Graduate Macro Theory II: Notes on Medium Scale DSGE Models

Eric Sims
University of Notre Dame

Spring 2011

1 Introduction

These notes introduce and describe a "medium scale" DSGE model. The model features Calvo price-setting but has capital, variable utilization, habit formation, and investment adjustment costs.

2 Households

Households in this model consume goods, supply labor, hold money, and save through both bonds and capital (the households own the capital stock). The household block is a little different from how we've been working thus far. We assume that there is variable utilization of capital, u_t. We assume that the household chooses u_t and then leases capital services, $\widehat{k}_t = u_t k_t$ to the firms, who in turn pay a rental rate, R_t, for those services. This ends up being a modeling assumption that gives rise to the same first order conditions we would otherwise get if the firms owned the capital stock. Second, we assume that there is habit formation in consumption. Third, we assume that there are investment adjustment costs. Fourth, we assume that there are is a preference shock to the utility of leisure.

The full household problem can be characterized as follows:

$$\max_{c_t,n_t,k_{t+1},u_t,M_{t+1},B_{t+1}} \quad E_0 \sum_{t=0}^{\infty} \beta^t \left(\ln(c_t - \gamma c_{t-1}) + \theta_t \frac{(1-n_t)^{1-\xi} - 1}{1-\xi} + \frac{\left(\frac{M_{t+1}}{p_t}\right)^{1-\nu} - 1}{1-\nu} \right)$$

s.t.

$$c_t + I_t + \frac{B_{t+1} - B_t}{p_t} + \frac{M_{t+1} - M_t}{p_t} \leq w_t n_t + R_t u_t k_t + i_t \frac{B_t}{p_t} + \text{Profit}_t - T_t$$

$$k_{t+1} = \left(1 - \frac{\tau}{2}\left(\frac{I_t}{I_{t-1}} - 1\right)^2\right) I_t + \left(1 - \delta_0 u_t^{\Delta}\right) k_t$$

Set this problem up as a Lagrangian:

$$\mathcal{L} = E_t \sum_{t=0}^{\infty} \beta^t \left\{ \ln(c_t - \gamma c_{t-1}) + \theta_t \frac{(1-n_t)^{1-\xi} - 1}{1-\xi} + \frac{\left(\frac{M_{t+1}}{p_t}\right)^{1-\nu} - 1}{1-\nu} + \dots \right.$$
$$\dots + \lambda_t \left(w_t n_t + R_t u_t k_t + (1+i_t)\frac{B_t}{p_t} + \text{Profit}_t - T_t - c_t - I_t - \frac{M_{t+1}}{p_t} + \frac{M_t}{p_t} - \frac{B_{t+1}}{p_t} \right) + \dots$$
$$\left. \dots + \mu_t \left(\left(1 - \frac{\tau}{2} \left(\frac{I_t}{I_{t-1}} - 1 \right)^2 \right) I_t + \left(1 - \delta_0 u_t^{\Delta}\right) k_t - k_{t+1} \right) \right\}$$

The first order conditions are as follows:

$$\frac{\partial \mathcal{L}}{\partial c_t} = 0 \Leftrightarrow \frac{1}{c_t - \gamma c_{t-1}} - E_t \frac{\beta\gamma}{c_{t+1} - \gamma c_t} = \lambda_t \quad (1)$$
$$\frac{\partial \mathcal{L}}{\partial n_t} = 0 \Leftrightarrow \theta_t (1-n_t)^{-\xi} = \lambda_t w_t \quad (2)$$
$$\frac{\partial \mathcal{L}}{\partial u_t} = 0 \Leftrightarrow \lambda_t R_t = \mu_t \Delta \delta_0 u_t^{\Delta - 1} \quad (3)$$
$$\frac{\partial \mathcal{L}}{\partial B_{t+1}} = 0 \Leftrightarrow \lambda_t = \beta E_t \lambda_{t+1} (1 + i_{t+1}) \frac{p_t}{p_{t+1}} \quad (4)$$
$$\frac{\partial \mathcal{L}}{\partial K_{t+1}} = 0 \Leftrightarrow \mu_t = \beta E_t \left(\lambda_{t+1} R_{t+1} u_{t+1} + \mu_{t+1} (1 - \delta_0 u_{t+1}^{\Delta}) \right) \quad (5)$$
$$\frac{\partial \mathcal{L}}{\partial M_{t+1}} = 0 \Leftrightarrow m_t^{-\nu} = \lambda_t - \beta E_t \lambda_{t+1} \frac{p_t}{p_{t+1}} \quad (6)$$
$$\frac{\partial \mathcal{L}}{\partial I_t} = 0 \Leftrightarrow \lambda_t = \mu_t \left(1 - \frac{\tau}{2} \left(\frac{I_t}{I_{t-1}} - 1 \right)^2 - \tau \left(\frac{I_t}{I_{t-1}} - 1 \right) \frac{I_t}{I_{t-1}} \right) + \beta E_t \mu_{t+1} \tau \left(\frac{I_{t+1}}{I_t} - 1 \right) \left(\frac{I_{t+1}}{I_t} \right)^2 \quad (7)$$

These conditions are just generalizations of things we've already seen. (1) is the definition of the marginal utility of income when there is habit formation. (2) is a standard labor supply condition. (3) actually turns out to be exactly the same first order condition for utilization that we saw before if $\lambda_t = \mu_t$ (which will hold in steady state and would hold if there are no adjustment costs) and if there were no monopolistic competition. (4) is the standard first order condition for bonds. (6) is the standard demand for real balances, $m_t \equiv \frac{M_{t+1}}{p_t}$. (6) and (7) are also the same first order conditions for capital and investment that would obtain if the firms owned the capital stock. We will define another variable to correspond to the price of capital (measured in terms of goods):

$$q_t = \frac{\mu_t}{\lambda_t} \quad (8)$$

This is the marginal value (measured in goods) of an additional unit of installed capital.

3 Firms

As before, production is split into two sectors – a competitive final goods sector and a monopolistically competitive intermediate goods sector.

3.1 Final Goods

The final good is a CES aggregate of intermediate goods. There are a continuum of intermediate goods indexed by j over the unit interval:

$$y_t = \left(\int_0^1 y_{j,t}^{\frac{\varepsilon-1}{\varepsilon}} dj \right)^{\frac{\varepsilon}{\varepsilon-1}} \tag{9}$$

As we have previously seen, profit maximization by the final goods firm implies a downward sloping demand curve for each intermediate good and an aggregate price index:

$$y_{j,t} = \left(\frac{p_{j,t}}{p_t} \right)^{-\varepsilon} y_t \tag{10}$$

$$p_t = \left(\int_0^1 p_{j,t}^{1-\varepsilon} dj \right)^{\frac{1}{1-\varepsilon}} \tag{11}$$

3.2 Intermediate Goods

Intermediate goods firms produce output using capital services and labor. There is a technology shifter that is common across firms. Define capital services as $\widehat{k}_t = u_t k_t$. The intermediate goods firm cannot choose utilization and capital separately (those are determined by the household). The production function is:

$$y_{j,t} = a_t \widehat{k}_{j,t}^{\alpha} n_{j,t}^{1-\alpha} \tag{12}$$

The intermediate goods firm cannot freely adjust its prices period by period. Following Calvo (1983), it faces a constant hazard, $1 - \phi$, of being able to adjust its price in any period. Hence, it will not necessarily be able to maximize profits every period. It will, however, nevertheless find it optimal to minimize costs regardless of the price of its good. Hence, we can break the problem up into two parts. The firms are price-takers in input markets, facing nominal wage $w_t p_t$ and nominal rental rate $R_t p_t$ (w_t and R_t are the real factor prices). The cost minimization problem is:

$$\min_{n_{j,t}, \widehat{k}_{j,t}} \quad w_t p_t n_{j,t} + R_t p_t \widehat{k}_{j,t}$$

$$\text{s.t.}$$

$$a_t \widehat{k}_{j,t}^{\alpha} n_{j,t}^{1-\alpha} \geq \left(\frac{p_{j,t}}{p_t} \right)^{-\varepsilon} y_t$$

Form a Lagrangian:

$$\mathcal{L} = -w_t p_t n_{j,t} - R_t p_t \widehat{k}_{j,t} + \varphi_{j,t} \left(a_t \widehat{k}_{j,t}^{\alpha} n_{j,t}^{1-\alpha} - \left(\frac{p_{j,t}}{p_t} \right)^{-\varepsilon} y_t \right)$$

