Christiano
Economics 416
Advanced Macroeconomics
Take home midterm exam.

1 Explaining Labor Market Volatility

The purpose of this question is to explore a labor market puzzle that has bedeviled business cycle researchers for years. The problem is to produce a sensible model that generates the amount of labor market volatility that we observe in the data. The first step is to show that there is a problem in the standard real business cycle model. The second step is to document a particular diagnosis of the problem, namely that it reflects excessive movement in the wage. The third step is to introduce firm/worker bargaining over the wage and show that this opens a possible route for solving the problem. The example is inspired by Hagedorn and Manovskii's 2008 AER paper in which they showed that if the unemployment compensation of the worker is high enough, then wages could be smooth enough and, hence, employment volatile enough, to match the data. Hagedorn and Manovskii's posited explanation has been criticized on the ground that real-world agents' outside option is not as great as Hagedorn and Manovskii's explanation requires. The problem is that if workers' outside option is reduced to levels that the critics argue is empirically plausible, then it is claimed (see Shimer's 2005 AER paper) that the bargaining model loses the ability to account for the volatility of labor markets. We will pursue these ideas further later in the course, by drawing attention to the observations in Hall and Milgrom's 2008 AER paper.

1.1 Real Business Cycle Model

Consider the following real business cycl model. At time t, the representative household maximizes

$$E_t \sum_{j=0}^{\infty} \beta^j \left[\log C_{t+j} + \psi \log\left(1 - N_{t+j}\right)\right],$$

subject to

$$C_t + K_{t+1} - (1-\delta) K_t \leq r_t K_t + w_t N_t,$$

for all t. Here, r_t denotes the rental rate of capital and w_t denotes the wage rate. The household is 'small' and takes the prices as given. There is a representative firm. In period t, the firm maximizes by choice of K_t and N_t, its profits:

$$Y_t - r_t K_t - w_t N_t,$$

subject to its technology:

$$Y_t = K_t^{\alpha} \left[\exp\left(a_t\right) N_t\right]^{1-\alpha},$$

where

$$a_t = \rho_a a_{t-1} + \varepsilon_t.$$

Here, ε_t is iid with mean zero and $E\varepsilon_t^2 = \sigma^2 = 0.01^2$, $\rho_a = 0.95$. Also, $\beta = 1.03^{-1/4}$, $\alpha = 0.36$, $\delta = 0.025$. Finally, assign a value to ψ which implies that $N_t = 1/3$ in steady state, given the setting of the other parameters. That is, the representative household works one-third of available time.

1.1.1 Questions

1. Use Dynare to solve the model and simulate 1,000 observations on log output and log employment (work with the 'periods=1000' command in stoch_simul). Detrend these two series using the HP filter. Compute the standard deviation of the result. Display the ratio of the standard deviation of (filtered) employment to the standard deviation of (filtered) output. This ratio is call the 'relative volatility of employment to output'.

2. Go to the web-based data base of the Federal Reserve Bank of St. Louis (FRED) and retrieve data on Real Gross Domestic Product (GDP) and employment, All Employees: Total nonfarm (take these data quarterly). Do to these data what you did to the model data. Display the relative volatility of employment to output in the data.

3. You will see that the relative volatility of employment is much higher in the data than it is in the model. This failing of the model has attracted a lot of attention. One interpretation is that it reflects wages rise too much in the wake of a shock that causes output to expand. Explore this hypothesis by returning to question 1. Fix the wage rate exogenously at its steady state value and assume that firms are always on their labor demand schedule, while households always supply all the labor that is demanded (this implies that sometimes they work more than they want). What happens to the volatility of employment with this change in the model? Much of the macro labor supply literature is about trying to reproduce the properties of this sticky wage model. But, economists prefer if they can arrive at this by some endogenous mechanism.

1.2 Real Business Cycle Model with Nash Bargaining

Assume that the representative household has a unit mass of workers. Each worker goes to the labor market. A fraction, N_t, of the workers meet a firm and are employed. The complementary fraction is unemployed. There is perfect insurance inside the household and each worker enjoyed the same level of consumption, C_t. Each employed worker brings home the wage, w_t, and each unemployed worker brings home an unemployment payment, D. The household problem is to maximize:

$$E_0 \sum_{t=0}^{\infty} \beta^t \log (C_t),$$

subject to

$$C_t + K_{t+1} - (1-\delta) K_t \leq w_t N_t + r_t K_t + (1 - N_t) D - T_t.$$

Here, T_t denotes taxes raised to pay for government unemployment payments. The government budget constraint is:

$$(1 - N_t) D = T_t.$$

Workers are instructed by households to be interested in the present discounted value of the resources they bring home. Workers do not suffer any personal disutility from working. Thus, a period t employed worker enjoys utility, V_t, where

$$V_t = w_t + E_t m_{t+1} \left[\rho V_{t+1} + (1-\rho)(f_{t+1} V_{t+1} + (1 - f_{t+1}) U_{t+1}) \right]. \quad (1)$$

