Tutorial on Economics and Econometrics of the New Keynesian Model

Lawrence Christiano

1 Simple New Keynesian Model

Following are the equations of the log-linearized simple New Keynesian model with no capital or working capital channel:

$$\pi_t = \beta E_t \pi_{t+1} + \kappa x_t \text{ (Phillips curve)}$$

$$x_t = -\left[r_t - E_t \pi_{t+1} - r_t^*\right] + E_t x_{t+1} \text{ (IS curve)}$$

$$r_t = \alpha r_{t-1} + (1-\alpha)\left[\phi_\pi \pi_t + \phi_x x_t\right] \text{ (policy rule)}$$

$$r_t^* = E_t\left(a_{t+1} - a_t\right) - \frac{1}{1+\varphi} E_t\left(\tau_{t+1} - \tau_t\right) \text{ (natural rate)}$$

$$y_t^* = a_t - \frac{1}{1+\varphi}\tau_t \text{ (natural output)}$$

$$x_t = y_t - y_t^* \text{ (output gap)}$$

$$\Delta a_t = \rho \Delta a_{t-1} + \varepsilon_t^a,\ \tau_t = \lambda \tau_{t-1} + \varepsilon_t^\tau$$

The above equations represent the equilibrium conditions of an economy, linearized about its steady state. In the economy, household preferences are given by:

$$E_0 \sum_{t=0}^{\infty} \left(\log C_t - \exp\left(\tau_t\right) \frac{N_t^{1+\varphi}}{1+\varphi} \right),\ \tau_t = \lambda \tau_{t-1} + \varepsilon_t^\tau,\ \varepsilon_t^\tau \tilde{}\, iid,$$

where C_t denotes consumption, τ_t is a time t preference shock and N_t denotes employment. The budget constraint of the household is:

$$P_t C_t + B_{t+1} \leq W_t N_t + R_{t-1} B_t + T_t,$$

where T_t denotes (lump sum) taxes and profits, P_t is the price level, W_t denotes the nominal wage rate and B_{t+1} denotes bonds purchased at time t which deliver a non-state-contingent rate of return, R_t, in period $t+1$.

Competitive firms produce a homogeneous output good, Y_t, using the following technology:

$$Y_t = \left[\int_0^1 Y_{i,t}^{\frac{\varepsilon-1}{\varepsilon}} di \right]^{\frac{\varepsilon}{\varepsilon-1}}, \ \varepsilon > 1,$$

where $Y_{i,t}$ denotes the i^{th} intermediate good, $i \in (0,1)$. The competitive firms takes the price of the final output good, P_t, and the prices of the intermediate goods, $P_{i,t}$, as given and chooses Y_t and Y_{it} to maximize profits. This results in the following first order condition:

$$Y_{i,t} = Y_t \left(\frac{P_t}{P_{i,t}} \right)^{\varepsilon}.$$

The producer of Y_{it} is a monopolist which takes the above equation as its demand curve. Note that if this demand curve is substituted back into the production function,

$$Y_t = \left[\int_0^1 Y_{i,t}^{\frac{\varepsilon-1}{\varepsilon}} di \right]^{\frac{\varepsilon}{\varepsilon-1}} = Y_t P_t^{\varepsilon} \left[\int_0^1 \left(P_{i,t}^{-\varepsilon} \right)^{\frac{\varepsilon-1}{\varepsilon}} di \right]^{\frac{\varepsilon}{\varepsilon-1}} = Y_t P_t^{\varepsilon} \left[\int_0^1 P_{i,t}^{(1-\varepsilon)} di \right]^{\frac{\varepsilon}{\varepsilon-1}},$$

or, after cancelling Y_t and rearranging,

$$\begin{aligned} P_t^{-\varepsilon} &= \left[\int_0^1 P_{i,t}^{(1-\varepsilon)} di \right]^{\frac{\varepsilon}{\varepsilon-1}} \\ P_t &= \left[\int_0^1 P_{i,t}^{(1-\varepsilon)} di \right]^{\frac{1}{1-\varepsilon}}. \end{aligned}$$

Thus, we get a simple expression relating the price of the aggregate good back to the individual prices.

