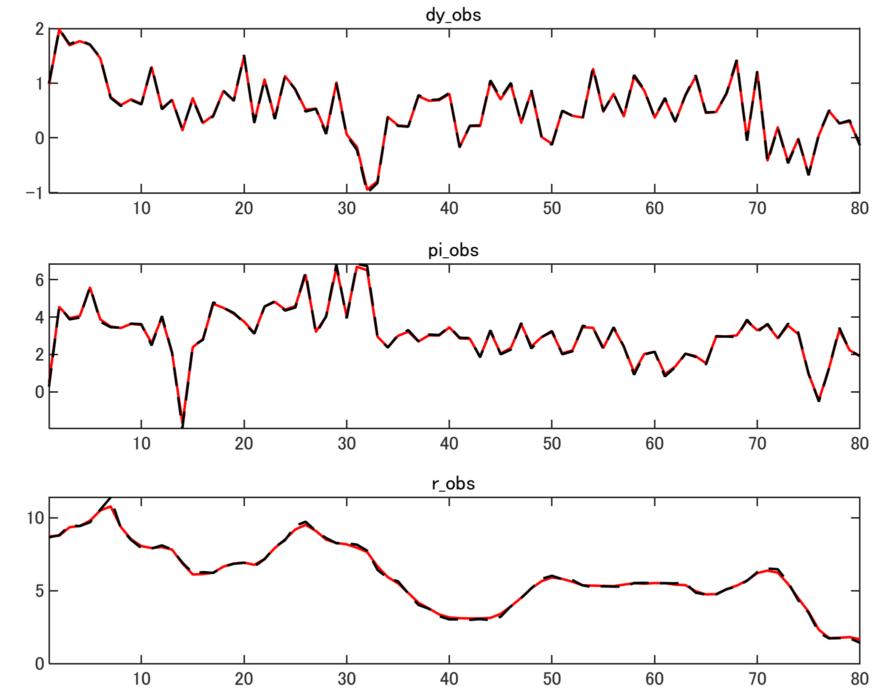
- The Brooks and Gelman (1998) convergence diagnostics (with two chains):
- There may be a lot of draws before the sample has “converged” that is when a draw from the chain is indistinguishable from a draw from the (invariant) posterior.
- The first $N_0 = 5000$ draws of the MH chain are dropped.



# Shocks

- “Smoothed” shocks and observed variables at the posterior mean



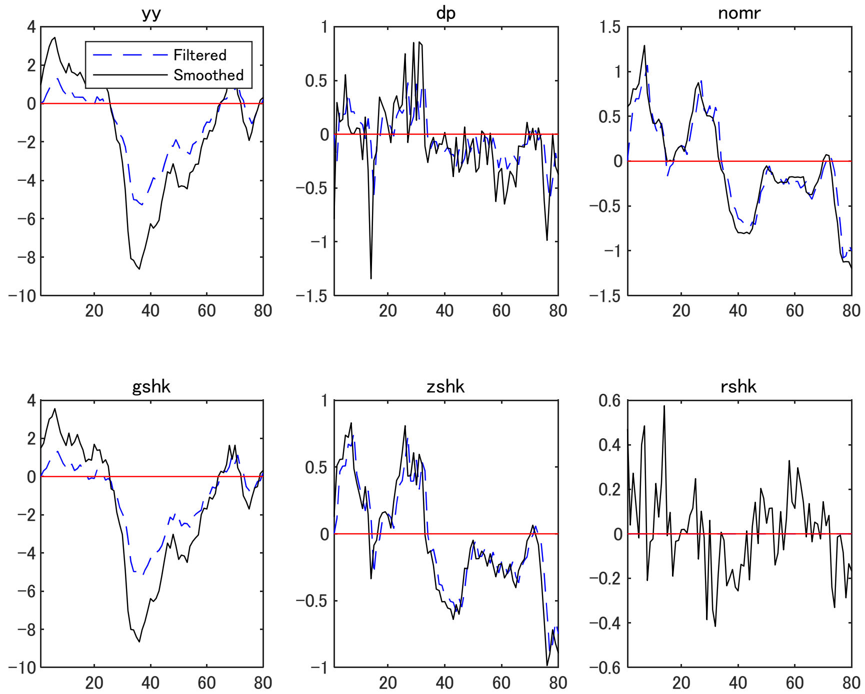
# Smoothed and filtered variables

- calib_smoother calculates smoothed and filtered (option) variables.
- Filtering