The first order conditions are:

$$\frac{\partial \mathcal{L}}{\partial n_{j,t}} = 0 \Leftrightarrow w_t = \frac{\varphi_{j,t}}{p_t}(1-\alpha) a_t \widehat{k}_{j,t}^{\alpha} n_{j,t}^{-\alpha} \tag{13}$$

$$\frac{\partial \mathcal{L}}{\partial \widehat{k}_{j,t}} = 0 \Leftrightarrow R_t = \frac{\varphi_{j,t}}{p_t} \alpha a_t \widehat{k}_{j,t}^{\alpha-1} n_{j,t}^{1-\alpha} \tag{14}$$

These conditions say to equate the real factor prices with real marginal cost times the marginal products. These conditions can be combined to derive an expression relating the ratio of capital services to labor to factor prices and α:

$$\frac{\widehat{k}_{j,t}}{n_{j,t}} = \frac{w_t}{R_t} \frac{\alpha}{1-\alpha} \tag{15}$$

Notice that none of the terms on the right hand side depend on j. Hence, the capital labor ratio will be equal across all firms, which in turn will be equal to the aggregate ratio: $\frac{\widehat{k}_{j,t}}{n_{j,t}} = \frac{\widehat{k}_t}{n_t}\ \forall j$. Since all firms will hire capital services and labor in the same ratio, this means that marginal cost will be equal across firms. Call $mc_t \equiv \frac{\varphi_{j,t}}{p_t} = \frac{\varphi_t}{p_t}\ \forall j$. Then:

$$mc_t = w_t^{1-\alpha} \left(\frac{1}{1-\alpha} \right)^{1-\alpha} \left(\frac{1}{\alpha} \right)^{\alpha} \frac{R_t^{\alpha}}{a_t} \tag{16}$$

Now that we've taken care of factor demand, let's consider the pricing problem of a firm that gets to update its price in period t. It wants to choose its prices to maximize the present discounted value of profits, where it discounts by the stochastic discount factor, $M_t = \beta\lambda_t$. Current (nominal) profits are:

$$\text{Profit}_{j,t} = p_{j,t} y_{j,t} - w_t p_t n_{j,t} - R_{j,t} p_t \widehat{k}_{j,t}$$

The first order conditions for optimal choice of capital and labor can be written:

$$w_t p_t = \varphi_t (1-\alpha) \frac{y_{j,t}}{n_{j,t}} \Rightarrow w_t p_t n_{j,t} = \varphi_t (1-\alpha) y_{j,t}$$

$$R_t p_t = \varphi_t \alpha \frac{y_{j,t}}{\widehat{k}_{j,t}} \Rightarrow R_t p_t \widehat{k}_{j,t} = \varphi_t \alpha y_{j,t}$$

This means I can write current profits as:

$$\text{Profit}_{j,t} = p_{j,t} y_{j,t} - \varphi_t y_{j,t}$$

The firm will want to maximize the real profits it returns to households, and so real profits can be written:

$$\frac{\text{Profit}_{j,t}}{p_t} = \frac{p_{j,t}}{p_t} y_{j,t} - mc_t y_{j,t}$$

In addition to the stochastic discount factor, firms will also discount future profits by ϕ^s, since this represents the probability that a price chosen at time t is still in effect at time s. The profit maximization problem can then be written:

$$\max_{p_{j,t}} \quad E_t \sum_{s=0}^{\infty} (\phi\beta)^s \lambda_{t+s} \left(\left(\frac{p_{j,t}}{p_{t+s}} \right)^{1-\varepsilon} y_{t+s} - mc_{t+s} \left(\frac{p_{j,t}}{p_{t+s}} \right)^{-\varepsilon} y_{t+s} \right)$$

I can re-write this problem in somewhat more compact fashion (so as to facilitate taking the derivative) as:

$$\max_{p_{j,t}} \quad E_t \sum_{s=0}^{\infty} (\phi\beta)^s \lambda_{t+s} p_{t+s}^{\varepsilon} y_{t+s} \left(p_{t+s}^{-1} p_{j,t}^{1-\varepsilon} - mc_{t+s} p_{j,t}^{-\varepsilon} \right)$$

Re-write this in terms of nominal marginal cost, $\varphi_{t+s} = mc_{t+s} p_{t+s}$, so as to simplify further:

$$\max_{p_{j,t}} \quad E_t \sum_{s=0}^{\infty} (\phi\beta)^s \lambda_{t+s} p_{t+s}^{\varepsilon-1} y_{t+s} \left(p_{j,t}^{1-\varepsilon} - \varphi_{t+s} p_{j,t}^{-\varepsilon} \right)$$

The first order condition is:

$$E_t \sum_{s=0}^{\infty} (\phi\beta)^s \lambda_{t+s} p_{t+s}^{\varepsilon-1} y_{t+s} \left((1-\varepsilon) p_{j,t}^{-\varepsilon} + \varepsilon \varphi_{t+s} p_{j,t}^{-\varepsilon-1} \right) = 0$$

This simplifies to:

$$p_t^{\#} = \frac{\varepsilon}{\varepsilon - 1} E_t \frac{\sum_{s=0}^{\infty} (\phi\beta)^s \lambda_{t+s} p_{t+s}^{\varepsilon} y_{t+s} mc_{t+s}}{\sum_{s=0}^{\infty} (\phi\beta)^s \lambda_{t+s} p_{t+s}^{\varepsilon-1} y_{t+s}} \tag{17}$$

Note that I have re-written the problem in terms of real marginal cost, not nominal, and have gone ahead and imposed that the rest price, $p_t^{\#}$, is the same across j, since nothing on the right hand side depends upon j.

We can write this recursively as:

$$\begin{aligned} p_t^{\#} &= \frac{\varepsilon}{\varepsilon - 1} E_t \frac{A_t}{D_t} \\ A_t &= \lambda_t p_t^{\varepsilon} y_t mc_t + \phi\beta E_t A_{t+1} \\ D_t &= \lambda_t p_t^{\varepsilon-1} y_t + \phi\beta E_t D_{t+1} \end{aligned}$$

Now define $\widehat{A}_t \equiv \frac{A_t}{p_t^{\varepsilon}}$ and $\widehat{D}_t \equiv \frac{D_t}{p_t^{\varepsilon-1}}$. Then we have:

$$\widehat{A}_t = \lambda_t y_t mc_t + \phi\beta E_t \left(\frac{p_{t+1}}{p_t}\right)^{\varepsilon} \widehat{A}_{t+1}$$
$$\widehat{D}_t = \lambda_t y_t + \phi\beta E_t \left(\frac{p_{t+1}}{p_t}\right)^{\varepsilon-1} \widehat{D}_{t+1}$$

This means that we can write the optimal reset price as:

$$p_t^{\#} = \left(\frac{\varepsilon}{\varepsilon-1}\right) \frac{p_t^{\varepsilon}}{p_t^{\varepsilon-1}} E_t \frac{\widehat{A}_t}{\widehat{D}_t}$$

This simplifies to:

$$\frac{p_t^{\#}}{p_t} = \left(\frac{\varepsilon}{\varepsilon-1}\right) E_t \frac{\widehat{A}_t}{\widehat{D}_t} \tag{18}$$

4 The Government

The government does three things in this model. The fiscal side of the government (1) consumes some private output, g_t; (2) levies lump sum taxes, T_t; and (3) issues debt, d_t, which pays interest i_t. On the monetary side, it sets nominal interest rates according to a Taylor type rule.

The Taylor rule is written as a partial adjustment rule in which nominal interest rates react positively to deviations of inflation from steady state inflation and positively to deviations in output growth from trend (trend output growth here is implicitly equal to zero since I have not modeled the trend, which ends up being innocuous):

$$i_{t+1} = \rho_i i_t + (1-\rho_i)\phi_\pi(\pi_t - \pi^*) + (1-\rho_i)\phi_y\left(\frac{y_t}{y_{t-1}} - 1\right) + \varepsilon_{i,t} \tag{19}$$

Recall the (somewhat unfortunate on my end) timing convention here – i_{t+1} is the interest rate on bonds *today* that pay off in period $t+1$. Hence it is known (and set) at time t. Given the nominal interest rate it chooses, the Fed adjusts the money supply so as to achieve equilibrium in the money market.