The i^{th} intermediate good firm uses labor, $N_{i,t}$, to produce output using the following production function:

$$Y_{i,t} = \exp(a_t) N_{i,t}, \ \Delta a_t = \rho \Delta a_{t-1} + \varepsilon_t^a,$$

where Δ is the first difference operator and ε_t^a is an iid shock. We refer to the time series representation of a_t as a 'unit root' representation. The i^{th} firm sets prices subject to Calvo frictions. In particular,

$$P_{i,t} = \begin{cases} \tilde{P}_t & \text{with probability } 1-\theta \\ P_{i,t-1} & \text{with probability } \theta \end{cases},$$

where $\tilde{P}_t$ denotes the price chosen by the $1-\theta$ firms that can reoptimize their price at time t. The i^{th} producer is competitive in labor markets, where it pays $W_t(1-\nu)$ for one unit of labor. Here, ν represents a subsidy which has the effect of eliminating the monopoly distortion on labor in the steady state. That is, $1-\nu = (\varepsilon-1)/\varepsilon$.

At this point it is interesting to observe that if the household and government satisfy their budget constraints and markets clear, then the resource constraint is satisfied (Walras' law). Optimization leads the households to satisfy their budget constraint as a strict equality:

$$\begin{aligned} P_t C_t + B_{t+1} &= W_t N_t + R_{t-1} B_t + T_t \\ &= W_t N_t + R_{t-1} B_t + \overbrace{\int_0^1 P_{i,t} Y_{i,t} - (1-\nu) W_t \int_0^1 N_{i,t} di}^{\text{profits}} - T_t^g, \end{aligned}$$

where T_t^g denotes lump sum taxes raised by the government (profits from the final good firms need not be considered, because they are zero). The government budget constraint is

$$\nu W_t N_t + B_{t+1}^g = T_t^g + R_{t-1} B_t^g,$$

where B_{t+1}^g denotes government purchases of bonds (i.e., 'lending', if positive and 'borrowing' if negative). Note that, clearing in the labor market implies

$$\int_0^1 N_{i,t} di = N_t.$$

By the fact that final good firms make zero profits,

$$\int_0^1 P_{i,t} Y_{i,t} = P_t Y_t.$$

Substituting the government budget constraint and the expressions for profits (using labor market clearing) back into the budget constraint:

$$\begin{aligned}
P_t C_t + B_{t+1} &= W_t N_t + R_{t-1} B_t + T_t \\
&= W_t N_t + R_{t-1} B_t + \overbrace{P_t Y_t - (1-\nu) W_t N_t - \overbrace{\left[-R_{t-1} B_t^g + \nu W_t N_t + B_{t+1}^g\right]}^{=T_t^g}}^{T_t = \text{profits, net of taxes}} \\
&= W_t N_t + R_{t-1} B_t + P_t Y_t - (1-\nu) W_t N_t + R_{t-1} B_t^g - \nu W_t N_t - B_{t+1}^g \\
&= R_{t-1} B_t + P_t Y_t + R_{t-1} B_t^g - B_{t+1}^g
\end{aligned}$$

or,

$$P_t C_t + \left(B_{t+1} + B_{t+1}^g\right) = R_{t-1}\left(B_t + B_t^g\right) + P_t Y_t.$$

But, clearing in the bond market requires

$$B_{t+1} + B_{t+1}^g = 0 \text{ for all } t.$$

So,

$$C_t = Y_t,$$

and the resource constraint is satisfied. Incidentally, in this model with lump sum taxes, the equilibrium allocations are independent of the time pattern of government debt. So, for convenience, we just set $B_t^g = 0$ and so market clearing requires $B_t = 0$. Of course, we could have B_t not equal to zero, so that there is positive volume in the debt market. However, this would not be an interesting theory of why there is debt and so we don't do this.

