Notes for Macroeconomics II, EC 607

Christopher L. House
University of Michigan

August 20, 2003

1. Basic Dynamic Optimization.

This is a summary of some basic mathematics for handling constrained optimization problems.[1] In macro, we deal with optimization over time. Sometimes the horizons for dynamic optimization problems are finite while sometimes they are infinite. Dynamic optimization problems come in two forms:

1. Discrete time ($t = 0, 1, 2, ...$)
2. Continuous time ($t \in \mathbb{R}$).

For discrete time problems we will often use simple Lagrangians. Consider the following simple growth problem:

$$\max \sum_{t=0}^{\infty} \beta^t u(c_t)$$

subject to the constraints:

$$k_t^{\alpha} \geq c_t + k_{t+1}$$

This can be dealt with using a standard Lagrangian:

$$\mathfrak{L} = \sum_{t=0}^{\infty} \beta^t u(c_t) + \sum_{t=0}^{\infty} \lambda_t \left[k_t^{\alpha} - c_t - k_{t+1}\right]$$

[1] A lot of this material is adapted from Takayama, *Mathematical Economics*, 2nd ed., Cambridge University Press, 1985. For further development of these ideas, the interested student is encouraged to read Chapters 0 and 1 of Takayama and see the references cited there for further explanations of these idea.

and the first order conditions are:

$$\partial c_t : \beta^t u'(c_t) - \lambda_t = 0$$

$$\partial k_{t+1} : -\lambda_t + \lambda_{t+1} \alpha k_{t+1}^{\alpha-1} = 0$$

Alternatively we could set up the Lagrangian as a current value problem as:

$$\mathfrak{L} = \sum_{t=0}^{\infty} \beta^t \left[u(c_t) + \lambda_t \left(k_t^\alpha - c_t - k_{t+1} \right) \right]$$

and the first order conditions are:

$$\partial c_t : u'(c_t) - \lambda_t = 0$$

$$\partial k_{t+1} : -\lambda_t + \beta \lambda_{t+1} \alpha k_{t+1}^{\alpha-1} = 0$$

These two formulations are identical – the only difference is in the interpretation of the shadow value λ_t. In the first set up, λ_t represents the marginal increase in the objective function $\sum_{t=0}^{T} \beta^t u(c_t)$ which would come about as a result of reducing the constraint in period t. In the second setup, λ_t represents the marginal increase in $u(c_t)$ from the same experiment. You are free to use either in the class.

The discrete time problem we just considered has an analogue in continuous time.

$$\max \int_0^\infty e^{-\rho t} u(c(t))$$

subject to the constraints:

$$\dot{k}(t) = k(t)^\alpha - c(t)$$

Note that k cannot jump (it's a state variable) while c can jump (it's a control variable). To solve these problems we usually set up the Hamiltonian function:

$$H = e^{-\rho t} u(c(t)) + \lambda(t) \left[k(t)^\alpha - c(t) \right]$$

The first order necessary conditions for this type of problem are:

$$\text{(for control variables) } H_c = 0$$

$$\text{(for state variables) } H_k = -\dot{\lambda}(t)$$

and

$$\dot{k}(t) = k(t)^{\alpha} - c(t)$$

As with the discrete time problems there is a "current value" Hamiltonian as well:

$$h = u(c(t)) + \theta(t)\left[k(t)^{\alpha} - c(t)\right]$$

the first order conditions are similar:

$$\text{(for control variables) } h_c = 0$$

$$\text{(for state variables) } h_k = \rho\theta - \dot{\theta}$$

and of course:

$$\dot{k}(t) = k(t)^{\alpha} - c(t)$$

What follows is a more detailed discussion of results from constrained optimization. You will not be tested on this material but you may want to skim it to get a feel for why we use the techniques we use. Of course, this discussion is very informal and the interested reader should look to other sources for a more complete analysis. Simon and Blume, and Takayama are good references.

2. Constrained Optimization:

A function is concave if the convex combination of the images of f lies below the function at the convex combination.

For a problem to be a concave problem, f must be concave and the constraint set must be convex. With Lagrangians, x is chosen to maximize L while λ is chosen to minimize L. We want to get all of the binding constraints attached to multipliers. QSP are also called foc or complementary slackness conditions.

λ's are referred to as Lagrange multipliers (shadow prices, shadow values of the constraints).

2.1. Basic Constrained Optimization:

Consider the following problem (P):

$$\max f(x)$$

subject to:

$$g_i(x) \geq 0 \text{ for } i = 1, ...m$$

where f and the g_i's are real valued (typically differentiable) functions defined on $x \in \mathbb{R}^N$.

Without the constraints (the g's) we would simply take derivatives of f and find the critical values at which $f_x = 0$. The presence of the constraints complicates matters a bit. The main simplifying result is the following theorem usually attributed to Kuhn and Tucker (see Takayama p. 75):

Theorem (Kuhn-Tucker) Let f, $g_1, g_2 ... g_m$ be real valued concave functions defined on $X \subseteq \mathbb{R}^N$, X convex. Assume that there exists $x \in X$ s.t. $g_i(x) > 0$ for $i = 1, ...m$ then x^* achieves a maximum of $f(x)$ subject to $g_i(x) \geq 0$ for $i = 1, ...m$ if and only if there exists a vector $\lambda^* = [\lambda_1^*, \lambda_2^*, ...\lambda_m^*]$ with $\lambda_i^* \geq 0$ such that

$$\mathfrak{L}(x, \lambda^*) \leq \mathfrak{L}(x^*, \lambda^*) \leq \mathfrak{L}(x^*, \lambda) \qquad (SP)$$

(i.e. (x^*, λ^*) is a saddle point) for all $x \in X$ and all $\lambda \geq 0$ where $\mathfrak{L}$ is the function:

$$\mathfrak{L}(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$$

Remarks: Note that the theorem works "both ways" so if there is a maximum (M) then it satisfies the saddle point condition (condition SP) while at the same time if you find a point that satisfies the saddle point condition then you have the solution to the maximization problem P. The λ's are referred to as "Lagrange multipliers" (shadow prices, shadow values of the constraints).

Also the constraints will satisfy :

1. $g_i(x) \geq 0$
2. $\sum_i \lambda_i g_i(x) = 0$
3. because all of the g_i and λ_i must be non-negative this implies that if a constraint holds with an inequality (i.e. $>$ rather than $\geq$) then the Lagrange multiplier associated with that constraint must be zero.
4. The condition that there exists an x with $g_i(x) > 0$ for $i = 1, ...m$ is known as Slater's condition.

So if the functions are concave then $SP \iff M$, Importantly, $SP \Longrightarrow M$ without any conditions on concavity and without Slater's condition. If the functions are differentiable then we have the ability to take derivatives and use calculus to solve the problem (we assume that X is open so that differentiation is unambiguous).

The quasi-saddle point conditions (often called first order conditions, or complementary slackness conditions) are:

(QSP) (Quasi-Saddle-Point Conditions): The QSP conditions are:

1. (recall that $x \in \mathbb{R}^N$)
$$\frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j} = 0 \text{ , for } j = 1...N$$

2. $$g_i(x) \geq 0$$

3. $$\lambda_i g_i(x) = 0$$

Theorem: If $f, g_i\ \forall i$ are differentiable then $SP \Longrightarrow QSP$ and if $f, g_i\ \forall i$ are all concave then $QSP \Longrightarrow SP$.

These two theorems are what underlies the practice of "setting up Lagrangians". Kuhn and Tucker have additional qualifications that can extend the results to allow for non-concave functions f, g. Usually some form of concave objectives or convex constraint sets are required. Their main results concern showing that the solution is a saddle point and showing that the QSP conditions are useful in finding the maximum. (again see Takayama for a detailed discussion).

2.1.1. Variations of the Lagrangians...

Sometimes you will see Lagrangians written differently or variations on the QSP conditions. For instance, sometimes the Lagrangian is written in terms of g_i's that are given as $0 \geq g_i(x)$. Authors that do this write the Lagrangian with negative signs on the multipliers:

$$\mathfrak{L}(x, \lambda) = f(x) - \sum_{i=1}^{m} \lambda_i g_i(x)$$

Obviously this amounts to the same thing.