The government budget constraint says that spending plus payment of interest on existing debt must equal revenue collection plus issuance of new debt plus seignorage revenue:

$$g_t + i_t \frac{d_t}{p_t} = T_t + \frac{d_{t+1} - d_t}{p_t} + \frac{M_{t+1} - M_t}{p_t} \tag{20}$$

The terms $\frac{M_{t+1}-M_t}{p_t}$ is seignorage revenue. The government essentially earns revenue by printing more money (in real terms). We assume that the central bank returns this revenue back to the fiscal authority. The model turns out to be Ricardian in the sense that it does not matter how the government finances its spending between taxes and bonds.

5 Exogenous Processes

The exogenous variables of the model are a_t, θ_t, and g_t. I assume that each of them follow AR(1) processes in the logs:

$$\ln a_t = \rho_a \ln a_{t-1} + \varepsilon_{a,t} \tag{21}$$

$$\ln \theta_t = (1-\rho_\theta)\ln \theta^* + \rho_\theta \ln \theta_{t-1} + \varepsilon_{\theta,t} \tag{22}$$

$$\ln g_t = (1-\rho_g)(\ln \omega + \ln y^*) + \rho_g \ln g_{t-1} + \varepsilon_{g,t} \tag{23}$$

Here it is assumed that the unconditional mean of a_t is 1 (equivalently the unconditional mean of the log is zero). The unconditional mean of θ_t is θ^*, and the unconditional mean of g_t is ωy^*, where $0 \leq \omega < 1$ is the steady state share of government spending in output. ρ_a, ρ_g, and ρ_θ are all assumed to be between 0 and 1. Note that M_{t+1} is *not* an exogenous variable once the model is closed with a Taylor rule – the central bank must endogenously adjust M_{t+1} so as to achieve its interest rate target, given money demand.

The model has four exogenous stochastic shocks: $\varepsilon_{i,t}$, the monetary policy shock; $\varepsilon_{a,t}$, the technology shock; $\varepsilon_{\theta,t}$, the preference shock; and $\varepsilon_{g,t}$, the government spending shock.

6 Aggregation

Let's begin by summarizing the aggregate demand block of the model and then move onto the supply block. Bond market-clearing requires that $d_t = B_t$ (i.e. government debt is held by households). This means that the government budget constraint can be written:

$$T_t = g_t + (1+i_t)\frac{B_t}{p_t} - \frac{B_{t+1}}{p_t} - \frac{M_{t+1}-M_t}{p_t}$$

Now go to the household budget constraint, and plug in for T_t:

$$c_t + I_t + \frac{B_{t+1}}{p_t} + \frac{M_{t+1}-M_t}{p_t} = w_t n_t + R_t u_t k_t + \text{Profit}_t + (1+i_t)\frac{B_t}{p_t} - \left(g_t + (1+i_t)\frac{B_t}{p_t} - \frac{B_{t+1}}{p_t} - \frac{M_{t+1}-M_t}{p_t}\right)$$

This simplifies to:

$$c_t + I_t = w_t n_t + R_t u_t k_t + \text{Profit}_t - g_t$$

Total profits must be equal to the sum of profits earned by intermediate good firms:

$$\text{Profit}_t = \int_0^1 \text{Profit}_{j,t} dj$$

Real profits earned by intermediate goods firms are:

$$\text{Profit}_{j,t} = \frac{p_{j,t}}{p_t} y_{j,t} - w_t n_{j,t} - R_t \widehat{k}_{j,t}$$

Substitute in the demand for each intermediate good:

$$\text{Profit}_{j,t} = \left(\frac{p_{j,t}}{p_t}\right)^{1-\varepsilon} y_t - w_t n_{j,t} - R_t \widehat{k}_{j,t}$$

Now plug this into the aggregate profit expression:

$$\text{Profit}_t = \int_0^1 \left(\left(\frac{p_{j,t}}{p_t}\right)^{1-\varepsilon} y_t - w_t n_{j,t} - R_t \widehat{k}_{j,t} \right) dj$$

Distribute the integral:

$$\text{Profit}_t = \int_0^1 \left(\frac{p_{j,t}}{p_t}\right)^{1-\varepsilon} y_t dj - \int_0^1 w_t n_{j,t} dj - \int_0^1 R_t \widehat{k}_{j,t} dj$$

Note that many of these terms can be taken out of the integrals:

$$\text{Profit}_t = y_t \frac{1}{p_t^{1-\varepsilon}} \int_0^1 p_{j,t}^{1-\varepsilon} dj - w_t \int_0^1 n_{j,t} dj - R_t \int_0^1 \widehat{k}_{j,t} dj$$

Now use the fact that (i) the aggregate price level is defined as $p_t^{1-\varepsilon} = \int_0^1 p_{j,t}^{1-\varepsilon} dj$, (ii) aggregate labor demand must equal supply, $\int_0^1 n_{j,t} dj = n_t$, and (iii) the aggregate supply of capital services must equal demand, $\int_0^1 \widehat{k}_{j,t} dj = u_t k_t = \widehat{k}_t$. Plugging these in, we get:

$$\text{Profit}_t = y_t - w_t n_t - R_t u_t k_t \tag{24}$$

Plugging this expression into the household budget constraint yields a standard looking aggregate accounting identity:

$$y_t = c_t + I_t + g_t \tag{25}$$

Now go to the supply side of the economy. The Calvo pricing assumption, while appearing rather ad-hoc and unrealistic, is used so frequently because it facilitates aggregation. Each period, firms have a fixed probability $1 - \phi$ of getting to update their price. Importantly, this probability is independent of past history. Since there are an infinite (i.e. continuum) number of firms, there will be exactly the fraction $1 - \phi$ firms who get to update in any period, and exactly the fraction ϕ firms who do not. Furthermore, since these firms are drawn randomly, looking at the distribution of any subset of firms is the same as looking at the distribution of all firms.

To see this concretely, recall the definition of the aggregate price index:

$$p_t = \left(\int_0^1 p_{j,t}^{1-\varepsilon} dj \right)^{\frac{1}{1-\varepsilon}}$$

Equivalently:

$$p_t^{1-\varepsilon} = \int_0^1 p_{j,t}^{1-\varepsilon} dj \tag{26}$$

The term on the right hand side is *both* the *sum* and the *average* of $p_{j,t}^{1-\varepsilon}$ over j. This is the advantage of defining firms as existing over the unit interval – there is total mass 1 of firms, so the sum and the average are the same thing. The price index is really the average, but defining firms as existing over the unit interval means this is equivalent to the sum.

So as to fix ideas and see this more clearly, let's consider a discrete time version of this, where there are N firms (think of N as being very large). Then the aggregate price index would be:

$$p_t^{1-\varepsilon} = \frac{1}{N}\sum_{j=1}^{N} p_{j,t}^{1-\varepsilon}$$

Under the assumptions of the model, at time t, the fraction $1-\phi$ of firms will adjust their price to the same level, $p_t^{\#}$, while the other fraction ϕ firms will be stuck with their previous price. This means that the price level in any period t is:

$$p_t^{1-\varepsilon} = \frac{1}{N}\sum_{j=1}^{(1-\phi)N} p_t^{\#1-\varepsilon} + \frac{1}{N}\sum_{(1-\phi)N+1}^{N} p_{j,t-1}^{1-\varepsilon}$$

This relies on the assumption that N is sufficiently large so that, when each firm has probability $(1-\phi)$ of getting to adjust, there are actually the fraction $(1-\phi)$ of total firms who get to adjust (this would only be the case in expectation for N finite). Since $p_t^{\#1-\varepsilon}$ does not depend on j, we can pull it out of the sum to get:

$$p_t^{1-\varepsilon} = (1-\phi)p_t^{\#1-\varepsilon} + \frac{1}{N}\sum_{(1-\phi)N+1}^{N} p_{j,t-1}^{1-\varepsilon}$$

Multiply and divide the second component on the right hand side by $(\phi N - 1)$ (which is the total number of firms in that sum). We can re-write this as:

$$p_t^{1-\varepsilon} = (1-\phi)p_t^{\#1-\varepsilon} + \frac{1}{N}(\phi N - 1)\frac{1}{\phi N - 1}\sum_{(1-\phi)N+1}^{N} p_{j,t-1}^{1-\varepsilon}$$