The Ramsey equilibrium for the model is the equilibrium associated with the optimal monetary policy. It can be shown that the Ramsey equilibrium is characterized by zero inflation, $\pi_t = 0$, at each date and for each realization of a_t and τ_t and that consumption and employment in the Ramsey equilibrium corresponds to their first best levels.[1] That is, C_t and N_t satisfy the resource constraint

$$C_t = \exp(a_t) N_t,$$

[1]For a discussion, see http://faculty.wcas.northwestern.edu/~lchrist/course/optimalpolicyhandout.pdf

and the condition that the marginal rate of substitution between consumption and labor equals the marginal product of labor

$$\frac{\text{marginal utility of leisure}}{\text{marginal utility of consumption}} = C_t \exp(\tau_t) N_t^{\varphi} = \exp(a_t).$$

Solving for N_t :

$$\log(N_t^*) = -\frac{\tau_t}{1+\varphi}, \ \log(C_t^*) = a_t - \frac{\tau_t}{1+\varphi},$$

where $*$ indicates that the variable corresponds to the Ramsey equilibrium. In the description of the model above, y_t denotes log output and y_t^* denotes log output in the Ramsey equilibrium, i.e., $\log(C_t^*)$. The gross interest rate in the Ramsey equilibrium, R_t^*, satisfies the intertemporal household first order condition,

$$1 = \beta E_t \frac{u_{c,t+1}^*}{u_{c,t}^*} \frac{R_t^*}{1+\pi_{t+1}^*},$$

where $u_{c,t}^*$ indicates the marginal utility of consumption in the Ramsey equilibrium. Also, $\pi_t^* = 0$. With our utility function:

$$1 = \beta E_t \frac{C_t^*}{C_{t+1}^*} R_t^* = \beta E_t \frac{R_t^*}{\exp\left[\Delta a_{t+1} - \frac{\tau_{t+1}-\tau_t}{1+\varphi}\right]} = \beta E_t \exp\left[\log(R_t^*) - \Delta a_{t+1} + \frac{\tau_{t+1}-\tau_t}{1+\varphi}\right],$$

Approximately, one can 'push' the expectation operator into the power of the exponential. Doing so and taking the log of both sides, one obtains:

$$0 = \log\beta + \log(R_t^*) - E_t \Delta a_{t+1} + E_t \frac{\tau_{t+1}-\tau_t}{1+\varphi},$$

or,

$$r_t^* = E_t \Delta a_{t+1} - E_t \frac{\tau_{t+1}-\tau_t}{1+\varphi},$$

where $r_t^* \equiv \log(R_t^* \beta)$, the log deviation of R_t^* from its value in the non-stochastic steady state. The variable, r_t^*, corresponds to the 'natural rate of interest' and y_t^* corresponds to the 'natural rate of output'.

2 Exercises

1. Before turning to the econometric part of the assignment, it is useful to study the economics of the simple NK model, by seeing how the model economy responds to a shock. Consider the following parameterization:

$$\begin{aligned} \beta &= 0.97, \ \phi_x = 0, \ \phi_\pi = 1.5, \ \alpha = 0, \ \rho = 0.2, \ \lambda = 0.5, \\ \varphi &= 1, \ \theta = 0.75, \ \sigma_a = \sigma_\tau = 0.02. \end{aligned}$$

   (a) In the case of the technology and preference shocks, use Dynare to compute the impulse response functions of the variables to each shock. The m file, plots.m, can be used for this purpose.

i. Consider the response of the economy to a technology shock and a preference shock. In each case, indicate whether the economy over- or under- responds to the shock, relative to their 'natural' responses. What is the economic intuition in each case?

ii. Replace the time series representation of a_t with

$$a_t = \rho a_{t-1} + \varepsilon_t^a.$$

How does the response of the economy to ε_t^a with this representation compare to the response to ε_t^a with the unit root representation?