If a problem has non-negativity constraints on x then we would add these constraints to the g's and set up a Lagrangian like:

$$\mathfrak{L}(x,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{N} \mu_j x_j$$

(j goes from 1 to N because $x \in R^N$). Then the QSP conditions would include:

$$\frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j} + \mu_j = 0$$

and

$$\mu_j x_j = 0$$

for all of the x_j's. Often you will see an alternative set of conditions:

1. $$\frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j} \leq 0 \text{ , for } j = 1...N$$

2. $$x_j \left[\frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j} = 0 \right] \text{ , for } j = 1...N$$

3. $$g_i(x) \geq 0$$

4. $$\lambda_i g_i(x) = 0$$

This amounts to the same thing however as you can see. If $x_j > 0$ then $\mu_j = 0$ and $\frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j} = 0$. Condition (2) above asserts that this is true here as well.

Example: Consider the problem of maximizing the quadratic

$$\max_x \left\{ -\frac{1}{2}(x-\bar{x})^2 \right\}$$

subject to the constraint

$$\phi \geq x$$

where ϕ is some number.

Clearly the solution is to set $x = \bar{x}$ or $x = \phi$ depending on whether ϕ is greater than or less than $\bar{x}$. Let's suppose that $\phi < \bar{x}$.

The Lagrangian is:

$$L = -\frac{1}{2}(x-\bar{x})^2 + \lambda[\phi - x]$$

the QSP conditions are:

$$\begin{aligned} -(x-\bar{x}) - \lambda &= 0 \\ \phi - x &\geq 0 \\ \lambda[\phi - x] &= 0 \end{aligned}$$

Either the constraint binds or it doesn't. Suppose that it doesn't, then $\lambda = 0$ (by the last equation). Then, by the first condition we have $x = \bar{x}$ but this violates the second condition since by assumption $\phi < \bar{x}$. As a result the constraint must bind $\lambda > 0$ and $x = \phi$. What is λ^*? Apparently $\lambda^* = \bar{x} - x > 0$ and increasing as ϕ gets lower and lower.

2.2. Discrete Time Lagrangians:

Often in macro, the problems we will tackle have the following form:

$$\max \sum_{t=0}^{T} \beta^t u(c_t, k_t)$$

subject to:

$$f(k_t) \geq c_t + k_{t+1}$$

with k_0 given and $T \in [1, 2, ...\infty]$ (note we allow for infinite time horizons).

The set up for this problem is exactly what you would expect given the Kuhn-Tucker results:

$$\mathfrak{L} = \sum_{t=0}^{T} \beta^t u\left(c_t, k_t\right) + \sum_{t=0}^{T} \lambda_t \left[f\left(k_t\right) - c_t - k_{t+1}\right]$$

with one λ_t for each constraint. Note that among the first order conditions is:

$$\beta^t u' = \lambda_t$$

which says that the shadow price of the constraint λ_t is $\beta^t \frac{\partial u}{\partial c}$.

Sometimes you will see this set up as a "current value" Lagrangian:

$$\mathfrak{L} = \sum_{t=0}^{T} \beta^t \left[u\left(c_t, k_t\right) + \lambda_t \left(f\left(k_t\right) - c_t - k_{t+1}\right)\right]$$

(so that the β^t operates on both the function to be maximized as well as on the Lagrange multiplier). This isn't the way that a normal Lagrangian would be set up but it is just as good - since you are simply multiplying the λ_t's by a fixed set of positive numbers, we can choose the transformed λ's so that $\beta^t \lambda$ is equal to the original shadow values. That said, the new λ's have a new interpretation. Now the first order condition is:

$$u' = \lambda_t$$

You can see why this is called a "current value" Lagrangian. It's because the shadow prices of consumption are stated in terms of current utility rather than utility discounted to the first period.

2.3. Continuous Time

Consider the problem:

$$\max_{u(t),x(t)} \int_0^\infty e^{-\rho t} f(x(t), u(t)) dt$$

subject to the constraints:

$$\dot{x}\left(t\right) = g(x(t), u(t))$$

Here x is a vector of "state variables" and u is a vector of "control variables". State variables cannot change at time t while the controls can jump (for example, capital cannot jump while investment can).

How would we solve this "normally"? We would form a Lagrangian:

$$\mathfrak{L} = \int_0^\infty e^{-\rho t} f(x(t), u(t)) dt + \int_0^\infty \lambda(t) \left[g(x(t), u(t)) - \dot{x}(t) \right] dt$$

and proceed.

We will not do this directly, rather we will define an auxiliary function H:

$$H(x(t), \lambda(t), t) = \max_{u(t)} \left\{ e^{-\rho t} f(x(t), u(t)) + \lambda(t) g(x(t), u(t)) \right\}$$

If we differentiate this with respect to the control variable then:

$$e^{-\rho t} \frac{\partial f}{\partial u} + \lambda(t) \frac{\partial g}{\partial u} = 0$$

This would give us a corresponding $u^*(\lambda, x, t)$ (note that this requires knowledge of the functions $\lambda(t)$, and $x(t)$). The Lagrangian can then be rewritten as:

$$\mathfrak{L} = \int_0^\infty \left[H(x(t), \lambda(t), t) - \lambda(t) \dot{x}(t) \right] dt$$

(we don't have a time path for x or for the shadow value of λ ... if we knew these we could simply find u^*). Next we will use a useful result from the calculus of variations:

Consider choosing $x(t)$ to maximize the sum (here x could be a vector):

$$\int M(x, \dot{x}, t) dt$$

The calculus of variations says that the optimal $x(t)$ must satisfy:[2]

$$\frac{\partial M}{\partial x(t)} - \frac{d\left[\frac{\partial M}{\partial \dot{x}}\right]}{dt} = 0$$

For our problem

$$M(x, \dot{x}, t) = H(x(t), \lambda(t), t) - \lambda(t) \dot{x}(t)$$

and the optimality condition says that for the variable x:

$$H_x + \frac{d[\lambda(t)]}{dt} = H_x + \dot{\lambda}(t) = 0$$

[2] The origins for this result are shown below in the section on the Calculus of Variations.

and the same with $\lambda(t)$:

$$H_\lambda - \dot{x} - \frac{d\,[0]}{dt} = H_\lambda - \dot{x} = 0$$

What is H_λ? Recall:

$$\begin{aligned} H\left(x(t), \lambda(t), t\right) &= \max_{u(t)} \left\{ e^{-\rho t} f(x(t), u(t)) + \lambda(t) g(x(t), u(t)) \right\} \\ &= e^{-\rho t} f(x, u^*\left(\lambda, x, t\right)) + \lambda g(x, u^*\left(\lambda, x, t\right)) \end{aligned}$$

where I have now dropped the time subscripts. Differentiate this with respect to λ and use the envelope theorem to get:

$$H_\lambda = g(x, u^*\left(\lambda, x, t\right))$$

Thus we have:

$$\begin{aligned} H_\lambda &= \dot{x} \\ H_\lambda &= g(x, u^*\left(\lambda, x, t\right)) = \dot{x} \end{aligned}$$

Which really just says that the constraint will bind.