Note that $\frac{1}{\phi N-1}\sum_{(1-\phi)N+1}^{N} p_{j,t-1}^{1-\varepsilon}$ is the average price of non-updating price-setters. Provided that N is sufficiently large, and, importantly, assuming that the non-updating price-setters are drawn randomly from the distribution of firms, it must be the case that the average of this subset must be equal to the population average:

$$\frac{1}{\phi N - 1}\sum_{(1-\phi)N+1}^{N} p_{j,t-1}^{1-\varepsilon} = \frac{1}{N}\sum_{j=1}^{N} p_{j,t-1}^{1-\varepsilon}$$

This is pretty intuitive. Suppose that you have a distribution of firm prices. What this is saying is that if you draw a large enough random sample from that distribution, then the mean of that sample is equal to the population mean. In the limiting case (which is what we get with the integrals and the continuum of firms), $N = \infty$, so this is valid. Using the definition of the aggregate price level, we can write this as:

$$\frac{1}{\phi N - 1} \sum_{(1-\phi)N+1}^{N} p_{j,t-1}^{1-\varepsilon} = p_{t-1}^{1-\varepsilon}$$

Now going back, we have:

$$p_t^{1-\varepsilon} = (1-\phi) p_t^{\#1-\varepsilon} + \frac{1}{N}(\phi N - 1) p_{t-1}^{1-\varepsilon}$$

The limit of $\frac{\phi N - 1}{N}$ as $N \to \infty$ is ϕ. So we are left with:

$$p_t^{1-\varepsilon} = (1-\phi) p_t^{\#1-\varepsilon} + \phi p_{t-1}^{1-\varepsilon} \tag{27}$$

This means that the current price level can be written as a convex combination of the reset price and the previous price level, just as assumed earlier.

Define gross inflation as:

$$\Pi_t = \frac{p_t}{p_{t-1}} = 1 + \pi_t \tag{28}$$

Similarly, define gross reset price inflation as:

$$\Pi_t^{\#} = \frac{p_t^{\#}}{p_{t-1}} = 1 + \pi_t^{\#} \tag{29}$$

Divide both sides of the equation for the aggregate price level by $p_{t-1}^{1-\varepsilon}$ to get:

$$\Pi_t^{1-\varepsilon} = (1-\phi)\Pi_t^{\#1-\varepsilon} + \phi \tag{30}$$

Now let's aggregate the production function. The demand for intermediate good j is given by:

$$y_{j,t} = \left(\frac{p_{j,t}}{p_t}\right)^{-\varepsilon} y_t$$

Plug in the production function for the intermediate good j, and take note of the fact that common factor markets imply that capital and labor will be hired in the same proportion, which will in turn be equal to the aggregate capital to labor ratio:

$$a_t \left(\frac{\widehat{k}_t}{n_t}\right)^{\alpha} n_{j,t} = \left(\frac{p_{j,t}}{p_t}\right)^{-\varepsilon} y_t$$

Now sum these up across the intermediate goods firms:

$$\int_0^1 a_t \left(\frac{\widehat{k}_t}{n_t}\right)^{\alpha} n_{j,t}dj = \int_0^1 \left(\frac{p_{j,t}}{p_t}\right)^{-\varepsilon} y_t dj$$

Use the fact that much of this does not depend upon j to simplify:

$$a_t \left(\frac{\widehat{k}_t}{n_t}\right)^{\alpha} \int_0^1 n_{j,t}dj = y_t \int_0^1 \left(\frac{p_{j,t}}{p_t}\right)^{-\varepsilon}$$

Now use the fact the aggregate labor demand must equal aggregate labor supply (e.g. $\int_0^1 n_{j,t}dj = n_t$) and simplify to get:

$$y_t = \frac{a_t \widehat{k}_t^{\alpha} n_t^{1-\alpha}}{v_t} \tag{31}$$

Where:

$$v_t = \int_0^1 \left(\frac{p_{j,t}}{p_t}\right)^{-\varepsilon} \tag{32}$$

v_t essentially measures a distortion introduced by dispersion in relative prices. As we have seen, there is already a distortion associated with the monopoly power of firms (so that there is less capital and labor in steady state than there would be without the monopoly power). This shows that there is an additional distortion associated with relative price fluctuations owing to price stickiness. In a flexible price model, we saw that all firms would choose the same price. Hence $v_t = 1$ at all times. With sticky prices, relative prices will fluctuate, and $v_t \neq 1$ in general. Nevertheless, to a first order approximation (about a zero inflation steady state), this term disappears.

This introduces another complication in terms of aggregation, since we again have j subscripts floating around. It turns out, however, that we can again appeal to Calvo pricing to write v_t in terms of only aggregate things.

$$\int_0^1 \left(\frac{p_{j,t}}{p_t}\right)^{-\varepsilon} = p_t^{\varepsilon} \int_0^1 p_{j,t}^{-\varepsilon}$$
$$\int_0^1 p_{j,t}^{-\varepsilon} = \int_0^{1-\phi} p_t^{\#-\varepsilon} + \int_{1-\phi}^1 p_{j,t-1}^{-\varepsilon}$$
$$\int_0^1 p_{j,t}^{-\varepsilon} = (1-\phi) p_t^{\#-\varepsilon} + \int_{1-\phi}^1 p_{j,t-1}^{-\varepsilon}$$

Now go back to the expression:

$$v_t = p_t^{\varepsilon} \int_0^1 p_{j,t}^{-\varepsilon} = (1-\phi)\left(\frac{p_t^{\#}}{p_t}\right)^{-\varepsilon} + p_t^{\varepsilon} \int_{1-\phi}^1 p_{j,t-1}^{-\varepsilon}$$

$$v_t = p_t^{\varepsilon} \int_0^1 p_{j,t}^{-\varepsilon} = (1-\phi)\left(\frac{p_t^{\#}}{p_{t-1}}\right)^{-\varepsilon}\left(\frac{p_{t-1}}{p_t}\right)^{-\varepsilon} + p_t^{\varepsilon} \int_{1-\phi}^1 p_{j,t-1}^{-\varepsilon}$$

$$v_t = (1-\phi)\Pi_t^{\#-\varepsilon}\Pi_t^{\varepsilon} + p_{t-1}^{-\varepsilon} p_t^{\varepsilon} \int_{1-\phi}^1 \left(\frac{p_{j,t-1}}{p_{t-1}}\right)^{-\varepsilon} dj$$

Again using the fact that we're randomly sampling from the entire distribution of firm prices, the last term simplifies to:

$$v_t = (1-\phi)\Pi_t^{\#-\varepsilon}\Pi_t^{\varepsilon} + \phi\Pi_t^{\varepsilon} v_{t-1} \tag{33}$$

We will want to redefine the expression for optimal reset price derived above in terms of inflation since we will in principle allow for non-zero steady state inflation. Recall that:

$$\frac{p_t^{\#}}{p_t} = \left(\frac{\varepsilon}{\varepsilon-1}\right) E_t \frac{\widehat{A}_t}{\widehat{D}_t}$$

Let's re-write this in terms of Π_t^* by multiplying and dividing the left hand side by p_{t-1} to get:

$$\Pi_t^{\#} = \Pi_t \left(\frac{\varepsilon}{\varepsilon-1}\right) E_t \frac{\widehat{A}_t}{\widehat{D}_t} \tag{34}$$

Now let's re-write the auxiliary variables $\widehat{A}_t$ and $\widehat{B}_t$ in terms of stationary variables:

$$\widehat{A}_t = \lambda_t y_t mc_t + \phi\beta E_t \Pi_{t+1}^{\varepsilon} \widehat{A}_{t+1} \tag{35}$$

$$\widehat{D}_t = \lambda_t y_t + \phi\beta E_t \Pi_{t+1}^{\varepsilon-1} \widehat{D}_{t+1} \tag{36}$$