Here, m_{t+1} denotes the household's stochastic discount factor and the object in square brackets indicates the various things that could happen to the worker in period $t+1$. These include that the worker will remain matched to the same firm in $t+1$ with probability ρ. With probability $1-\rho$ the worker will separate from the firm. In that case, there are two possibilities. With probability f_{t+1} the worker matches immediately with another firm. With probability $1 - f_{t+1}$ the worker goes into unemployment in $t+1$.

The value of being an unemployed worker is:

$$U_t = D + E_t m_{t+1} \left[f_{t+1} V_{t+1} + (1 - f_{t+1}) U_{t+1} \right]. \quad (2)$$

A worker who is unemployed in period t receives a payment, D. In period $t+1$ the period t unemployed worker is employed with probability f_{t+1} and is unemployed with probability $1 - f_{t+1}$.

There are two types of firms in this variant of the model. There are the firms (we'll call them *RBC firms*) that look just like their cousins in the real business cycle model. They operate the Cobb-Douglas production function. To do so, they rent capital, K_t, and a second input which we denote by h_t. The RBC firms hire K_t and h_t, in competitive markets. The second type of firm, we call them the *bargaining firms*, are endowed with the knowledge of how to convert one unit of labor power into one unit of h_t. RBC firms and bargaining firms interact in competitive markets.

We now discuss bargaining firms in greater detail. The value to the firm of an employed worker is denoted J_t :

$$J_t = \vartheta_t - w_t + \rho E_t m_{t+1} J_{t+1}, \quad (3)$$

where ϑ_t denotes the (competitively determined) market value of the one unit of h_t produced by the worker and sold by the bargaining firm.

The number of employed workers in period t is denoted l_t and this evolves as follows:

$$l_t = (\rho + x_t) l_{t-1}. \quad (4)$$

Here, ρ corresponds to the exogenous rate at which employed workers are separated from their firms at the end of the period. Also, x_t denotes the hiring rate so that the number of new hires in period t is equal to $x_t l_{t-1}$. Note that the job finding rate is given by,

$$f_t = \frac{x_t l_{t-1}}{1 - \rho l_{t-1}}, \tag{5}$$

where the numerator is the number of workers that are newly-hired at the beginning of time t, while the denominator is the number of workers who are searching for work at the end of time $t-1$. The denominator term is arrived at as follows. The number of workers that are unemployed during period $t-1$ is $1 - l_{t-1}$. In addition, $(1-\rho)\, l_{t-1}$ workers become separated from their firm at the end of $t-1$. Thus, the total number of people searching at the end of $t-1$ is

$$1 - l_{t-1} + (1-\rho)\, l_{t-1} = 1 - \rho l_{t-1}.$$

A firm that wishes to meet with a worker can do so by paying a fixed cost, κ. Free entry implies that bargaining firms cannot make profits by hiring a worker, so that

$$J_t = \kappa$$

We assume that the wage is determined by Nash bargaining:

$$J_t = \frac{1-\eta}{\eta}(V_t - U_t).$$

Here, η is the share of the total surplus, $J_t + V_t - U_t$, given to workers. The object, J_t, is the surplus of the firm. It is what the firm gets by employing the worker (i.e, J_t), minus what it gets if it does not employ the worker (i.e., nothing). Similarly for the surplus of the workers.

The resource constraint in this economy is:

$$C_t + \kappa x_t l_{t-1} + K_{t+1} - (1-\delta)\, K_t \leq K_t^{\alpha} \left[\exp(a_t)\, h_t\right]^{1-\alpha},$$

for all t. Here, $\kappa x_t l_{t-1}$ denotes the goods purchased by bargaining firms to address their hiring costs. Also, h_t is the quantity of input goods purchased by RBC firms from bargaining firms. Clearing in that market requires

$$l_t = h_t.$$

To save notation, you may just as well use l_t wherever h_t appears, so that the resource constraint is:

$$C_t + \kappa x_t l_{t-1} + K_{t+1} - (1-\delta)\, K_t \leq K_t^{\alpha} \left[\exp(a_t)\, l_t\right]^{1-\alpha}.$$