2.3.1. Summary So Far:

We began with the Lagrangian:

$$\mathfrak{L} = \int_0^\infty e^{-\rho t} f(x(t), u(t)) dt + \int_0^\infty \lambda(t) \left[g(x(t), u(t)) - \dot{x}\left(t\right) \right] dt$$

defined the auxiliary function H:

$$H\left(x(t), \lambda(t), t\right) = \max_{u(t)} \left\{ e^{-\rho t} f(x(t), u(t)) + \lambda(t) g(x(t), u(t)) \right\}$$

and saw that the solution satisfies:

1.
$$\frac{\partial \left\{ e^{-\rho t} f(x(t), u(t)) + \lambda(t) g(x(t), u(t)) \right\}}{\partial u} = 0$$

   or:

$$e^{-\rho t} \frac{\partial f(t)}{\partial u} + \lambda(t) \frac{\partial g(t)}{\partial u} = 0$$

2.

$$-H_x = \dot{\lambda}(t)$$

3.

$$\dot{x} = g(x, u^* (\lambda, x, t))$$

(we also have the boundary condition - a transversality condition):

4.

$$\lim_{t \to \infty} \lambda (t) \, x(t) = 0$$

2.3.2. Recipe

A convenient way to proceed is to take the problem as given:

$$\max_{u(t), x(t)} \int_0^{\infty} e^{-\rho t} f(x(t), u(t)) dt$$

subject to the constraints:

$$\dot{x} \, (t) = g(x(t), u(t))$$

and immediately define the Hamiltonian:

$$H = e^{-\rho t} f(x(t), u(t)) + \lambda(t) g((x(t), u(t))$$

now the solution will satisfy:

$$H_u = 0$$

$$-H_x = \dot{\lambda}$$

and

$$\dot{x} \, (t) = g(x(t), u(t))$$

2.3.3. Current Value vs. Present Value:

As with the Lagrangians, it is often useful to convert this into a "current value" Hamiltonian by re-normalizing the Lagrange multipliers. Let :

$$\theta(t) = e^{\rho t} \lambda \, (t)$$

then with H defined as:

$$H = e^{-\rho t} f(x(t), u(t)) + e^{-\rho t} \theta(t) g((x(t), u(t))$$

The solution satisfied:

$$H_u = 0$$
$$-H_x = \dot{\lambda}$$

and

$$\dot{x}\,(t) = g(x(t), u(t))$$

which here requires only a change for $\dot{\lambda}$:

$$\begin{aligned} \lambda &= e^{-\rho t}\theta \\ \dot{\lambda} &= -\rho e^{-\rho t}\theta + e^{-\rho t}\dot{\theta} \end{aligned}$$

so that $H_x = -\rho e^{-\rho t}\theta + e^{-\rho t}\dot{\theta}$.

Then we could alternatively use the current value Hamiltonian:

$$h = f(x(t), u(t)) + \theta(t) g((x(t), u(t))$$

which requires the first order conditions:

1. $$h_u = 0$$

2. $$\dot{x}\,(t) = g(x(t), u(t))$$

3. $$h_x = \rho\theta - \dot{\theta}$$

2.4. Calculus of Variations

The basic calculus of variations problem is to choose a function $x\,(t)$ to maximize a sum:

$$J = \int_a^b f(x(t), \dot{x}(t), t)dt$$

we assume that f is continuously differentiable. Let X be the set of all real-valued continuously differentiable functions on $[a, b]$. We are looking for a particular $x \in X$. Suppose we have found the right function x^* and now consider a displacement function $h \in X$ with $h\,(a) = h\,(b) = 0$ and define a new function

$$x_\varepsilon(t) = x^*(t) + \varepsilon h(t)$$

under the assumption that x^* is optimal, J is maximized when ε is zero. This implies that

$$\frac{\partial\left[J_\varepsilon\right]}{\partial\varepsilon}=0$$

when $\varepsilon = 0$.

Notice that

$$\frac{\partial\left[J_\varepsilon\right]}{\partial\varepsilon}=\frac{\partial\left[\int_a^b f(x^*(t)+\varepsilon h(t),\dot{x}^*(t)+\varepsilon\dot{h}(t),t)dt\right]}{\partial\varepsilon}$$

and:

$$\frac{\partial\left[J_\varepsilon\right]}{\partial\varepsilon}=\int_a^b\left(\frac{\partial f(x^*)}{\partial x}h\right)dt+\int_a^b\left(\frac{\partial f(x^*)}{\partial\dot{x}}\dot{h}\right)dt$$

evaluated at the optimal x^*.

Integrate the second sum by parts ($\int u\ dv = vu - \int v\ du$). $u = f_{\dot{x}}(t)$ so $du = \frac{\partial f_{\dot{x}}}{\partial t}dt$, $dv = \dot{h}dt$ so $v = h$:

$$\int_a^b\left(\frac{\partial f(x^*)}{\partial\dot{x}}\dot{h}\right)dt = h(t)f_{\dot{x}}(t)|_a^b-\int_a^b h(t)\frac{\partial f_{\dot{x}}}{\partial t}dt$$

The first is zero because $h(a) = h(b) = 0$. Thus:

$$\frac{\partial\left[J_\varepsilon\right]}{\partial\varepsilon}=\int_a^b\left(\frac{\partial f(x^*)}{\partial x}h(t)\right)dt-\int_a^b h(t)\frac{\partial f_{\dot{x}}}{\partial t}dt$$

this must be zero so:

$$\int_a^b\left[\frac{\partial f(x^*)}{\partial x}-\frac{\partial f_{\dot{x}}}{\partial t}\right]h(t)dt=0$$

This must be true *for any displacement* $h(t)$. Thus,

$$\frac{\partial f(x^*)}{\partial x}-\frac{\partial f_{\dot{x}}}{\partial t}=0$$

Formally, this is called an Euler equation.[3]

These variational problems are similar to the "one shot deviations" principles considered in economics.

[3]Note $f_{\dot{x}}(x^*) = f_{\dot{x}}\left(x^*(t),\dot{x}^*(t),t\right)$ so $\frac{\partial f(x^*)}{\partial x}-\frac{\partial f_{\dot{x}}}{\partial t}=0$ implies that:

$$f_x-\frac{\partial f_{\dot{x}}}{\partial x}\frac{\partial x}{\partial t}-\frac{\partial f_{\dot{x}}}{\partial\dot{x}}\frac{\partial\dot{x}}{\partial t}-\frac{\partial f_{\dot{x}}}{\partial t}=0$$

which is a 2nd order differential equation with two boundary conditions.

Example: Minimum Distance Problem

$$J = \int_a^b \sqrt{1 + \dot{x}(t)^2} dt$$

Here $f = \sqrt{1 + \dot{x}(t)^2}$, the associated Euler equation is:

$$\frac{\partial f(x^*)}{\partial x} - \frac{\partial f_{\dot{x}}}{\partial t} = 0$$

there is no $x(t)$ term so $\frac{\partial f(x^*)}{\partial x} = 0$. Note also that

$$f_{\dot{x}} = \frac{\dot{x}}{(1 + \dot{x}^2)^{\frac{1}{2}}}$$

The euler equation then implies that

$$\frac{d}{dt} \left[\frac{\dot{x}}{(1 + \dot{x}^2)^{\frac{1}{2}}} \right] = 0$$

so that $\left[\frac{\dot{x}}{(1+\dot{x}^2)^{\frac{1}{2}}} \right]$ is a constant. This implies that $\dot{x}$ is constant (a straight line).

3. Dynamic Programming

3.1. Basic Problem:

The basic class of problems we want to attack are recursive problems of the form:

$$\max \sum_{t=0}^{\infty} \beta^t u(c_t, s_t)$$

subject to the transition equation

$$s_{t+1} = g(s_t, c_t)$$

We saw in class that this problem obeyed a functional equation called "Bellman's Equation":

$$V(s) = \max_c \{u(c, s) + \beta V(g[c, s])\}$$

This equation is a functional equation meaning that it maps functions into other functions. If you start with a function f_0 you can get a new function f_1 via:

$$f_1(s) = \max_c \{u(c, s) + \beta f_0(g[c, s])\}$$

I claimed in class that this functional mapping was a contraction mapping and that therefore there is a unique fixed point that solves it. In other words, there is only one function V that maps into itself. Furthermore, if we begin with an arbitrary function that is not V and then iterate on the Bellman equation (i.e. start with $f_0 \neq V$ and get f_1 as above, then put in f_1 and get f_2 etc...) then the resulting equations will converge to the true value function (i.e. $f_n \to V$). In addition, typically V will be concave and differentiable in the interior of the state space.

The solution to a Dynamic Programming problem is a "policy function" which tells the agent what he or she should do if they find themselves in state s. Denote this policy function as $h(s)$; $h : S \to C$ where S is the state space and C is the control space. If you are in state $s \in S$ then $h(s) = c \in C$ is the choice of the control you should make (this is the "right move" to make in this situation).

Because h is the optimal rule, V will satisfy:

$$V(s) = u(h(s), s) + \beta V(g[h(s), s])$$

(note that the max operator is gone since h is already taking care of the maximization).

3.2. Euler Equations in a Dynamic Programming Context

Many problems can be set up with either Lagrangians or with Dynamic Programming. For these problems the Euler equations that you get from the Lagrangians will also pop up from the Dynamic Programming setup.