7 Equilibrium Conditions

The entire set of equilibrium conditions is given by:

$$y_t = c_t + I_t + g_t \tag{37}$$

$$y_t = \frac{a_t \widehat{k}_t^{\alpha} n_t^{1-\alpha}}{v_t} \tag{38}$$

$$\widehat{k}_t = u_t k_t \tag{39}$$

$$\Pi_t = \left((1-\phi)\Pi_t^{\#1-\varepsilon} + \phi\right)^{\frac{1}{1-\varepsilon}} \tag{40}$$

$$v_t = (1-\phi)\Pi_t^{\#-\varepsilon}\Pi_t^{\varepsilon} + \phi\Pi_t^{\varepsilon} v_{t-1} \tag{41}$$

$$\Pi_t^{\#} = \Pi_t \left(\frac{\varepsilon}{\varepsilon - 1}\right) E_t \frac{\widehat{A}_t}{\widehat{D}_t} \tag{42}$$

$$\widehat{A}_t = \lambda_t y_t mc_t + \phi\beta E_t \Pi_{t+1}^{\varepsilon} \widehat{A}_{t+1} \tag{43}$$

$$\widehat{D}_t = \lambda_t y_t + \phi\beta E_t \Pi_{t+1}^{\varepsilon-1} \widehat{D}_{t+1} \tag{44}$$

$$w_t = mc_t(1-\alpha)\left(\frac{\widehat{k}_t}{n_t}\right)^{\alpha} \tag{45}$$

$$R_t = mc_t \alpha \left(\frac{\widehat{k}_t}{n_t}\right)^{\alpha-1} \tag{46}$$

$$\frac{1}{c_t - \gamma c_{t-1}} - E_t \frac{\beta\gamma}{c_{t+1} - \gamma c_t} = \lambda_t \tag{47}$$

$$\theta_t (1-n_t)^{-\xi} = \lambda_t w_t \tag{48}$$

$$\lambda_t R_t = \mu_t \Delta \delta_0 u_t^{\Delta - 1} \tag{49}$$

$$\lambda_t = \beta E_t \lambda_{t+1}(1+i_{t+1})\Pi_{t+1}^{-1} \tag{50}$$

$$\mu_t = \beta E_t \left(\lambda_{t+1} R_{t+1} u_{t+1} + \mu_{t+1}(1 - \delta_0 u_{t+1}^{\Delta})\right) \tag{51}$$

$$m_t^{-\nu} = \lambda_t - \beta E_t \lambda_{t+1} \Pi_{t+1}^{-1} \tag{52}$$

$$\lambda_t = \mu_t \left(1 - \frac{\tau}{2}\left(\frac{I_t}{I_{t-1}} - 1\right)^2 - \tau\left(\frac{I_t}{I_{t-1}} - 1\right)\frac{I_t}{I_{t-1}}\right) + \beta E_t \mu_{t+1} \tau \left(\frac{I_{t+1}}{I_t} - 1\right)\left(\frac{I_{t+1}}{I_t}\right)^2 \tag{53}$$

$$k_{t+1} = \left(1 - \frac{\tau}{2}\left(\frac{I_t}{I_{t-1}} - 1\right)^2\right) I_t + \left(1 - \delta_0 u_t^{\Delta}\right) k_t \tag{54}$$

$$q_t = \frac{\mu_t}{\lambda_t} \tag{55}$$

$$\ln a_t = \rho_a \ln a_{t-1} + \varepsilon_{a,t} \tag{56}$$

$$\ln \theta_t = (1-\rho_\theta)\ln\theta^* + \rho_\theta \ln\theta_{t-1} + \varepsilon_{\theta,t} \tag{57}$$

$$\ln g_t = (1-\rho_g)(\ln\omega + \ln y^*) + \rho_g \ln g_{t-1} + \varepsilon_{g,t} \tag{58}$$

$$i_{t+1} = \rho_i i_t + (1-\rho_i)\phi_\pi(\pi_t - \pi^*) + (1-\rho_i)\phi_y\left(\frac{y_t}{y_{t-1}} - 1\right) + \varepsilon_{i,t} \tag{59}$$

This is a system of 23 equations in 23 variables. The 23 variables are stacked into the vector X_t as follows:

$$X_t = \left[y_t \ c_t \ I_t \ g_t \ a_t \ \widehat{k}_t \ n_t \ u_t \ k_t \ v_t \ \Pi_t \ \Pi_t^{\#} \ \widehat{A}_t \ \widehat{D}_t \ mc_t \ w_t \ R_t \ \lambda_t \ \mu_t \ \theta_t \ m_t \ i_{t+1} \ q_t \right]$$

The 23 equations are as follows. (37) is the aggregate accounting identity. (38) is the aggregate production function. (39) is the definition of capital services. (40) is the definition of aggregate price dispersion. (41) describes the evolution of aggregate inflation. (42) is the equation for reset price inflation. (43)-(44) are the auxiliary terms that show up in the reset price inflation term, written in recursive form. (45) is the labor demand curve, and (46) is the capita demand curve. (47) defines the marginal utility of income. (48) is the labor supply curve. (49) is the condition for optimal capital utilization. (50) is the Euler equation for bonds. (51) is the Euler equation for capital. (52) is the demand for real money balances. (53) is the Euler equation for investment. (54) is the capital accumulation equation. (55) defines q_t as the ratio of the marginal utility of capital to the marginal utility of income, which in turn says how many goods an extra unit of capital is worth. (56) is the exogenous process for technology. (57) is the exogenous process for the preference shifter. (58) is the exogenous process for government spending. (59) is the Taylor rule.

8 The Steady State and Parameter Values

We need to analyze the non-stochastic steady state of the model in order to pin down some parameter values. First, begin, with (53). Since investment will be constant in the steady state, it must be the case that $\mu^* = \lambda^*$, which implies that $q^* = 1$. Now go to (49), the first order condition for utilization. Using this fact about the steady state values of the Lagrange multipliers, we can find the steady state rental rate for capital services:

$$R^* = \Delta \delta_0 u^{*\Delta - 1}$$

We will normalize $u^* = 1$ (which in turn means that $\widehat{k}^* = k^*$). This implies that $R^* = \Delta\delta_0$. Now go to (51), the first order condition for capital. Using the above facts, we can find another expression for the steady state rental rate:

$$R^* = \frac{1}{\beta} - (1 - \delta_0) \tag{60}$$

These two equations imply a parametric restriction of Δ:

$$\Delta = \frac{\frac{1}{\beta} - 1}{\delta_0} + 1 \tag{61}$$

This is exactly the same restriction as we had earlier, and for plausible values ends up mean that $\Delta = 1.41$.

From the Taylor rule, the only way for the nominal interest rate to be constant is for $\pi_t = \pi^*$. In other words, in the steady state, actual inflation must equal the central bank's target for inflation.

Now go to (43) and (44) and let's solve for steady state values of the auxiliary variables, $\widehat{A}$ and $\widehat{D}$:

$$\widehat{A}^* = \frac{\lambda^* y^* mc^*}{1 - \phi\beta(1+\pi^*)^{\varepsilon}}$$
$$\widehat{B}^* = \frac{\lambda^* y^*}{1 - \phi\beta(1+\pi^*)^{\varepsilon-1}}$$

Note that for these expressions to make sense, we must have $\phi\beta(1+\pi^*)^{\varepsilon} < 1$. We will maintain this assumption. Now plug this into (42):

$$1 + \pi^{\#*} = (1+\pi^*)\frac{\varepsilon}{\varepsilon - 1} mc^* \frac{1 - \phi\beta(1+\pi^*)^{\varepsilon-1}}{1 - \phi\beta(1+\pi^*)^{\varepsilon}}$$

Now we can go to (41) to find an expression for $1 + \pi^{\#*}$:

$$(1+\pi^*)^{1-\varepsilon} = (1-\phi) * (1+\pi^*)^{1-\varepsilon} + \phi$$
$$1 + \pi^{\#*} = \left(\frac{(1+\pi^*)^{1-\varepsilon} - \phi}{1-\phi}\right)^{\frac{1}{1-\varepsilon}}$$

We can then use this to solve for steady state marginal cost:

$$mc^* = \frac{\varepsilon - 1}{\varepsilon}\left(\frac{(1+\pi^*)^{1-\varepsilon} - \phi}{1-\phi}\right)^{\frac{1}{1-\varepsilon}} \frac{1}{1+\pi^*}\frac{1 - \phi\beta(1+\pi^*)^{\varepsilon}}{1 - \phi\beta(1+\pi^*)^{\varepsilon-1}} \tag{62}$$

Now that we have an expression for steady state marginal cost and an expression for the steady state rental rate on capital, we can find an expression for the steady state capital to labor ratio from (46). We assume a value of α that is standard based on labor's share (i.e. $\alpha = \frac{1}{3}$:

$$\frac{k^*}{n^*} = \left(\frac{\alpha mc^*}{R^*}\right)^{\frac{1}{1-\alpha}} \tag{63}$$