Finally, there is the discount factor, m_{t+1} :

$$m_{t+1} = \beta \frac{C_t}{C_{t+1}}.$$

1.2.1 Questions

1. The parameters of the RBC part of the model are, as before,

$$\alpha = 0.36,\ \beta = 1.03^{-1/4},\ \rho_a = 0.95,\ E\varepsilon_t^2 = 0.01^2,\ \delta = 0.025.$$

   Set the persistence of job matches, ρ, to 0.90. The other parameters of the bargaining part of the model,

$$D, \kappa, \eta,$$

   should be set so that, given the other parameter values, the following is true in steady state:

$$\begin{array}{rl} \frac{\kappa x l}{GDP} = 0.01 & \text{hiring costs as a fraction of GDP } (\equiv C + \delta K) \\ \frac{D}{w} = 0.98 & \text{replacement ratio} \\ u = 1 - l = 0.055 & \text{unemployment rate} \end{array}$$

   Display formulas for the steady state of the model, including for D, κ, η. Report the steady state values of $K, C, x, w, \vartheta, U, V, f, r$.

2. Generate 1,000 observations on log GDP and log employment in the model, HP filter the result as you did in the RBC model and compute the standard deviations of the result. Does this model do a better job at accounting for the observed volatility of labor? To help answer this question, compute and display on the same graph the impulse response to technology shock implied by the RBC model and the RBC model with bargaining.

3. Discuss the role played in the analysis of the high replacement ratio. Do this by repeating 2 with replacement ratios of 0.99 and 0.97.

2 Monte Carlo Markov Chain

The idea here is to explore the accuracy of the Metropolis-Hastings algorithm for approximating a distribution. We'll specify a density function for a single random variable. Since we know the density function we're trying to approximate, we'll be able to assess the quality of the approximation provided by the MH algorithm. In applying the algorithm, we'll need to compute a second derivative. Do this numerically. Here is a formula for a function, $f(x)$:

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

for ε small (for example, you could set $\varepsilon = 0.000001$.)

In this exercise we make the test of the MH algorithm pretty tough by specifying that the true density function is bimodal. At the same time, the test will be very weak because we are working with a one-dimensional random variable. We we consider a mixture of normal distribution. This density function is a linear combination of two normals, where the weights in the linear combination are denoted π and $1-\pi$. The object, π, is the probability that a variable is drawn from the first Normal distribution, and $1-\pi$ is the probability of drawing from the second Normal distribution. Suppose the i^{th} Normal has mean and variance, μ_i and σ_i^2, respectively, $i=1,2$. In addition, suppose

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

1. Let x denote the mixture of Normals random variable. Graph its density over the range, $x=-0.15$ to $x=0.15$. Specify a very fine grid of values of x so that you get a very accurate graph of the mixture of Normals. If the grid is fine enough then you can pinpoint the value of x, x^*, where the density is highest by simply identifying the value of x on your grid that produces highest density (i.e., you can just use the max operator in MATLAB).

2. Compute the second derivative of the density function around the mode, x^*. Apply the algorithm described in the handout to compute a sequence, $x^{(1)}, x^{(2)}, ..., x^{(M)}$, where $x^{(1)} = x^*$. Choose the scalar in the jump distribution so that you get a rejection rate of 23 percent when $M = 1,000$. Graph the histogram of $x^{(1)}, x^{(2)}, ..., x^{(M)}$. Scale the heights of the histogram bars so that the sum of the areas of the histogram rectangles equals unity. Only with this scaling will the histogram be comparable to the true density function, the mixture of Normals. Graph the properly scaled historgram on the same graph with the density of the true mixture of Normals you graphed in question 1. Don't make the graph a bar chart, make it a line graph so it looks like a density function. Does the MH algorithm produce a good-looking approximation?

3. In question 2, you should have found that the histogram is rather choppy. Now increase M to 100,000. You should find that you get an amazingly accurate approximation. Try different values of k, fixing M at, say 5,000. Does k affect accuracy much?

4. Center the MCMC algorithm around the second, lower, peak of the density function. Use $M = 100,000$. Does it make much difference that you did not center things on the actual mode?

5. Draw the Laplace approximation of the density function around the mode of the mixture of Normals (see the later sheets on the Bayesian inference handout for the Laplace approximation). How well does that approximate the distribution?