(b) Do the calculations with $\phi_\pi = 0.99$. What sort of message does Dynare generate, and can you provide the economic intuition for it? (In this case, there is 'indeterminacy', which means a type of multiplicity of equilibria...this happens whenever $\phi_\pi < 1$.) Provide intuition for this result.

(c) Return to the parameterization, $\phi_\pi = 1.5$. Now, insert r_t into the Cavlo pricing equation. Redo the calculations and note how Dynare reports indeterminacy again. Provide economic intuition for your result.

(d) Explain why it is that when the monetary policy rule is replaced by the $r_t = r_t^*$, the natural equilibrium (i.e., Ramsey) is a solution to the equilibrium conditions. Explain why the natural equilibrium is not the only solution to the equilibrium conditions (i.e., the indicated policy rule does not support the natural equilibrium uniquely). Verify this result computationally in Dynare.

(e) Now replace the monetary policy rule with

$$r_t = r_t^* + \alpha\left(r_{t-1} - r_{t-1}^*\right) + (1-\alpha)\left[\phi_\pi \pi_t + \phi_x x_t\right].$$

Explain why the natural equilibrium is a solution to the equilibrium conditions with this policy. Verify computationally that this policy rule uniquely supports the natural equilibrium (in the sense of satisfying determinacy), as long as ϕ_π is large enough. Provide intuition. Conclude that the Taylor rule uniquely supports the natural equilibrium if the natural rate of interest is included in the rule.

(f) Consider the following alternative representation for the technology shock:

$$a_t = \rho a_{t-1} + \xi_t^0 + \xi_{t-1}^1,$$

where both shocks are iid, so that the sum is iid too. Here, we assume agents see ξ_t^0 at time t and they see ξ_{t-1}^1 at $t-1$. Thus, agents have advance information (or, 'news') about the future realization of a shock. Introduce this change into the code and set $\rho = 0.2$. Verify that when there is a shock to ξ_t^1, inflation falls contemporaneously and the output gap jumps. Provide intuition for this apparently contradictory result. What happens when the natural rate of interest is introduced in the policy rule?

2. We now explore the MCMC algorithm and the Laplace approximation in a simple example. Technical details about both these objects are discussed in lecture notes.[2] One

[2] See http://faculty.wcas.northwestern.edu/~lchrist/course/Gerzensee_2013/estimationhandout.pdf

practical consideration not mentioned in the notes is relevant for the case in which the pdf of interest is of a non-negative random variable. Since the jump distribution is Normal, a negative candidate, x, is possible (see the notes for a detailed discussion of x and the 'jump distribution'). As a result, we should assign a zero value to the density of a Weibull over negative random variables when implementing the MCMC algorithm.

Hopefully, it is apparent that the MCMC algorithm is quite simple, and can be programmed by anyone with a relatively small exposure to MATLAB. A useful exercise to understand how the algorithm works, is to use it to see how well it approximates a simple known function. Thus, consider the Weibull probability distribution function (pdf),

$$ba^{-b}\theta^{b-1}e^{-\left(\frac{\theta}{a}\right)^b},\ \theta \geq 0,$$

where a, b are parameters. (For an explanation of this pdf, see the MATLAB documentation for $wblpdf(\theta, a, b)$.) Consider $a = 10$, $b = 20$. Graph this pdf over the grid, $[7, 11.5]$, with intervals 0.001 (i.e., graph g on the vertical axis, where $g = wblpdf(x, 10, 20)$, and x on the horizontal axis, where $x = 7 : .001 : 11.5$). Compute the mode of this pdf by finding the element in your grid with the highest value of g. Let f denote the log of the Weibull density function and compute the second derivative of f at the mode point numerically, using the formula,

$$f''(x) = \frac{f(x + 2\varepsilon) - 2f(x) + f(x - 2\varepsilon)}{4\varepsilon^2},$$