Given this, we can recover an expression for the steady state wage:

$$w^* = mc^*(1-\alpha)\left(\frac{k^*}{n^*}\right)^{\alpha} \tag{64}$$

Now let's find an expression for steady state v^* from (40):

$$v^* = \frac{(1-\phi)(1+\pi^*)^{\varepsilon}(1+\pi^{\#*})^{-\varepsilon}}{1 - \phi(1+\pi^*)^{\varepsilon}} \tag{65}$$

Go to the expression for the marginal utility of income, simplified and evaluated at steady state:

$$\lambda^* = \frac{1}{c^*}\frac{1-\beta\gamma}{1-\gamma} \tag{66}$$

Use this to simplify the household's first order condition for labor supply:

$$\theta^*(1-n^*)^{-\xi} = \lambda^* w^* = \frac{1}{c^*}\frac{1-\beta\gamma}{1-\gamma} mc^*(1-\alpha)\left(\frac{k^*}{n^*}\right)^{\alpha}$$

Multiply both sides by n^* to find an expression for $\frac{c^*}{n^*}$:

$$\theta^* n^*(1-n^*)^{-\xi} = \lambda^* w^* = \frac{n^*}{c^*}\frac{1-\beta\gamma}{1-\gamma} mc^*(1-\alpha)\left(\frac{k^*}{n^*}\right)^{\alpha}$$

Now, from the exogenous expression for government spending and the capital accumulation equation evaluate in steady state, we can write the aggregate accounting identity as:

$$y^* = \frac{c^* + \delta_0 k^*}{1-\omega}$$

Divide both sides by n^*:

$$\frac{y^*}{n^*} = \frac{\frac{c^*}{n^*} + \delta_0 \frac{k^*}{n^*}}{1-\omega}$$

Using properties of the aggregate production function, we can simplify the left hand side:

$$\frac{1}{v^*}\left(\frac{k^*}{n^*}\right)^{\alpha} = \frac{\frac{c^*}{n^*} + \delta_0 \frac{k^*}{n^*}}{1-\omega}$$

Use this to solve for the consumption to employment ratio in steady state:

$$\frac{c^*}{n^*} = \frac{(1-\omega)}{v^*}\left(\frac{k^*}{n^*}\right)^{\alpha} - \delta_0 \frac{k^*}{n^*}$$

We can pick $\omega = 0.2$ to match the average share of government spending in output. Given this parameter, we now have a value for the consumption to employment ratio. Given this, and given a target of $n^* = \frac{1}{3}$ to match the average behavior of hours worked, we can solve for a value of θ^* in terms of things which are known and the parameter ξ, which governs the labor supply elasticity and γ, which governs the degree of internal habit formation:

$$\theta^* = \frac{1}{n^*(1-n^*)^{-\xi}}\frac{n^*}{c^*}\frac{1-\beta\gamma}{1-\gamma} mc^*(1-\alpha)\left(\frac{k^*}{n^*}\right)^{\alpha} \tag{67}$$

Then given n^*, we also have y^* from the production side of the economy:

$$y^* = \frac{1}{v^*}\left(\frac{k^*}{n^*}\right)^{\alpha} n^* \tag{68}$$

Given n^*, we can also easily retrieve the steady state values of capital and consumption from the expressions for the ratios of those variables to hours. Finally, we can get the steady state nominal interest rate from (50):

$$1 + i^* = \frac{1}{\beta}(1+\pi^*) \tag{69}$$

From (52), we can then solve for steady state real money balances (I will set $\nu = 1$ so that preferences over money balances are log):

$$m^* = \left(\lambda^* \left(1 - \frac{\beta}{1+\pi^*}\right)\right)^{-\frac{1}{\nu}} \tag{70}$$

To get numerical values for the steady state, I need to specify some numerical values for some of the parameters. Other parameters do not influence the steady state. I set $\alpha = \frac{1}{3}$ and $\delta 0 = 0.025$. I set $\beta = 0.99$, and $\pi^* = 0.005$. This is a quarterly calibration. This implies that the average annualized inflation rate is about 2 percent, while the average annualized nominal interest rate is about 6 percent (inflation and interest rates are always quoted at annual rates), both of which are in line with US data. These values imply that $\Delta = 1.4040$. I set $\gamma = 0.75$. This parameter plays only a very minor role in the steady state but I need a value to compute it. I set $\phi = 0.8$; this says that firms change their prices once every 5 quarters on average. I set $\varepsilon = 6$, implying steady state markups of around 20 percent. I set $\omega = 0.2$, which is the average share of government spending in output. Finally, I set $\xi = 1$, so that preferences over leisure are log. These parameters imply a value of $\theta^* = 1.9358$ to hit steady state hours of 0.33.

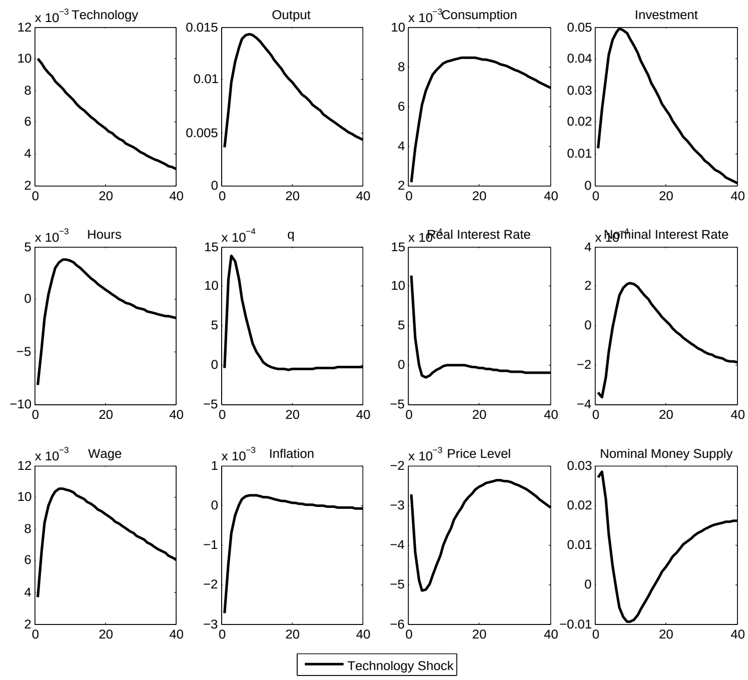
Given this set of parameters, I get the following steady state values of the variables of the model:

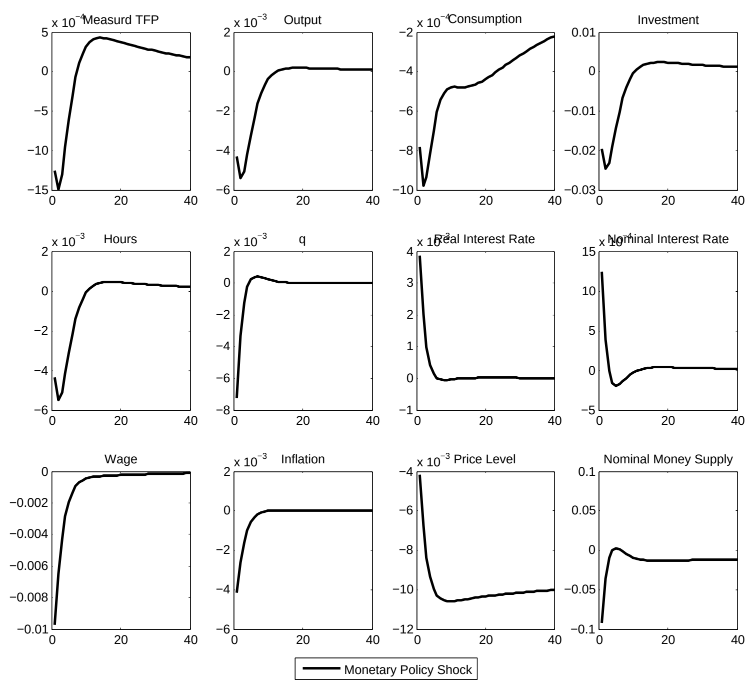
y	0.9077	v	1.0018	λ	1.8789
c	0.5482	π	0.005	μ	1.8789
I	0.1780	$\pi^{\#}$	0.0266	θ	1.9358
g	0.1815	$\widehat{A}$	7.7230	m	35.6582
a	1	$\widehat{D}$	9.0723	i	0.0152
$\widehat{k}$	7.1206	mc	0.8329	q	1
n	0.33	w	1.5377		
u	1	R	0.0351		