$$\hat{\alpha}_{t|t} = E[\alpha_t | \Omega_t]$$

- Smoothing

$$\hat{\alpha}_{t|N} = E[\alpha_t | \Omega_N]$$

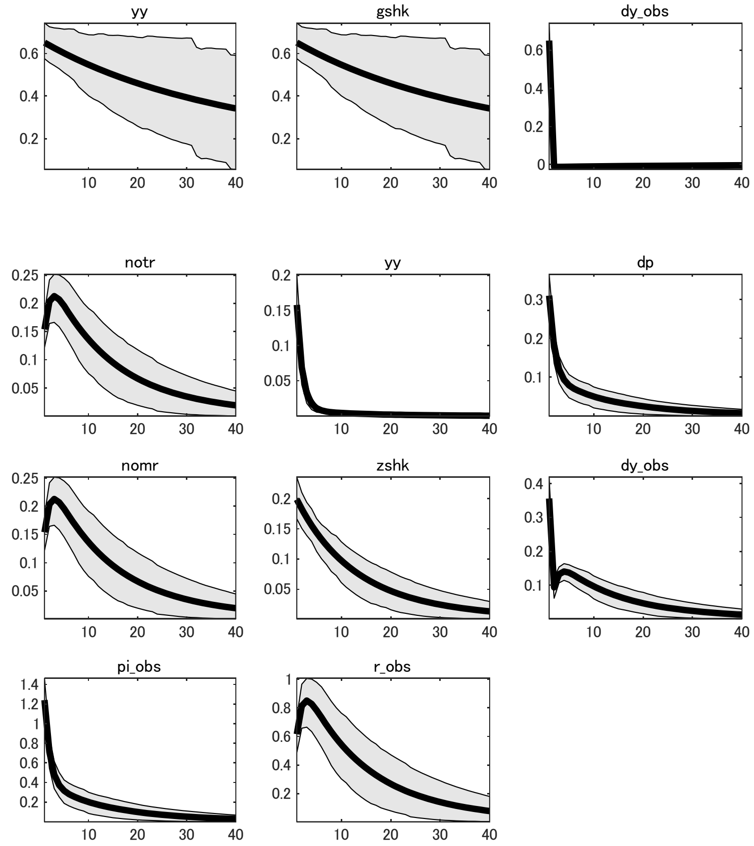


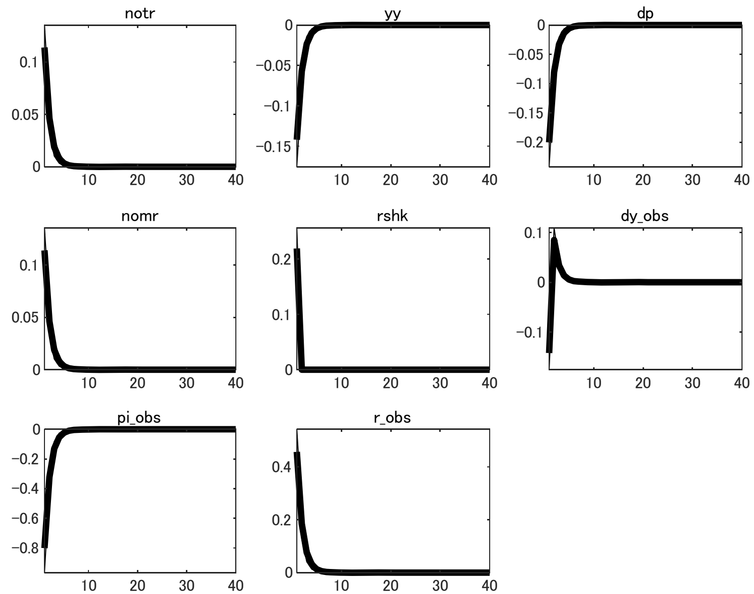

```
calib_smoother(datafile = us, filtered_vars);
```