Let's see how this is done.

Note that because V is optimal then (provided that we are not constrained in our choice of c) we must have:

$$\frac{\partial u}{\partial c} + \beta \frac{\partial V}{\partial s}(s') \frac{\partial g}{\partial c} = 0 \tag{1}$$

The policy function must satisfy this first order condition.

But what is $\frac{\partial V}{\partial s}$? Because V was unknown, it seems unlikely that we will know $\frac{\partial V}{\partial s}$ and this makes the interpretation of this first order condition difficult. Luckily we can often find $\frac{\partial V}{\partial s}$.

Because $h(s)$ is the (unknown) optimal policy function, then, as above:

$$V(s) = u(h(s), s) + \beta V(g[h(s), s])$$

If we differentiate this with respect to s we get:

$$\frac{\partial V}{\partial s}(s) = \frac{\partial u}{\partial s} + \frac{\partial u}{\partial c}\frac{\partial h}{\partial s} + \beta \frac{\partial V}{\partial s}(s') \left[\frac{\partial g}{\partial s} + \frac{\partial g}{\partial c}\frac{\partial h}{\partial s} \right]$$

Now group the terms with the $\frac{\partial h}{\partial s}$ on them.

$$\frac{\partial V}{\partial s}(s) = \frac{\partial u}{\partial s} + \beta \frac{\partial V}{\partial s}(s') \frac{\partial g}{\partial s} + \left\{ \beta \frac{\partial V}{\partial s}(s') + \frac{\partial u}{\partial c} \right\} \frac{\partial h}{\partial s}$$

The term in the brackets must be zero (see equation (1) – the first order condition from the choice of c). So,

$$\frac{\partial V}{\partial s}(s) = \frac{\partial u}{\partial s} + \beta \frac{\partial V}{\partial s}(s') \frac{\partial g}{\partial s}$$

Well this tells me a little bit more but it looks like if I want to know $\frac{\partial V}{\partial s}(s)$ I will need to know $\frac{\partial V}{\partial s}(s')$ so I'm still stuck – *unless* I was lucky enough to have $\frac{\partial g}{\partial s} = 0$.

In class I said we could set up the problem so that $\frac{\partial g}{\partial s}$ dropped out. This essentially requires writing the Bellman Equation so that we choose s' rather than c.

More formally, the transition function is:

$$s' = g(c, s)$$

If this can be inverted, then we could write c as a function of s and s':

$$c = \rho(s, s')$$

Then we could write the Bellman equation as:

$$V(s) = \max_{s'} \{u(\rho(s, s'), s) + \beta V(s')\}$$

where we are now choosing next period's state variable rather than today's control variable.

Let's take our first order condition again (now with respect to s' instead of c):

$$\frac{\partial u}{\partial c}\frac{\partial \rho}{\partial s'} + \beta \frac{\partial V}{\partial s}(s') = 0$$

Again, I need to know $\frac{\partial V}{\partial s}$. Let's try the same procedure as before. Let h be the policy function so that:

$$V(s) = u(\rho(s, h(s)), s) + \beta V(h(s))$$

Differentiating and using the first order condition to get rid of the terms with $\frac{\partial h}{\partial s}$ (i.e. the envelope theorem) gives me:

$$\frac{\partial V}{\partial s}(s) = \frac{\partial u}{\partial s} + \frac{\partial u}{\partial c}\frac{\partial \rho}{\partial s}$$

This implies that the first order condition can be written as:

$$\frac{\partial u}{\partial c}\frac{\partial \rho}{\partial s'} + \beta \left[\frac{\partial u}{\partial s}(s', c') + \frac{\partial u}{\partial c}(s', c')\frac{\partial \rho}{\partial s}\right] = 0$$

so that there is no reference to the unknown functions V or h.

3.3. Examples

To clarify things, let's look at a couple of examples.

3.3.1. Example 1: The Growth Model:

The growth model requires the consumer to maximize

$$\sum_{t=0}^{\infty} \beta^t u(c_t)$$

subject to:

$$k_{t+1} = k_t (1-\delta) + k_t^{\alpha} - c_t$$

This gives rise to the Bellman equation:

$$V(k) = \max_{c} \{u(c) + \beta V(k(1-\delta) + k^{\alpha} - c)\}$$

Instead, let's set it up so they choose next period's state variable k':

$$V(k) = \max_{k'} \{u(k(1-\delta) + k^{\alpha} - k') + \beta V(k')\}$$

The first order condition (for the choice of k') is:

$$-u'(c) + \beta \frac{\partial V}{\partial k}(k') = 0$$

and the derivative of the value function is:

$$\frac{\partial V}{\partial k}(k) = u'(c)\left[(1-\delta) + \alpha k^{\alpha}\right]$$

(I can ignore the effects of the policy function – i.e. the terms that involve $\frac{\partial k'}{\partial k}$ – because of the envelope theorem). Plugging this into the first order condition (to get rid of the terms with the "V's" on them) gives me:

$$u'(c) = \beta u'(c')\left[(1-\delta) + \alpha (k')^{\alpha}\right]$$

which is the standard Euler equation for the growth model.

3.3.2. Example 2: Consumption Under Uncertainty:

Now consider the uncertainty case in which a consumer get's stochastic income:

$$w_t = \bar{w} + \varepsilon_t$$

where $\varepsilon_t \sim i.i.d.$. The consumer's utility is:

$$E_0 \left[\sum_{t=0}^{\infty} \beta^t u(c_t) \right]$$

subject to:

$$c_t + b_t = w_t + b_{t-1}(1+r)$$

The state is b_{t-1} and w_t (i.e. the assets and income you enter the period with, respectively).

Let's redefine existing assets as:

$$A = b_{t-1}(1+r)$$

so that saving today is $b_t = \frac{A'}{1+r}$. Then the Bellman equation is:

$$V(A, w) = \max_{A'} \left\{ u\left(A + w_t - \frac{A'}{1+r} \right) + \beta E\left[V(A', w') \right] \right\}$$

(**note that I am again picking tomorrows state**).[4]

The first order condition is:

$$\partial A' : -u'(c) \frac{1}{1+r} + \beta E\left[\frac{\partial V}{\partial A}(A', w') \right] = 0$$

and the derivative of V with respect to A is simply:

$$\frac{\partial V}{\partial A}(A, w) = u'(c)$$

(again note that I am using the envelope theorem to write this). Plugging in to get rid of the terms with the V's gives :

$$u'(c) \frac{1}{1+r} = \beta E\left[u'(c') \right]$$

which is our standard stochastic Euler equation (see your class notes from the 2^{nd} or 3^{rd} class).

---

[4]Note that this expectation is taken with respect to the distribution of ε.

$$V(A, w) = \max_{A'} \left\{ u\left(A + w_t - \frac{A'}{1+r} \right) + \beta \int \left[V(A', \bar{w} + \varepsilon) f(\varepsilon) d\varepsilon \right] \right\}$$

(assuming that ε has a density).

3.4. Attack Strategy:

For value functions where the choices are "smooth" – i.e. the controls are not constrained, or if they are, the constraints never bind and don't influence the problem, the following approach is often rewarding.

1. Look at the problem. Decide which variables are "state" variables and which ones are "control" variables.

2. Write down the Bellman Equation (it is probably good to write it so that you are choosing next period's state rather than today's control – see the discussion above).

3. Take the derivative w.r.t. the choice variable (either the control or next period's state depending on how you set it up) and set it equal to zero.

4. Take a derivative of V with respect to the state and plug in to get rid of the term's with the V's on them.

At the end of this you should have a well defined Euler equation which is subject to standard interpretation.

3.5. One last technical note:

YOU CAN IGNORE THIS LAST SECTION IF YOU WANT

The principle of optimality says that if $V(s)$ is the maximum utility I can expect given that I am in state s then it the function $V(s)$ must satisfy:

$$V(s) = \max_c \{u(c, s) + \beta V(g[c, s])\}$$

A moment ago, I considered a case with uncertainty in which the Bellman equation took a form similar to:

$$V(s, \varepsilon) = \max_c \{u(c, s, \varepsilon) + \beta E[V(g(c, s, \varepsilon), \varepsilon')]\}$$

and the definition of an expected value implies that:

$$V(s, \varepsilon) = \max_c \left\{u(c, s, \varepsilon) + \beta \int V(g(c, s, \varepsilon), \varepsilon') f(\varepsilon') d\varepsilon'\right\}$$

This looked harmless enough at the time but there is one slight problem.