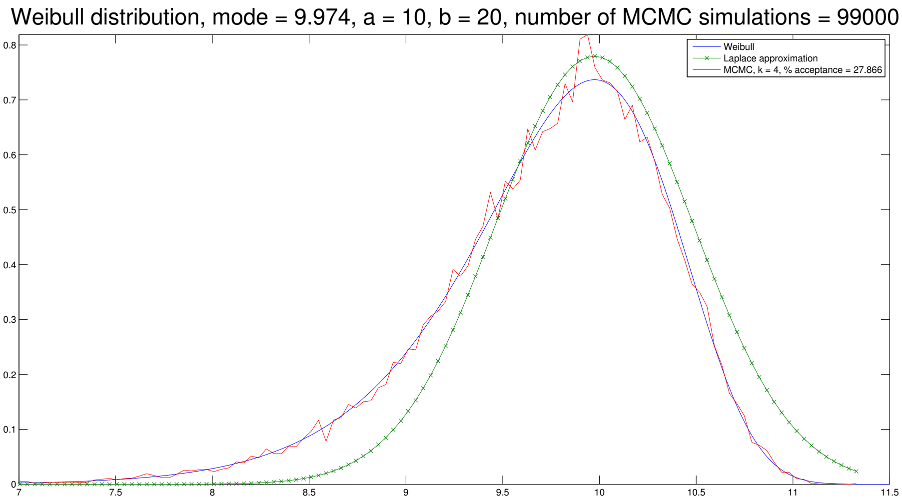
for ε small (for example, you could set $\varepsilon = 0.000001$.) Here, x denotes θ^* and f denotes the log of the output of the MATLAB function, $wblpdf$. Set $V = -f''(\theta^*)^{-1}$.[3]

Set $M = 1,000$ (a very small number!) and try $k = 2, 4, 6$. Which implies an acceptance rate closer to the recommended value of around 0.23? Choose the value of k that gets closest to that acceptance rate and note that the MCMC estimate of the distribution is quite volatile. Change M to 10,000. If you have time (now, the simulations takes time) try $M = 100,000$. Note how the MCMC estimate of the distribution is starting to smooth out. When I set $M = 100,000$ and $k = 4$, I obtained (see the MATLAB code MCMC.m, with the parameter iw set to unity) the following result:

[3] The strategy for computing the mode of the Weibull and f'' in the text are meant to resemble what is done in practice, when the form of the density function is unknown. In the case of the Weibull, these objects are straightforward to compute analytically. In particular,

$$f'(\theta) = \frac{b-1}{\theta} - b\left(\frac{\theta}{a}\right)^{b-1}\frac{1}{a},\ f''(\theta) = -\frac{b-1}{\theta^2} - (b-1)\,b\left(\frac{\theta}{a}\right)^{b-2}\frac{1}{a^2}.$$

and the mode of f is $\theta^* = ((b-1)/b)^{1/b}\,a$.