# Bayesian IRFs

- The parameter draws can be transformed into other statistics of interest $h(\theta)$.
  - Bayesian inference for IRFs can be implemented first converting each $\theta$ to $h(\theta)$ and computing means and credible intervals.

```
estimation(
…
irf = 40,
bayesian_irf
);
```



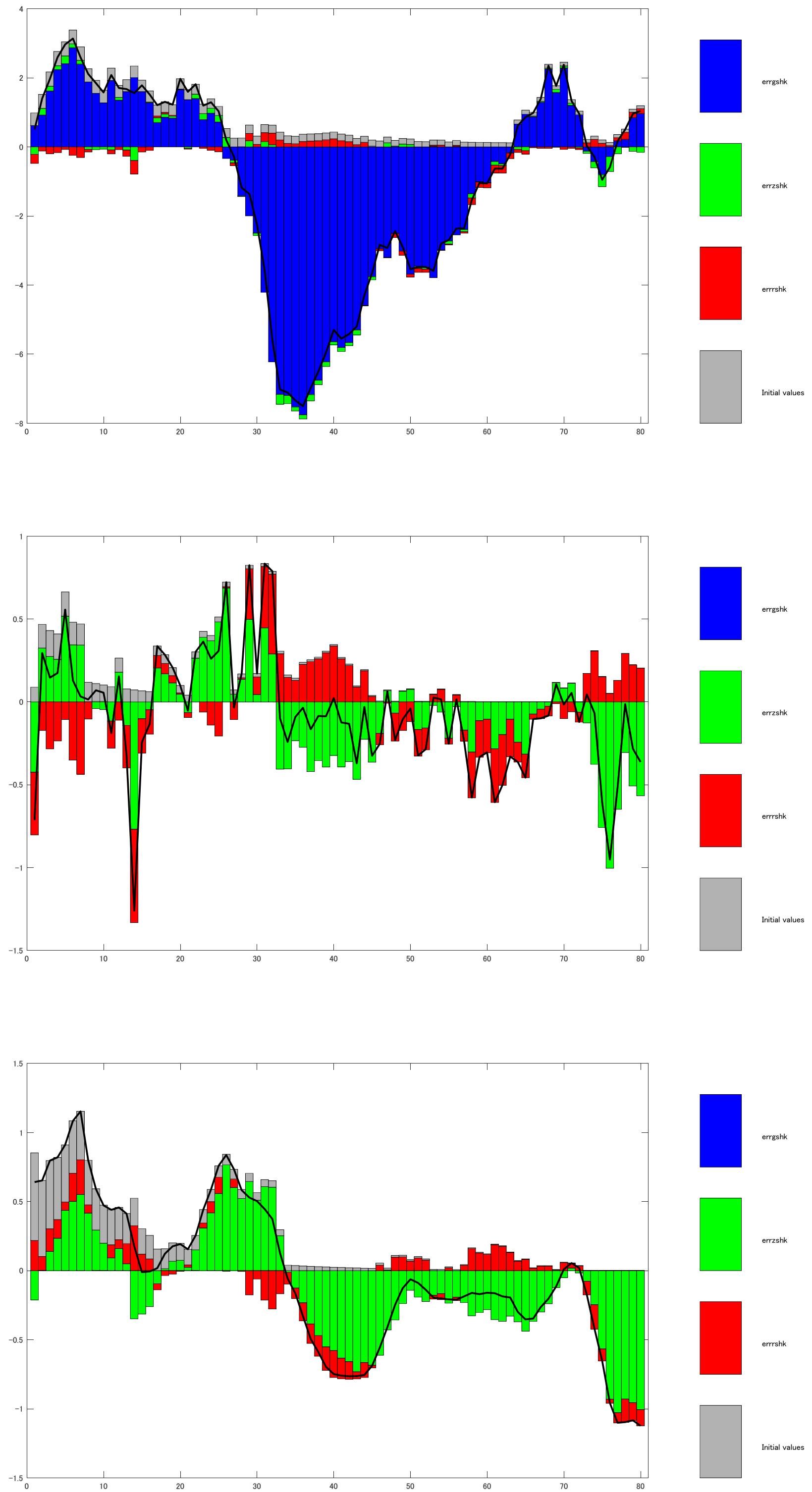
# Historical decomposition

- Each state variable at $t$ can be decomposed into the sequence of each shock

$$x_t = P(\theta)^t x_0 + Q(\theta)\sum_{i=1}^{t} P(\theta)^{t-i}\epsilon_i$$

- Note that the effect of initial values of $x_0$ diminishes as $t \to \infty$.

```
shock_decomposition yy dp nomr;
```



# A basic RBC model

# A Basic RBC Model

- The social planner's problem is given by

$$\max_{\{C_t, N_t, I_t\}} E_0 \sum_{t=0}^{\infty} \beta^t [\ln C_t + \chi \ln(1 - N_t)]$$

subject to

$$C_t + I_t = z_t K_t^{\alpha} N_t^{1-\alpha} = Y_t$$
$$K_{t+1} = (1 - \delta) K_t + I_t$$
$$\ln z_t = \rho \ln z_{t-1} + \sigma \varepsilon_t, \qquad \varepsilon_t \sim iid \mathcal{N}(0,1)$$

for $t = 0,1, \ldots$, taking $K_0$ and $z_{-1}$ as given.

- The optimal allocation $\{C_t, N_t, K_{t+1}, I_t, Y_t\}$ satisfies the following system of nonlinear difference equations:

$$\frac{1}{C_t} = E_t\left\{\frac{\beta}{C_{t+1}}\left[\alpha z_{t+1} K_{t+1}^{\alpha-1} N_{t+1}^{1-\alpha} + 1 - \delta\right]\right\}$$