Suppose that the function $V(s)$ (the true value function) is not measurable. In this case the integral $\int V(g(c, s, \varepsilon), \varepsilon') f(\varepsilon') d\varepsilon'$ will not be well defined and the value function will not satisfy the Bellman equation.

In cases like this we are basically screwed. We can safely ignore them since they are "pathological" in the sense that they almost never arise in economic problems. See Stokey, Lucas and Prescott [1989] if you are a glutton for punishment.

4. Vector Autoregressions (VARs)

Often we will find solutions to linear models given in the following form:

$$Y_t = AY_{t-1} + B\varepsilon_t$$

where ε_t is a vector of exogenous shocks with a know variance/covariance matrix $E\left[\varepsilon_t^2\right] =$. We want to be able to find several features of the moments of Y without simulation.

4.1. Unconditional Variance

The first question that arises in this context concerns the unconditional variance of Y itself. That is, what is the long run variance of each of the components of Y without any knowledge of past Y's.

Recall that the variance of a vector is found as:

$$V[y] = E\left[\left(y - \mu_y\right)\left(y - \mu_y\right)'\right]$$

(not to be confused with $y'y$ which is the sum of squares)[5]

If we assume that the mean of Y is zero (which is also quite common), then the variance of our process is given by $E\left[Y_t Y_t'\right]$. This leads to the following relationship:

$$\begin{aligned} E\left[Y_t Y_t'\right] &= E\left[\left(AY_{t-1} + B\varepsilon_t\right)\left(AY_{t-1} + B\varepsilon_t\right)'\right] \\ &= E\left[\left(AY_{t-1} + B\varepsilon_t\right)\left(Y_{t-1}'A' + \varepsilon_t'B'\right)\right] \\ &= E\left[\left(AY_{t-1}Y_{t-1}'A' + B\varepsilon_t\varepsilon_t'B' + B\varepsilon_t Y_{t-1}'A' + AY_{t-1}\varepsilon_t'B'\right)\right] \end{aligned}$$

[5]... i.e., let:

$$y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

then:

$$yy' = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\left(y_1 \ y_2\right) = \begin{pmatrix} y_1^2 & y_1 y_2 \\ y_2 y_1 & y_2^2 \end{pmatrix}$$

which when expectations are taken gives us variances (on the diagonal) and covariances (off the diagonal). On the other hand,

$$y'y = \left(y_1 \ y_2\right)\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = y_1^2 + y_2^2$$

which is the sum of squares (sum of variances)).

since $E[\varepsilon_t] = 0$ we have[6]:

$$E[Y_t Y_t'] = AE[Y_{t-1}Y_{t-1}']A' + BE[\varepsilon_t \varepsilon_t']B'$$

Denote the variance of Y as V. Then:

$$V = AVA' + B\Sigma B'$$

the V that solves this equation is the unconditional variance of Y.

To solve for this matrix use the following vectorization trick:

$$vec(ABC) = [C' \otimes A]\, vec(B)$$

Now, letting $\Sigma^B = B\Sigma B'$ we have:

$$V = AVA' + \Sigma^B$$

using the $vec()$ operator:

$$vec(V) = [A \otimes A]\, vec(V) + vec(\Sigma^B)$$

and

$$vec(V)\{I - [A \otimes A]\} = vec(\Sigma^B)$$
$$vec(V) = vec(\Sigma^B)[I - (A \otimes A)]^{-1}$$

(a simple "devec" program can be used to reform V).

4.1.1. Practical Problems

Notice that the solution involves inverting the matrix $I - (A \otimes A)$. Often this is fine but there may be cases in which the matrix is very large. For instance, if there are 20 variables in Y then the matrix $I - (A \otimes A)$ will be 400×400. This is a lot for any computer – there are 160,000 elements in the matrix.

It is sometimes better to use an approximation that avoids this large matrix. Recall that for a univariate case,

$$y_t = \rho y_{t-1} + \varepsilon_t$$

[6] Note that $E[\varepsilon_t Y_{t-1}] = 0$ – past Y's are not correlated with current shocks – although current shocks are correlated with future Y's so $E[\varepsilon_t Y_{t-1}] \neq 0$ in general.

we would have:

$$\sigma_y^2 = \rho^2 \sigma_y^2 + \sigma_\varepsilon^2$$

and we could find:

$$\sigma_y^2 = \frac{\sigma_e^2}{1-\rho^2}$$

Alternatively we could use the following:

$$\begin{aligned}
\sigma_y^2 &= \sigma_\varepsilon^2 + \rho^2 \sigma_y^2 = \sigma_\varepsilon^2 + \rho^2 \left(\sigma_\varepsilon^2 + \rho^2 \sigma_y^2 \right) \\
&= \sigma_\varepsilon^2 + \rho^2 \sigma_\varepsilon^2 + \rho^4 \left(\sigma_\varepsilon^2 + \rho^2 \sigma_y^2 \right)
\end{aligned}$$

etc. to get:

$$\sigma_y^2 = \sum_{j=0}^{\infty} \left(\rho^2 \right)^j \sigma_\varepsilon^2$$

which again adds up to $\frac{\sigma_e^2}{1-\rho^2}$.

In the vector case, it turns out that addition of matrices is easier than inverting very large matrices. Thus we will form:

$$\begin{aligned}
V &= {}^{B} + AVA' = {}^{B} + A \left({}^{B} + AVA' \right) A' = ... \\
&= {}^{B} + A\ {}^{B} A' + AA\ {}^{B} A'A'
\end{aligned}$$

On the computer, it is easier to form the variance as follows: define $V_1 = {}^{B}$. Then define $V_{j+1} = AV_jA'$. Do a loop with 100 or so iterations and the variance you end up with will be very close to the correct V.

4.2. Autocorrelations

The j^{th} autocorrelation of Y is defined to be:

$$\Gamma_j = E \left[\left(Y_t - \mu \right) \left(Y_{t-j} - \mu \right)' \right]$$

Finding these autocorrelation matrices is relatively easy. Note that Γ_0 is just V. Furthermore notice that Γ_1 is given by

$$\begin{aligned}
\Gamma_1 &= E \left[Y_t Y_{t-1}' \right] = E \left[\left(AY_{t-1} + B\varepsilon_t \right) Y_{t-1}' \right] \\
&= E \left[AY_{t-1} Y_{t-1}' + B\varepsilon_t Y_{t-1}' \right]
\end{aligned}$$

Taking expectations gives:

$$\Gamma_1 = AE \left[Y_t Y_{t-1}' \right] = AV$$

Furthermore, note that $\Gamma_j = A\Gamma_{j-1}$:

$$\begin{aligned}
\Gamma_j &= E\left[Y_t Y'_{t-j}\right] = E\left[\left(AY_{t-1} + B\varepsilon_t\right)Y'_{t-j}\right] \\
&= E\left[AY_{t-1}Y'_{t-j} + B\varepsilon_t Y'_{t-j}\right] \\
&= E\left[AY_{t-1}Y'_{t-j}\right]
\end{aligned}$$

which is just $A\Gamma_{t-j}$. Thus,

$$\Gamma_j = A^j V$$

Note that the autocorrelation matrices are not symmetric (i.e. $\Gamma_j \neq \Gamma_{-j}$). To see this note that Γ_{-1} is given by:

$$\Gamma_{-1} = E\left[Y_t Y'_{t+1}\right]$$

we can't use the trick we used before since $E\left[Y_{t+1}\varepsilon_t\right] \neq 0$. Instead, we can go in "reverse":

$$\Gamma_{-1} = E\left[Y_t Y'_{t+1}\right] = E\left[Y_t\left(AY_t + B\varepsilon_{t+1}\right)'\right]$$

taking expectations gives us:

$$\Gamma_{-1} = VA'$$

and in general:

$$\Gamma_{-j} = V\left(A'\right)^j$$

If I take the transpose of this (and recall that V is symmetric) then I have:

$$\Gamma'_{-j} = A^j V = \Gamma_j$$

5. Solving Linear Rational Expectations Models.

Many models can be written in the following form:

$$E_t \left[\begin{pmatrix} C_{t+1} \\ K_{t+1} \end{pmatrix} \right] = M \begin{pmatrix} C_t \\ K_t \end{pmatrix}$$

where C_t is an $n \times 1$ vector of free variables at time t and K_t is an $m \times 1$ vector of exogenous or state variables at date t. The free variables consist of anything that can be changed at date t depending on the particular state K_t the economy is in. Here, expectations are conditional on information available at date t.