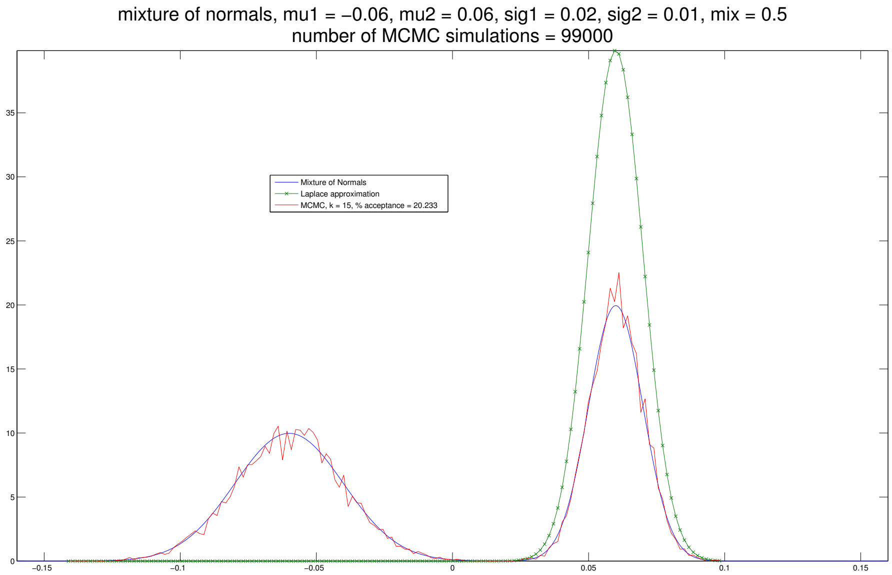


Note how well the MCMC approximation works. The Laplace approximation assigns too much density near the mode, and lacks the skewness of the Weibull. Still, for practical purposes the Laplace may be workable, at least as a first approximation in the initial stages of a research project. This could be verified in the early stages of the project by doing a run using the MCMC algorithm and comparing the results with those of the Laplace approximation. In practice, posterior distributions may not be as skewed as the Weibull is.

We subject the MCMC algorithm to a much tougher test if we posit that the true distribution is bimodal, as in the case of a mixture of two Normals. Suppose the i^{th} Normal has mean and variance, μ_i and σ_i^2, respectively, $i = 1, 2$. Suppose also that the $i = 1$ Normal is selected with probability, π, and the $i = 2$ normal is selected with probability $1 - \pi$. In addition, suppose

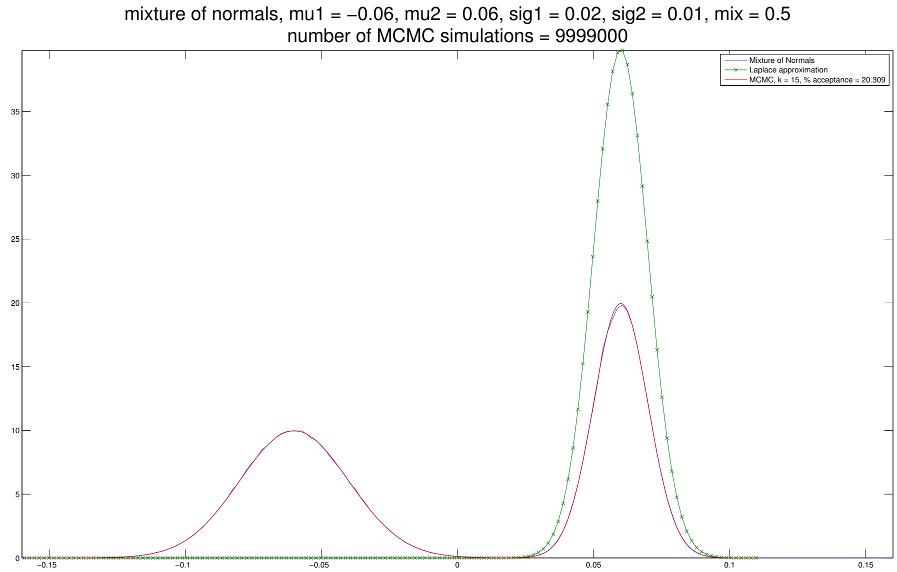
$$\mu_1 = -0.06,\ \mu_2 = 0.06,\ \sigma_1 = 0.02,\ \sigma_2 = 0.01,\ \pi = 1/2.$$

The mode of this distribution is the mode of the Normal with $i = 2$. If we apply exactly the same MCMC algorithm applied above, with $M = 100,000$ and $k = 15$, we obtain the following result:



These results (produced by running MCMC.m with iw set to zero) are comparable in accuracy to what was reported for the Weibull distribution. Taken together these sets of results suggest that the MCMC algorithm is quite good. It is not surprising that the Laplace approximation does poorly in this second example. It does a Normal approximation around the mode on the right. Because it ‘thinks’ that all the density is around that right mode and that density must integrate to unity, it follows that the Laplace approximation must rise up much higher than the right mode. To verify that the MCMC distribution in fact is converging to the right answer, the MCMC was run a second time with $M = 10,000,000$. The results are displayed in the following figure. Note that it is almost impossible to distinguish between the actual and the MCMC-generated distributions, so that the MCMC algorithm has roughly converged to the right answer. It is hard to say whether this bimodal example is empirically realistic. These kind of posterior distributions have not been reported in the literature. Of course, this may simply be that the MCMC has failed to find them even though they do exist.[4]