$$\frac{\chi C_t}{1-N_t} = (1-\alpha) z_t K_t^{\alpha} N_t^{-\alpha}$$

and

$$C_t + I_t = Y_t$$
$$Y_t = z_t K_t^{\alpha} N_t^{1-\alpha}$$
$$K_{t+1} = (1-\delta) K_t + I_t$$

- The usual transversality condition also holds.
- In a decentralized equilibrium in which households own capital and make real investment, the rental rate and the wage rate are given by

$$R_{kt} = \alpha z_t K_t^{\alpha-1} N_t^{1-\alpha}$$
$$w_t = (1-\alpha) z_t K_t^{\alpha} N_t^{-\alpha}$$

# Steady State

- We solve for the non-stochastic steady state in which $z_t = 1$ for all $t$.
- By the Euler equation,

$$1 = \beta[\alpha K^{\alpha-1} N^{1-\alpha} + 1 - \delta]$$

which follows that

$$\frac{K}{N} = \left(\frac{\alpha}{1/\beta - (1-\delta)}\right)^{\frac{1}{1-\alpha}}$$

- Also,

$$I = \delta K = \frac{\delta K}{N} N \text{ and } Y = \left(\frac{K}{N}\right)^{\alpha} N$$

- Thus, given $N$, we can solve for $K$, $I$, and $Y$. Finally, $C$ comes from the resource constraint.
- We use the remaining equation for $\chi$:

$$\chi = (1-\alpha)\left(\frac{K}{N}\right)^{\alpha}\frac{1-N}{C}$$

Given other parameter values, there is a one-to-one mapping between $\chi$ and $N$. We calibrate a parameter using an endogenous variable as the input.