Since $\begin{pmatrix} C_t \\ K_t \end{pmatrix}$ is $(m+n) \times 1$, the matrix M, must be $(m+n) \times (m+n)$ square.

Alternatively, we could write:

$$\begin{pmatrix} C_{t+1} \\ K_{t+1} \end{pmatrix} = M \begin{pmatrix} C_t \\ K_t \end{pmatrix} + \quad _{t+1}$$

where $\;_{t+1}$ is a vector including expectational errors (like $C_{t+1} - E_t [C_{t+1}]$) plus any shocks to the state (e.g. if A_{t+1} were the technology parameter we would have the technology shock ε_{t+1}^A in $\;_{t+1}$).

5.1. Example Part I.

Consider a simple RBC model:

$$\max E_0 \sum_{t=0}^{\infty} \beta^t \ln c_t$$

subject to

$$c_t + k_{t+1} = A_t k_t^\alpha + (1-\delta) k_t$$
$$A_{t+1} = (1-\rho) + \rho A_t + \varepsilon_{t+1}^A$$

The state here is the 2×1 vector: $\begin{bmatrix} k_t \\ A_t \end{bmatrix}$ and the free variable is: $[c_t]$.

The first order condition for the choice of c_t is:

$$\frac{1}{c_t} = \beta E_t \left\{ \frac{1}{c_{t+1}} \left[\alpha A_{t+1} k_{t+1}^{\alpha-1} + (1-\delta) \right] \right\}$$

The (nonstochastic) steady state of the model is given by:

$$
\begin{aligned}
r &= \frac{1}{\beta} - (1-\delta) = \alpha A k^{\alpha-1} \\
k &= \left(\frac{\alpha A}{r}\right)^{\frac{1}{1-\alpha}} = \left(\frac{\alpha}{r}\right)^{\frac{1}{1-\alpha}} \\
y &= k^{\alpha} \\
c &= y - \delta k
\end{aligned}
$$

Thus the solution to the model will satisfy the following three equations:

$$
\begin{aligned}
1 &: \frac{1}{c_t} = \beta E_t\left[\frac{1}{c_{t+1}}\left\{(1-\delta) + \alpha A_{t+1} k_{t+1}^{\alpha-1}\right\}\right] \\
2 &: k_{t+1} = k_t(1-\delta) + A_t k_t^{\alpha} - c_t \\
3 &: A_{t+1} = (1-\rho) + \rho A_t + \varepsilon_{t+1}
\end{aligned}
$$

"Log linearize" every equation ($t+1$ on the right):

$$
\begin{aligned}
1 &: E_t\left[-\tilde{c}_{t+1} + \beta r\left(\tilde{A}_{t+1} + (\alpha-1)\tilde{k}_{t+1}\right)\right] = -\tilde{c}_t \\
2 &: \tilde{k}_{t+1} = \frac{1}{\beta}\tilde{k}_t - \frac{c}{k} c_t + k^{\alpha-1}\tilde{A}_t \\
3 &: \tilde{A}_{t+1} = \rho\tilde{A}_t + \tilde{\varepsilon}_{t+1}
\end{aligned}
$$

where the last one also implies that:

$$
E_t\left[\tilde{A}_{t+1}\right] = \rho\tilde{A}_t
$$

This gives us the following system:

$$
\begin{bmatrix} -1 & \beta r(\alpha-1) & \beta r \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\left[\begin{pmatrix} \tilde{c}_{t+1} \\ \tilde{k}_{t+1} \\ \tilde{A}_{t+1} \end{pmatrix}\right]
=
\begin{bmatrix} -1 & 0 & 0 \\ -\frac{c}{k} & \beta^{-1} & k^{\alpha-1} \\ 0 & 0 & \rho \end{bmatrix}
\begin{pmatrix} \tilde{c}_t \\ \tilde{k}_t \\ \tilde{A}_t \end{pmatrix}
+
\begin{pmatrix} E_t\left[c_{t+1}\right] - c_{t+1} \\ E_t\left[k_{t+1}\right] - k_{t+1} \\ \varepsilon_{t+1}^{A} \end{pmatrix}
$$

or since the expected values of the expectational discrepancies and the technology shock are zero,

$$
\begin{bmatrix} -1 & \beta r(\alpha-1) & \beta r \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
E_t\left[\begin{pmatrix} \tilde{c}_{t+1} \\ \tilde{k}_{t+1} \\ \tilde{A}_{t+1} \end{pmatrix}\right]
=
\begin{bmatrix} -1 & 0 & 0 \\ -\frac{c}{k} & \beta^{-1} & k^{\alpha-1} \\ 0 & 0 & \rho \end{bmatrix}
\begin{pmatrix} \tilde{c}_t \\ \tilde{k}_t \\ \tilde{A}_t \end{pmatrix}
$$

which is

$$B_1 E_t \left[\begin{pmatrix} \tilde{c}_{t+1} \\ \tilde{k}_{t+1} \\ \tilde{A}_{t+1} \end{pmatrix} \right] = B_2 \begin{pmatrix} \tilde{c}_t \\ \tilde{k}_t \\ \tilde{A}_t \end{pmatrix}$$

We can invert B_1 to get:

$$E_t \left[\begin{pmatrix} \tilde{c}_{t+1} \\ \tilde{k}_{t+1} \\ \tilde{A}_{t+1} \end{pmatrix} \right] = B_1^{-1} B_2 \begin{pmatrix} \tilde{c}_t \\ \tilde{k}_t \\ \tilde{A}_t \end{pmatrix} = M \begin{pmatrix} \tilde{c}_t \\ \tilde{k}_t \\ \tilde{A}_t \end{pmatrix}$$

which is in "standard" form.

5.2. Diagonalization and the "Policy Function"

We are given the system:

$$E_t \left[\begin{pmatrix} C_{t+1} \\ K_{t+1} \end{pmatrix} \right] = M \begin{pmatrix} C_t \\ K_t \end{pmatrix}$$

We might be tempted to announce victory right away since we could plot out a realization of the variables from any starting position $\begin{pmatrix} C_t \\ K_t \end{pmatrix}$. This is wrong though since we can't know *a priori* any such starting position. We can only know a beginning state K_t. The free/ choice variables C_t will depend on the state $C_t = C(K_t)$. It is this policy function that we search for know.

step 1 Since M is square, we can break it up as follows:

$$M = \Gamma \Lambda \Gamma^{-1}$$

where Γ is a matrix of eigenvectors (columns) and Λ is the associated diagonal matrix of eigenvalues. In other words:

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & ... & 0 \\ 0 & 0 & 0 & \lambda_{m+n} \end{pmatrix} \text{ and } \Gamma = (v_1, v_2, ...v_{m+n})$$

where each v_i is the column eigenvector $((m+n) \times 1)$ associated with eigenvalue λ_i. A useful fact is that you can freely rearrange the eigenvalues and the eigenvectors while leaving $\Gamma\Lambda\Gamma^{-1}$ unchanged. That is if $\Gamma\Lambda\Gamma^{-1}$ as above, then the matrices:

$$\Lambda_2 = \begin{pmatrix} \lambda_i & 0 & 0 & 0 \\ 0 & \lambda_j & 0 & 0 \\ 0 & 0 & ... & 0 \\ 0 & 0 & 0 & \lambda_k \end{pmatrix} \text{ and } \Gamma_2 = (v_i, v_j, ...v_k)$$

will also solve $\Gamma_2\Lambda_2\Gamma_2^{-1} = \Gamma\Lambda\Gamma^{-1} = M$. To see this, recall the definition of an eigenvector. An eigenvector of the matrix B is a column vector v that satisfies:

$$Bv = \lambda v$$

where λ is a scalar (called the eigenvalue). Typically, if B is $n \times n$, there will be n distinct eigenvectors (with associated eigenvalues). Notice that since each will satisfy this equation:

$$Bv_1 = \lambda_1 v_1, \; Bv_2 = \lambda_2 v_2, \; ... \; Bv_n = \lambda_n v_n$$

Let $\Gamma_1 = [v_1, v_2, ...v_n]$ for an arbitrary (exhaustive) ordering of the vectors v_i. Then, no matter how the eigenvalues / vectors are ordered, the following statement must hold with equality:

$$B\Gamma_1 = \Gamma_1 \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

As a result,

$$B = \Gamma_1 \Lambda \Gamma_1^{-1}$$

for any ordering.