3 Bayesian Estimation I

Suppose the analyst has a sample of T iid Normal observations, $y = [y_1, ..., y_T]'$. To make the problem simple, suppose the analysis knows the value of the variance, σ^2. The likelihood of these observations is

$$P(y|\mu) = \frac{1}{(2\pi\sigma^2)^{T/2}} \exp\left[-\frac{1}{2}\sum_{t=1}^{T}\frac{(y_t-\mu)^2}{\sigma^2}\right].$$

The analyst does not know the value of μ. But, before seeing y, she had a prior about it, $P(\mu)$. This prior takes the following form:

$$P(\mu) = \frac{1}{2\pi\eta^2} \exp\left[-\frac{1}{2}\frac{(\mu-m)^2}{\eta^2}\right],$$

where m denotes the mean of the prior and its variance is η^2. The posterior over μ, after observing y, is denoted $P(\mu|y)$, where

$$P(\mu|y) \propto P(y|\mu)\, P(\mu),$$

where $\propto$ means 'is proportional to'. Let μ^* denote the mode of the posterior distribution, i.e.,

$$\mu^* = \arg\max_{\mu} P(\mu|y).$$

The maximum likelihood estimator of μ, $\hat{\mu}$, is

$$\hat{\mu} = \arg\max_{\mu} P(y|\mu).$$

1. Show that

$$\hat{\mu} = \bar{y},\ \ \bar{y} \equiv \frac{1}{T}\sum_{t=1}^{T} y_t.$$

2. Show that μ^* can be written in the form:

$$\mu^* = \alpha m + (1-\alpha)\bar{y},$$

where

$$\begin{aligned} \alpha &\rightarrow 0 \text{ as } T \rightarrow \infty \\ \alpha &\rightarrow 1 \text{ as } \eta \rightarrow 0. \end{aligned}$$

That is, as the precision of the prior increases ($\eta \rightarrow 0$), the data are ignored. As there are more data ($T \rightarrow \infty$) the prior is ignored. Derive an explicit formula for α in terms of T, σ, η.

4 Bayesian Estimation II

Consider the NK model described in class, in which the technology shock process has a difference stationary representation. Assign the following parameter values:

$$\begin{aligned} \beta &= 0.97,\ \phi_x = 0.15,\ \phi_\pi = 1.5,\ \alpha = 0.8,\ \rho = 0.9,\ \lambda = 0.5,\ \delta = 0.2, \\ \varphi &= 1,\ \theta = 0.75,\ \sigma_a = \sigma_\tau = 0.02. \end{aligned}$$

Generate $T = 4,000$ artificial observations on the variables using Dynare. Make sure to include output growth in the model, as well as inflation. (Output growth is the first difference of the log of the output gap, x_t, plus the first difference of the log of first best output.) These are generated by the stoch_simul command by including periods=4000 in the argument list of that command. You can find the simulated variables in an $N \times 4000$ matrix after the stoch_simul command in the matrix, oo_.endo_simul. The rows of this matrix correspond to the variables in the var command. Save that matrix as a MATLAB file. You need a Dynare program which loads the MATLAB file with the simulated data and estimates the model using only the inflation and output growth data.

The Dynare code for estimation is the same as the code you used to simulate the data, up to the stoch_simul command. Instead of the stoch_simul command, you use the estimate command. Before the estimate command you must have a varobs command which indicates which variables are to be used in the estimation, and another command where you specify the parameters to be estimated and the associated priors. Estimate the four parameters that govern the exogenous shocks. (All other parameters should be set to their true values, though for fun you could study the case where the econometrician makes the wrong assumption about the value of the other parameters.) Standard priors for standard deviations are the inverted gamma distribution. The best way to assess the reasonableness of this is to stare at the graph of the prior when Dynare generates it. A standard prior for autoregressive coefficients that you think ought to be positive is the beta distribution, because it is bounded on the interval between 0 and 1. An (incomplete) example of the required code is supplied as part of this take home exam.

1. 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 and set $M = 500$ (M corresponds to mh_replic). Adjust the value of k, so that you get a reasonable acceptance rate. Display all graphs in your answer to the midterm. 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 (this is a general feature of NK models: the dynamics of the data are relatively invariant to things on the supply side, it's almost like those models had dropped the labor supply curve). Graph the priors and posteriors. When you do the estimation, be sure to include mode_check as an argument. This triggers the graphing of the posterior and the likelihood for each parameter, holding the others at their mode. This allows you to verify visually that the mode has, in fact been achieved. It also allows you to see separately the role of the likelihood (i.e., the function of the data) in determining the curvature of the posterior. For example, in the case of λ the posterior is primarily driven by the prior, so you should see the

likelihood being very flat.

2. Redo 1., 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!

3. Note how the priors on σ_a and ρ have 'shoulders' on the right side. Redo 1., with $M = 4,000$. Note that the posteriors are now smoother. Actually, $M = 4,000$ is a small number of replications to use in practice.

4. 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.

5. Repeat 1. using 4,500 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.