[4] An early paper by Thomas Sargent suggests that bimodality may be generic in dynamic macroeconomic models. He displays an example in which a parameterization in which persistence reflects the effects of endogenous mechanisms is hard to distinguish econometrically from a parameterization in which persistence reflects the persistence of shocks. See, Sagent, 1978, ”Estimation of Dynamic Labor Demand Schedules under Rational Expectations,” Journal of Political Economy, Vol. 86, No. 6, Dec., pp. 1009-1044.



3. From here on, consider the following alternative parameterization, which is more appealing than the one in question 1 from an empirical point of view:

$$\begin{aligned}\beta &= 0.97,\ \phi_x = 0.15,\ \phi_\pi = 1.5,\ \alpha = 0.8,\ \rho = 0.9,\ \lambda = 0.5,\\ \varphi &= 1,\ \theta = 0.75,\ \sigma_a = \sigma_\tau = 0.02.\end{aligned}$$

Generate $T = 4,000$ artificial observations on the 'endogenous' (in the sense of Dynare) variables of the model. These are the variables in the 'var' list. The mod file provided, cggsim.mod, has 6 variables. Before doing the simulation, you should add the growth rate of output to the equations of the model and to the var list (call it 'dy'.) That way, Dynare will also simulate output growth. The variables simulated by Dynare are placed in the $n \times T$ matrix, oo_.endo_simul.[5] The n rows of oo_.endo_simul correspond to the $n = 7$ variables in var, *listed in the order in which you have listed them in the var statement* from the first to the last row. To verify the order that Dynare puts the variables in, see how they are ordered in M_.endo_names in the Dynare-created file, cggsim.m.

Now do Bayesian estimation, using the inverted gamma distribution as the prior on the two standard deviations and the beta distribution as the prior on the two autocorrelations.

(a) Set the mean of the priors over the parameters to the corresponding true values. Set the standard deviation of the inverted gamma to 10 and of the beta to 0.04. (It's hard to interpret these standard deviations directly, but you will see graphs of the priors, which are easier to interpret.) Use 30 observations in the estimation. Adjust the value of k, so that you get a reasonable acceptance rate. I found that $k = 1.5$ works well. Have a look at the posteriors, and notice how, with one exception, they are much tighter than the priors. The exception is lambda, where the posterior and prior are very similar. This is evidence that there is little information in the data about lambda.

[5] Here, endo_simul is the matrix, which is a 'field' in the structure, oo_.

(b) Redo (a), but set the mean and standard deviation of the prior on lambda equal to 0.95 and 0.04, respectively. Note how the prior and posterior are again very similar. There is not much information in the data about the value of lambda!

(c) Note how the priors on σ_a and ρ have faint 'shoulders' on the right side. Redo (a), with $M = 4,000$ (M is mh_replic, which controls the number of MCMC replications). Note that the posteriors are now smoother. Actually, $M = 4,000$ is a small number of replications to use in practice.

(d) Now set the mean of the priors on the standard deviations to 0.1, far from the truth. Set the prior standard deviation on the inverted gamma distributions to 1. Keep the observations at 30, and see how the posteriors compare with the priors. (Reset $M = 1,000$ so that the computations go quickly.) Note that the posteriors move sharply back into the neighborhood of 0.02. Evidently, there is a lot of information in the data about these parameters.

(e) Repeat (a) with 4,000 observations. Compare the priors and posteriors. Note how, with one exception, the posteriors are 'spikes'. The exception, of course, is lambda. Still, the difference between the prior and posterior in this case indicates there is information in the data about lambda.