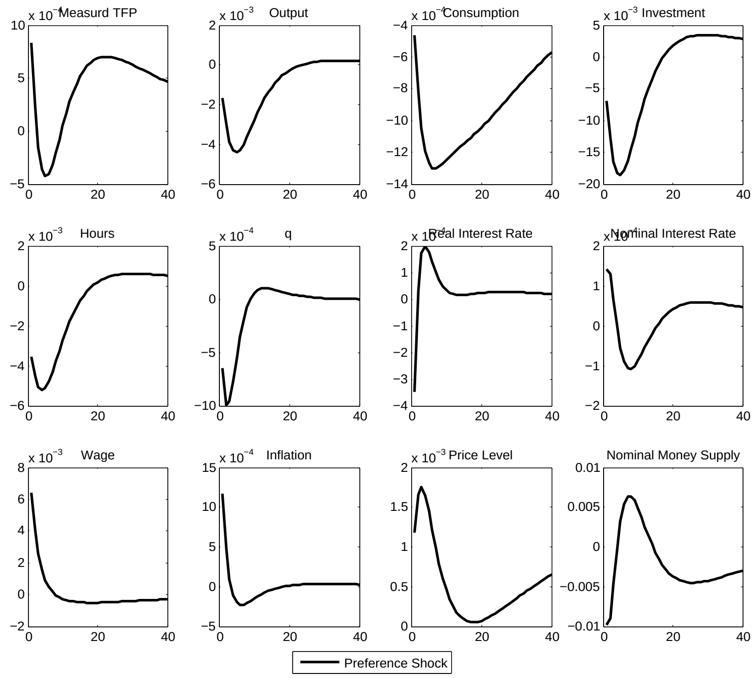
There are other parameters for which I need values. I set $\tau = 0.5$. I set $\rho_a = 0.97$ and $\sigma_{\varepsilon_a} = 0.01$. I set $\rho_\theta = \rho_g = \rho_i = 0.85$ and $\sigma_{\varepsilon_i} = \sigma_{\varepsilon_\theta} = \sigma_{\varepsilon_g} = 0.01$. I set $\phi_\pi = 1.5$ and $\phi_y = 0.5$.

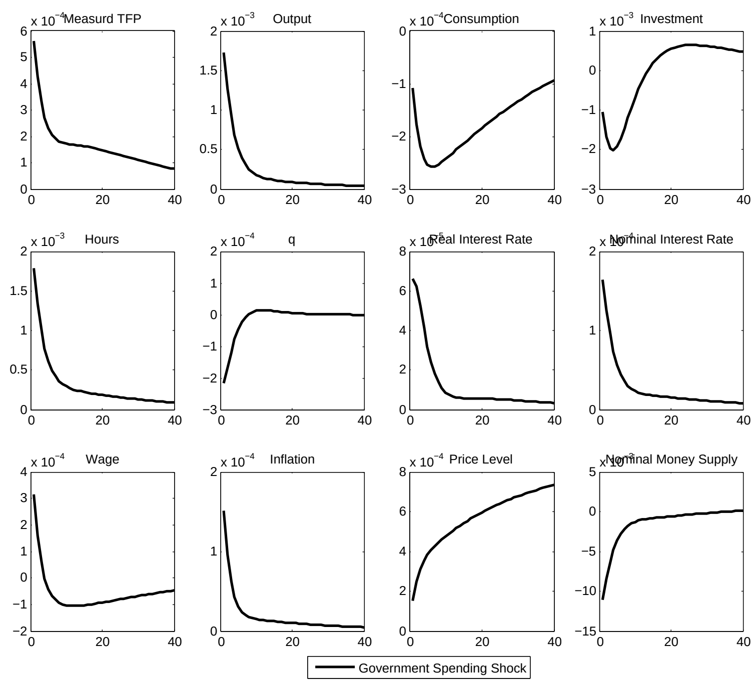
9 Impulse Responses

Below are impulse responses to each of the shocks. Discussions of each will follow.


Technology
Output
Consumption
Investment
Hours
q
Real Interest Rate
Nominal Interest Rate
Wage
Inflation
Price Level
Nominal Money Supply
Technology Shock


Measurd TFP
Output
Consumption
Investment
Hours
q
Real Interest Rate
Nominal Interest Rate
Wage
Inflation
Price Level
Nominal Money Supply
Monetary Policy Shock





In response to the technology shock, output, consumption, and investment all rise immediately and then follow hump-shaped responses. Hours actually fall immediately impact. Inflation falls. The fall in hours does not happen in response to this shock in a standard RBC model. Both the sticky prices (and the imperfect monetary policy) and the investment adjustment costs contribute to this hours decline. The price of capital does little on impact but then rises substantially.

Next consider the impulse responses to the monetary policy shock. We immediately observe that output, hours, investment, consumption, and hours all immediately decline in response to a "tightening" and follow hump-shaped patterns thereafter. As we would expect, the money supply contracts and inflation and the price level fall. Both real and nominal interest rates rise. A final thing to note is that measured TFP – failing to account for the endogenous response of capital utilization – falls with output. These responses are fairly similar to what obtains from estimating VARs to identify monetary policy shocks. Indeed, Christiano, Eichebnaum, and Evans (2005) hailed a model somewhat similar to this as a success because it could match the impulse responses from the data.

Next, consider the preference shock. An increase in θ means people like to work less. The reduction in labor supply leads to a reduction in output. Consumption falls, although not as much as the fall in output; hence investment falls. The price of capital declines, consistent with the observed decline in investment. Inflation and the price level rise; the money supply contracts.

Finally, consider the responses to the government spending shock. The mechanism through which the spending shock stimulates output is through a negative wealth effect – hence output and hours rise, while consumption falls. For this specification of the persistence of government spending, we see that investment declines on impact. Hence q falls. We see that inflation and the price level rise.

The model seems to fit the data pretty well in an unconditional sense as well. The relative volatilities are pretty close to those in the data (though again hours are not quite volatile enough relative to the data). The contemporaneous correlations among variables are pretty close to those in the data – the real interest rate is mildly countercyclical here (as in the data), inflation is mildly procyclical (as it is in the data), q is procyclical (as it is in the data), and money is procyclical (as it is in the data).

10 Estimating Free Parameters

As noted immediately above, the "fit" of the model is pretty good. This suggests that the model may be a fairly accurate description of the actual economy, which means that one may be able use the model to inform policy advice.

This fit is pretty good in spite of the fact that I haven't really picked the parameters to maximize "fit". Picking parameters to make the model fit the data as well as possible is the essence of econometric estimation. Just as one can estimate the parameters of a linear regression, one can (with some effort) estimate the parameters of a fairly complicated model such as this. Here I very briefly describe how one might go about doing this. I consider four different ways of estimating the model: (1) maximum likelihood, (2) method of moments, (3) simulated method of moments, and (4) indirect inference.

Before proceeding it is helpful to revisit our terminology on linearized DSGE models. Let s_t be a $m \times 1$ vector of state variables, and let y_t be a $n \times 1$ vector of control variables. Let x_t be a

$q \times 1$ vector of observable variables (which could in principle be either states or controls, though it is typical in this literature to assume that states are not observed). The state space representation of the model can be written:

$$
\begin{aligned}
s_t &= As_{t-1} + B\varepsilon_t \\
y_t &= \Phi s_t \\
x_t &= Cs_t
\end{aligned}
$$

Here ε_t is a $r \times 1$ vector of shocks, with $r \leq m$. Hence B is $m \times r$, and A is $m \times m$. Φ is $n \times m$. C is $q \times m$, and its values can easily be taken from Φ (or are 1s if some part of the state are observed).

It is typical to calibrate a subset of the model's parameters based on long run moments. Parameters that are usually calibrated include β, δ_0, α, etc.. Let Θ be a $v \times 1$ vector of the remaining parameters of the model (e.g. the Calvo parameter, the habit formation parameter, the parameter governing the adjustment cost function). The objective is to pick Θ to maximize some measure of goodness of fit of the model (given above by the state space representation) to the data.