With this in mind let's assume that $M = \Gamma\Lambda\Gamma^{-1}$ where Λ is a diagonal matrix of eigenvalues ordered by their absolute value. That is let Λ be given by:

$$\Lambda = \begin{pmatrix} \underset{i \times i}{J_1} & 0 \\ 0 & \underset{j \times j}{J_2} \end{pmatrix}$$

where J_1 contains eigenvalues that are less than *or equal to* 1 in absolute value[7] and J_2 contains the eigenvalues that are greater than one in absolute value. Here

[7] If the eigenvalue is complex then we would require it to be inside the unit circle.

there are i stable eigenvalues and $j = m + n - i$ unstable ones. Γ is a matrix of column eigenvectors arranged in the same order as the eigenvalues of Λ.

Consider the matrix Γ^{-1} as:

$$\Gamma^{-1} = \begin{pmatrix} \underset{i\times n}{G_{11}} & \underset{i\times m}{G_{12}} \\ \underset{j\times n}{G_{21}} & \underset{j\times m}{G_{22}} \end{pmatrix}$$

We now have:

$$E_t \left[\begin{pmatrix} C_{t+1} \\ K_{t+1} \end{pmatrix} \right] = \Gamma \Lambda \Gamma^{-1} \begin{pmatrix} C_t \\ K_t \end{pmatrix} \tag{2}$$

5.2.1. Solution Part I.

Premultiply the system (2) by Γ^{-1}. This gives us:

$$E_t \left[\Gamma^{-1} \begin{pmatrix} C_{t+1} \\ K_{t+1} \end{pmatrix} \right] = \Lambda \Gamma^{-1} \begin{pmatrix} C_t \\ K_t \end{pmatrix}$$

Define the matrices Z_{1t} and Z_{2t} as follows:

$$\begin{pmatrix} Z_{1t} \\ Z_{2t} \end{pmatrix} \equiv \Gamma^{-1} \begin{pmatrix} C_t \\ K_t \end{pmatrix} = \begin{pmatrix} G_{11} C_t + G_{12} K_t \\ G_{21} C_t + G_{22} K_t \end{pmatrix}$$

We now have the system:

$$E_t \left[\begin{pmatrix} Z_{1t+1} \\ Z_{2t+1} \end{pmatrix} \right] = \begin{pmatrix} \underset{i\times i}{J_1} & 0 \\ 0 & \underset{j\times j}{J_2} \end{pmatrix} \begin{pmatrix} Z_{1t} \\ Z_{2t} \end{pmatrix}$$

Notice that the evolution of the transformed variables Z_{1t} and Z_{2t} are governed only by J_1 and J_2 respectively. We can iterate the system forward to see that:

$$E_t \left[Z_{1t+T} \right] = (J_1)^T Z_{1t}$$

and

$$E_t \left[Z_{2t+T} \right] = (J_2)^T Z_{2t}$$

Notice that we can solve for the path of Z_1 independently of the path of Z_2. This is because of the diagonalization. Also, the expected values of the Z_1 series will converge to zero (since $J_1 << 1$) while the expected value of Z_2 will diverge to $\pm\infty$ for non zero Z_2 (since $J_2 >> 1$).[8]

No solution will allow such a divergence. Consequently, we must have $Z_{2t} = 0$ $\forall t$. What this means is that

$$Z_{2t} = G_{21}C_t + G_{22}K_t = 0$$

in **every** period (otherwise the system blows up). This equivalently means that:

$$G_{21}C_t = -G_{22}K_t$$

If G_{21} is square then we can (generically) find:

$$C_t = -\left[G_{21}\right]^{-1} G_{22}K_t$$

This tells us how the free variables (C_t) must be set according to the values of the predetermined (state) variables (K_t).

What does it mean for G_{21} to be square? It means that $j = n$ or that the number of unstable roots in Λ is exactly equal to the number of "non-predetermined" variables (C). For every unstable root, I have a "non-predetermined" variable that I can use to "zero" it. This property of a dynamic system corresponds exactly to the "saddle path" property of growth models.[9]

The resulting system is stable.

5.2.2. Solution Part II.

Now, armed with the linear policy rule $-\left[G_{21}\right]^{-1} G_{22}$ (which is $n \times m$) we can solve the system.

We would like the system in a VAR form:

$$Y_t = AY_{t-1} + B\delta_t$$

(with$Y_t = \begin{pmatrix} C_t \\ K_t \end{pmatrix}$ and a shock vector $\delta = n + m \times 1$).

[8] This is not quite correct since we do allow for the presence of a unit root in J_1. In such a case, some of the elements of Z_1 will not converge to zero in expected value.

[9] In the Ramsey model there is one predetermined variable (K_t) and one "choice" variable (C_t). The system is a saddle path – this implies that there is a unique value for C that is picked for each value of K.

Recall that the system will satisfy:

$$E_t\left[\left(\begin{array}{c} C_{t+1} \\ K_{t+1} \end{array}\right)\right] = M\left(\begin{array}{c} C_t \\ K_t \end{array}\right)$$

We already know that if we know K_t, we can figure out C_t from $C_t = -\left[G_{21}\right]^{-1} G_{22}K_t$. Let's rewrite M ($n+m \times n+m$) as follows:

$$M = \left(\begin{array}{cc} \underset{n\times n}{H_{11}} & \underset{n\times m}{H_{12}} \\ \underset{m\times n}{H_{21}} & \underset{m\times m}{H_{22}} \end{array}\right)$$

In expectation K_t satisfies:

$$E\left[K_t\right] = H_{21}C_{t-1} + H_{22}K_{t-1}$$

and furthermore, the *actual* value of K_t will satisfy:

$$K_t = H_{21}C_{t-1} + H_{22}K_{t-1} + \varepsilon_t$$

where ε_t are any shocks to the state vector (technology shocks, government spending shocks, taste shocks, depreciation shocks, etc.). Since $C_{t-1} = -\left[G_{21}\right]^{-1} G_{22}K_{t-1}$ we can write:

$$K_t = \left[-H_{21}\left[G_{21}\right]^{-1} G_{22} + H_{22}\right] K_{t-1} + \varepsilon_t$$

As a result, the "lower right part" of A is given by: $\left[-H_{21}\left[G_{21}\right]^{-1} G_{22} + H_{22}\right]$. (In fact, we could simply leave A as M on the bottom "row"). The bottom "row" of B is given by the identity matrix.

To get C_t we need to know K_t (which we do now). Specifically

$$C_t = -\left[G_{21}\right]^{-1} G_{22}\left[H_{21}C_{t-1} + H_{22}K_{t-1}\right] - \left[G_{21}\right]^{-1} G_{22}\varepsilon_t.$$

In expectation C_t also satisfies:

$$E_{t-1}\left[C_t\right] = H_{11}C_{t-1} + H_{12}K_{t-1}$$

so that we can write $A = M$ and $B = \left[\begin{array}{cc} 0 & -\left[G_{21}\right]^{-1} G_{22} \\ 0 & I \end{array}\right]$.