- By this approach, we can avoid solving a system of nonlinear equations for endogenous variables (Even though Dynare can solve it numerically)

# Calibration

- To parameterize and solve the model numerically, we use a calibration methodology.
- We assign parameter values based on either long-run steady-state relationship or moments from other studies such as microeconomic evidence.
    - $\alpha = 0.33$: consistent with the fact that labor share is about 2/3 in the U.S.
    - $\beta = 0.99$: the annual interest rate is $\frac{1}{\beta} - 1 = 4\%$.
    - $\delta = 0.025$: the annual depreciation rate is about $10\%$.
    - $N = 0.33$: households work 8 out of 24 hours a day.
    - Given $N$ and other parameter values, we can pin down $\chi = 1.75$ and all other steady-state values.

# An example with Dynare

- We solve and estimate the basic RBC model using Dynare.
- We set the parameter values

$$\alpha = 0.33, \beta = 0.99, \delta = 0.025, \chi = 1.75,$$
$$\rho = 0.95, \sigma = 0.01$$

which implies $N = 0.33$.

- We then solve the model using Dynare by log-linearization and obtain simulated data.

# Log-linearization by Dynare

- We use a change of variable $y_t = \ln x_t$, and Taylor expansion $f(x_t) = f(e^{y_t}) = 0$ around the steady state.

$$f(e^{y_t}) \approx f(e^{y}) + f'(e^{y})e^{y}(y_t - y)$$
$$\Leftrightarrow f(x_t) \approx f(x) + f'(x)x(\ln x_t - \ln x)$$

- Dynare computes the deviation from the steady state as the output of stoch_simul

$$\hat{x}_t = \ln x_t/x = \ln x_t - \ln x$$

Note: Dynare cannot deal with a variable if its steady state is zero. Thus, we use a change of variable

$$\tilde{z}_t = \exp z_t$$

$$\Leftrightarrow z_t = \ln \tilde{z}_t$$

If the steady state of $z_t$ is 0, that of $\tilde{z}_t$ is 1

Let $A_t = \exp a_t = e^{a_t} \Leftrightarrow a_t = \ln A_t$. Then

$$Y_t = A_t K_t^{\alpha}$$

$$\ln A_{t+1} = \rho_A \ln A_t + \varepsilon_{t+1}^{A}$$

Or,

$$Y_t = e^{a_t} K_t^{\alpha}$$

$$a_{t+1} = \rho_A a_t + \varepsilon_{t+1}^{A}$$

Dynare cannot deal with the latter! ($a = 0$ at the steady state)

# Dynare code: rbcdata.mod

```
var ly lc lk li lh lw Rk z;
varexo e;
parameters beta delta chi alpha rho;

[param,ss] = calibration;

alpha = param(1);
beta = param(2);
delta = param(3);
chi = param(4);
rho     = 0.95;
sigma   = 0.01;
```

- First we define variables and set parameters.
- As a separated file, Calibration.m sets the parameters using the steady state conditions.

```
model;
(1/exp(lc)) = beta*(1/exp(lc(+1)))*(1+alpha...
*(exp(lk)^(alpha-1))*exp(z(+1))*exp(lh(+1))^(1-alpha)-delta);
chi*exp(lc)/(1-exp(lh)) = exp(lw);
exp(lw) = exp(z)*(1-alpha)*exp(lk(-1))^alpha*exp(lh)^(-alpha);
exp(lc)+exp(li) = exp(ly);
exp(ly) = exp(z)*(exp(lk(-1))^alpha)*(exp(lh))^(1-alpha);
exp(li) = exp(lk)-(1-delta)*exp(lk(-1));
exp(Rk) = alpha*exp(ly)/exp(lk(-1));
z = rho*z(-1)+e;
end;
```