10.1 Maximum Likelihood

Assume that the shocks of the model are drawn from some distribution (almost always normal, so I will go with that). At time t, the density of x_t conditional on available information is:

$$
f(x_t \mid x_{t-1}, \Theta) = N(C\widehat{s}_{t|t-1}, C\Sigma_{t|t-1}C')
$$

Here $\widehat{s}_{t|t-1}$ denotes your time $t-1$ expected value of the state in time t. There is uncertainty over this for two reasons – first because of shocks realized at time t, but also because you may not observe the full state. $\Sigma_{t|t-1}$ is the mean square error of the forecast of the state. This is given by:

$$
\Sigma_{t|t-1} = E\left((s_t - \widehat{s}_{t|t-1})(s_t - \widehat{s}_{t|t-1})'\right)
$$

The likelihood function is:

$$
\mathcal{L}(\Theta) = -(qT/2)\ln(2\pi) - (1/2)\sum_{t=1}^{T} \ln \mid \Omega_t \mid -(1/2)\sum_{t=1}^{T} u_t'\Omega_t^{-1}u_t
$$

Where (recall that q is the number of observable variables):

$$
\begin{aligned}
u_t &= x_t - C\widehat{s}_{t|t-1} \\
\widehat{s}_{t|t-1} &= A\widehat{s}_{t-1|t-1} \\
\Omega_t = E(u_t u_t') &= C\Sigma_{t|t-1}C'
\end{aligned}
$$

The parameter estimates, $\widehat{\Theta}$, are those that maximize the likelihood function. One would typically use the Kalman filter to recursively form estimates of $\widehat{s}_{t-1|t-1}$, which is necessary to construct the likelihood. The Kalman filter is a (relatively) simple method for forecasting an unobserved state given observables and some knowledge of the evolution of the state.

There is an issue in terms of how many variables can be used as observables. In particular, it must be the case that $q \leq r$ – i.e. the number of observables must be less than or equal to the number of shocks. The reason is relatively straightforward – if $q > r$, then Ω_t will be singular, and Ω_t^{-1} will not exist. Recall that Ω_t is just the variance-covariance matrix of innovations (time t forecast errors); these innovations are just linear combinations of structural shocks. If you have more innovations then shocks then some innovations will be exact linear combinations of the others. This is a generic problem that manifests itself in these models, and goes by the term *stochastic singularity*. The basic idea is that there are *lots* of variables in which we are interested as economists but our models typically only contain a small number of shocks. If you want to use more observables to estimate the model, you can either (i) add more shocks or (ii) add measurement errors to the observed variables.

A modified form of estimation by maximum likelihood is Bayesian maximum likelihood. Bayesian estimation essentially imposes prior distributions on the parameters. For example, you might say that the prior distribution of the Calvo parameter is a gamma distribution with mean 0.75 and standard deviation 0.1. Bayesian estimation can essentially be thought of as maximizing the likelihood of the model plus a penalty, where the penalty is based on the strength of the prior (i.e. the standard deviation of the prior distribution). If the strength of the prior is very strong, then the estimates will come out near the mean of the prior distribution (essentially by construction). A "diffuse" prior is one that places equal probability on each possible realization (basically there is no strength of the prior). If the prior distributions are all diffuse than Bayesian MLE is equivalent to MLE.

I am no expert on Bayesian statistics, so take what follows with a grain of salt. Frequentists believe that there is a data generating process with "true" (i.e. fixed) parameters governing it, and our objective is to come up with estimates (and confidence intervals) of those parameters. The frequentist treats the observed sample as random (i.e. observed with noise). The Bayesian essentially flips this around. The Bayesian says there is no "true" value of a parameter, only distributions of that parameter. The Bayesian treats the observed sample as "true" (i.e. not random), and uses that sample to update the (random) distribution of the underlying parameters. In the example above, if your prior distribution of the Calvo parameter had mean 0.75 but the data really prefer 0.5, the posterior distribution of this parameter will "move" closer to 0.5 – the extent to which it moves closer depends on (i) how much the data really like 0.5 and (ii) how strong your prior was.

So basically the frequentist treats parameters as true and samples as random, whereas the Bayesian treats parameters as random and samples as true. An important implication of this is that frequentists produce point estimates of parameters, whereas Bayesian produce "posterior distributions" of parameters. The closes thing to a point estimate is the mode/mean/median of the posterior distribution. These will be the same thing if the shocks are normally distributed. Why would one want to go the Bayesian route? It is a way of essentially bringing extra information to bear on problems and can therefore help with sticky issues of identification. For example, suppose that the data "like" Calvo parameters of 0.25 and 0.75. If you have outside information that the parameter is much closer to 0.75, you can impose that prior and help get "better" estimates of the parameters. The cost of this is, of course, that if your prior distributions are bad this will affect your posterior distributions.

10.2 Generalized Method of Moments

We can also estimate the parameters of the model via the generalized method of moments (GMM). Let $g(\Theta)$ be (analytical) moments from the model (means, standard deviations, covariances, whatever). In the linearized versions of the models we have seen it is relatively easy to compute analytical (i.e. population) moments. We can compute these same moments in the data; call these empirical moments m^*. Let there be p moments and h parameters; for identification of all parameters we need $p \geq h$. GMM is the solution to the following problem:

$$\min_{\Theta} \quad (g(\Theta) - m^*)' \, W \, (g(\Theta) - m^*)$$

Here W is some $p \times p$ weighting matrix. If it were an identity matrix, the problem would essentially be to pick the parameters to minimize the sum of squared "errors", where the errors are defined as the deviations from the model moments and the data moments. Under certain regularity conditions, this will yield consistent estimates of Θ. But efficiency depends on the choice of W. What is typically used is something like the inverse of the variance-covariance matrix of the moments. Basically, this leads to the objective being to minimize the weighted sum of squared errors, where the weights are bigger the smaller is the variance of the particular moment in question. This essentially just weights the moments that are more precise bigger.

I like the idea of method of moments because I think of estimation in terms of matching moments (i.e. I'm more of a frequentist). It also has the advantage that there is a pretty clear way to see whether the model fits the data well or not – you can eyeball whether the moments from the model at the estimated parameter vector are close to those in the data. With maximum likelihood you (or at least I do) have a difficult time interpreting what exactly the maximized likelihood means. With a method of moments interpretation it is just easier to see whether the model fits the data well or not.

10.3 Simulated Method of Moments

The idea behind simulated method of moments is exactly the same as GMM, except the model moments are based on simulations of the model as opposed to analytical moments. Let $\widehat{g}(\Theta)$ be the vector of moments calculated for a simulation of the model at parameter vector Θ, where there are N data observations in the simulation (there are T observations in the observed data). The objective function is:

$$\min_{\Theta} \quad \left(\widehat{g}(\Theta) - m^*\right)' W \left(\widehat{g}(\Theta) - m^*\right)$$

There is a slight difference in what the optimal weighting matrix is. In particular, the optimal weighting matrix is:

$$W = \left(\Sigma(1 + T/N)\right)^{-1}$$

Here Σ is a consistent estimate of the variance-covariance matrix of the moments of the model. $(1 + T/N)$ is an adjustment factor based on the sample size of simulation. As that sample used in the simulation gets big relative to the sample size of the actual data, the simulated moments are essentially population moments, and the weighting matrix goes to the same as it is in the GMM case.

Why would one use simulated method of moments instead of GMM? You would do this if it were difficult/impossible to calculate analytical moments. In a linearized solution of the model, this usually isn't an issue. But if you solved the model using global solution techniques (like value function iteration), calculation of analytical moments would be impossible.

10.4 Indirect Inference

Indirect inference is essentially the same thing as simulated method of moments, except the empirical targets are not moments in the usual sense. Instead, these targets are "moments" in a loose sense. For example, one could estimate a regression of output on hours. This regression has no structural interpretation; it is reduced form. The OLS coefficient estimate is, however, nevertheless a function of the structural parameters of the model. Another object of interest may be impulse responses estimated from a VAR. This is the approach taken in Christiano, Eichenbaum, and Evans (2005). They estimated a VAR on actual data, then simulated data from their model and estimated the same VAR, and picked the parameter vector to match the responses estimated on the simulated data look as close as possible to those from the actual data. Mechanically the estimation is the same as simulated method of moments.