- Then we input model equations.
- We set $\tilde{X}_t = \ln X_t$ so that $X_t = e^{\tilde{X}_t}$

```
initval;
  lk = log(ss(1));
  lc = log(ss(2));
  lh = log(ss(3));
  li = log(ss(4));
  ly = log(ss(5));
  lw=log(ss(6));
  Rk=log(ss(7));
  ly_l = ly-lh;
  z = 0;
end;

steady;
```

- We solve for the steady state values numerically.
- Initial values are given so that we find the solution more likely.

```
check;

shocks;
var e = sigma^2;
end;

stoch_simul(periods=1000, order = 1);

datatomfile('simuldataRBC',[]);
```

- Check the Blanchard-Khan condition holds and the REE solution is unique.
- Finally, we simulate the model for 1,000 periods and save the simulated data.

# Appendix

# Notes on Inverse Gamma (Ingvar Strid)

- There are two representations of the inverse Gamma distribution.
- The inverse Gamma-2 distribution is based on the variance

$$\sigma^2 \sim IG_2(s, \nu)$$

with density function

$$p(\sigma^2|\nu, s) = C_g^{-1}\left(\frac{\nu}{2}, \frac{2}{s}\right)(\sigma^2)^{-\frac{1}{2}(\nu+2)} e^{-\frac{s}{2\sigma^2}}$$

where $C_g\left(\frac{\nu}{2}, \frac{2}{s}\right) = \Gamma\left(\frac{\nu}{2}\right)\left(\frac{2}{s}\right)^{\frac{\nu}{2}}$.

# Notes on Inverse Gamma (Ingvar Strid)

- The inverse Gamma-1 distribution is based on the standard deviation

$$\sigma \sim IG_2(s, \nu)$$

The density function can be obtained by change-of variables. If $Y = g(X) = X^{1/2}$, the density of $Y$ is given by

$$f(y) = \left| \frac{d}{dy} \left( g^{-1}(y) \right) \right| f_X\left( g^{-1}(y) \right)$$
$$= 2y \times f_X\left( g^{-1}(y) \right)$$

Then we have

$$p(\sigma | \nu, s) = 2\sigma \times C_g^{-1}\left( \frac{\nu}{2}, \frac{2}{s} \right) (\sigma^2)^{-\frac{1}{2}(\nu+2)} e^{-\frac{s}{2\sigma^2}}$$
$$= 2C_g^{-1}\left( \frac{\nu}{2}, \frac{2}{s} \right) \sigma^{-\nu-1} e^{-\frac{s}{2\sigma^2}}$$

# Notes on Inverse Gamma (Ingvar Strid)

The log density is

$$\ln p(\sigma|s,\nu) = \ln(2) - \ln\Gamma\left(\frac{\nu}{2}\right) - \frac{\nu}{2}\ln\left(\frac{2}{s}\right) - (\nu+1)\ln(\sigma) - \frac{s}{2\sigma^2}$$

- By setting $s = ba^2$ and $\nu = b$, we have the one in HS.

# Notes on Inverse Gamma (Ingvar Strid)

- In Wikipedia, the inverse Gamma-2 distribution of $x = \sigma^2$ is alternatively parameterized as

$$p(x|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-(\alpha+1)} e^{-\frac{\beta}{x}}$$

- By setting $\alpha = \frac{\nu}{2}$ and $\beta = s/2$, we get

$$p(x|s,\nu) = C_g\left(\frac{\nu}{2},\frac{2}{s}\right) x^{-\frac{1}{2}(\nu+2)} e^{-\frac{s}{2x}}$$

where $C_g\left(\frac{\nu}{2},\frac{2}{s}\right) = \Gamma\left(\frac{\nu}{2}\right)\left(\frac{2}{s}\right)^{\frac{\nu}{2}}$.