Problem: While the VAR form:

$$Y_t = MY_{t-1} + \left[\begin{array}{cc} 0 & -\left[G_{21}\right]^{-1} G_{22} \\ 0 & I \end{array}\right]\varepsilon_t$$

is mathematically correct, it is numerically prone to problems. The reason for this is that the policy function $P = -G_{21}^{-1}G_{22}$ is not going to be exact. Thus the unstable elements in Z_2 will not all be numerically zero. One immediate consequence of this is the fact that C_t and thus K_{t+1} will both be wrong to a very small degree. Unfortunately, the system cannot tolerate errors in C. Since the roots that operate on Z_2 will explode for any deviation from zero the system will eventually blow up.

To correct for this we change the VAR form to:

$$Y_t = AY_{t-1} + B\varepsilon_t$$

with B as before but with A given as:

$$A = \begin{bmatrix} 0 & -[G_{21}]^{-1} G_{22} \left[-H_{21} [G_{21}]^{-1} G_{22} + H_{22}\right] \\ 0 & \left[-H_{21} [G_{21}]^{-1} G_{22} + H_{22}\right] \end{bmatrix}$$

Again C_t and K_{t+1} will have errors because of the small numerical errors in the policy function. BUT importantly, the errors will not build up as they did before. The reason for this is that the errors are "transmitted" to the future by being embedded entirely in an error to K_{t+1}. Luckily the system is stable for any value of K so these errors will not accumulate.

Note that since the system solves :

$$E_{t-1}[Y_t] = AY_{t-1}$$

and Y_t actually obeys:

$$Y_t = AY_{t-1} + B\varepsilon_t$$

It must be the case that $B\varepsilon_t$ is the vector of expectational errors and exogenous shocks to the system.

5.3. Summary:

1. count free (n) and state (m)
2. type in M
3. MATLAB eig(M)
4. count stable, unstable
5. form G_{ij} and then get policy function $[G_{21}]^{-1} G_{22}$
6. form VAR representation.

5.4. "Redundant" Variables.

That's great. We have a solution for the basic RBC model. But suppose that we wanted to do some simple modifications to the model. For instance, suppose that we wanted to look at plots (impulse responses) of wages or output. These weren't in our original model but we might want to add them so we could look at there behavior.

Additionally, we might want to augment the model with variable labor (an important feature of business cycles).

Both of these modifications present a special problem because they typically introduce equations that are not intrinsically forward looking. That is, these new equations will have zeros in the corresponding row of B_1 (in our program). This will imply that B_1^{-1} does not exist and we will have difficulty expressing the solution in "standard" form:

$$E_t\left[Y_{t+1}\right] = MY_t$$

Consider the vector of variables Y (the variables we care about) as being made up of two pieces:

$$Y_t = \begin{pmatrix} y_t \\ x_t \end{pmatrix}$$

where y is a vector containing all of the forward looking variables (the states – like K, Z, G,... etc., and the co-states – like C_t, (in sticky price models P_t^*)...etc.) and x contains all of the redundant variables. These could be derived from knowledge of the equilibrium behavior of y BUT solving for the equilibrium behavior of y may involve some of the x's (e.g. labor supply is not forward looking but it will alter the choices / dynamics in the model).

5.4.1. System Reduction

Let's write down the system (all of the equations) as:

$$AE_t\left[Y_{t+1}\right] = BY_t$$

Clearly A will have a lot of zeros in it since many of the elements of Y_t are not forward looking. We can decompose A and B as follows:

$$\begin{pmatrix} a_{11} & a_{12} \\ 0 & 0 \end{pmatrix} E_t \begin{bmatrix} y_{t+1} \\ x_{t+1} \end{bmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{bmatrix} y_t \\ x_t \end{bmatrix}$$

Note that since the bottom "row" of Y_{t+1} is zero we know that:

$$0 = b_{21}y_t + b_{22}x_t$$

Since b_{22} is square, this implies that:

$$x_t = Fy_t$$

where $F = -(b_{22})^{-1}b_{21}$. Let's now rewrite the remaining parts of the system:

$$AE_t[Y_{t+1}] = BY_t$$

$$\begin{pmatrix} a_{11} & a_{12} \\ 0 & 0 \end{pmatrix} E_t \begin{bmatrix} y_{t+1} \\ Fy_{t+1} \end{bmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{bmatrix} y_t \\ Fy_t \end{bmatrix}$$

implies that:

$$[a_{11} + a_{12}F]\, E[y_{t+1}] = [b_{11} + b_{12}F]\, y_t$$

Since $a_{11} + a_{12}F \neq 0$ we can invert to get:

$$E[y_{t+1}] = My_t$$

where $M = [a_{11} + a_{12}F]^{-1}[b_{11} + b_{12}F]$. Now that we have reduced the system to only forward looking variables we can solve the system for a VAR form as before:

$$y_t = Ay_{t-1} + B_0\varepsilon_t$$

How do we get the full VAR though?

5.4.2. VAR form:

We need to rebuild the system in Y_t. This is not that hard since we have the equilibrium law of motion for the states and the co-states y_t and we know that the static variables solve $x = Fy$. Thus:

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} A_0 & 0 \\ FA_0 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} B_0 & 0 \\ FB_0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix}$$

which is a VAR:

$$Y_t = A_1Y_{t-1} + B_1\eta_t$$

Note that we have to separate the states from the costates when we order the subvector y and we have to separate the static variables x.

6. "Proof" of the Hayashi Theorem[10]

In general marginal q and average q (or Tobin's q) will be different just like the average product of labor is different from the marginal product of labor.

Hayashi (1982) showed that if both the production function and the adjustment cost function were homogeneous of degree 1, then the two q's would be equal. (average q = marginal q).

Think about an optimal path of K, N, I and the current value of V. If we had twice the capital now then the new optimal path would involve doubling the N and I at every point in time (this follows from the CRS assumptions). If we do this though, we will inevitably double the value of the firm (twice the revenue, twice the costs ... twice the profits). As a result, the current value of the firm is proportional to current K.

$$PV(K) = \zeta K$$

where ζ is a number.

This tells us that $\frac{\partial V}{\partial K} = \zeta$ and that $\frac{V}{K} = \zeta$ which is by definition Tobin's average q.

To prove Hayashi's theorem more formally, differentiate $q(t) K(t)$ with respect to time:

$$\frac{\partial [q(t) K(t)]}{\partial t} = K(t) \dot{q}(t) + q(t) \dot{K}(t)$$

Recall that the dynamic equation governing $\dot{q}$ is given by:

$$\dot{q} = (r + \delta) q - F_K + \phi\left(\frac{I}{K}\right) + \frac{I}{K}\phi'\left(\frac{I}{K}\right)$$

where we are using our "specialized" adjustment cost function:

$$C(I, K) = K\phi\left(\frac{I}{K}\right)$$

with ϕ an arbitrary convex function (so that dC/dK is $\phi(I/K) - K\phi(I/K)\frac{I}{K^2}$ as above). Plug this into the equation above to get:

$$\frac{\partial [q(t) K(t)]}{\partial t} = K(t)\left[(r + \delta) q - F_K + \phi\left(\frac{I}{K}\right) - \frac{I}{K}\phi'\left(\frac{I}{K}\right)\right] + q(t)\dot{K}(t)$$

[10] This is addapted from the construction in Barro and Sala-i-Martin. See also Basu's notes from last year and of course Hayashi [1982].

Use

$$\dot{K} = I - \delta K$$

and

$$q = 1 + \phi'\left(\frac{I}{K}\right)$$

to get:

$$\begin{aligned}
\frac{\partial\left[q(t)K(t)\right]}{\partial t} &= K(t)\left[(r+\delta)q - F_K + \phi\left(\frac{I}{K}\right) - \frac{I}{K}\phi'\left(\frac{I}{K}\right)\right] + \left[1 + \phi'\left(\frac{I}{K}\right)\right]\left[I - \delta K\right] \\
&= K(r+\delta)q - KF_K + K\phi\left(\frac{I}{K}\right) - I\phi'\left(\frac{I}{K}\right) + I + I\phi'\left(\frac{I}{K}\right) - \delta K - \delta K\phi'\left(\frac{I}{K}\right) \\
&= K(r+\delta)q - KF_K + K\phi\left(\frac{I}{K}\right) + I - \delta K - \delta K\phi'\left(\frac{I}{K}\right)
\end{aligned}$$

because F is CRS in K and N we have:

$$F(K,N) = F_N N + F_K K$$

so that:

$$F_K K = F(K,N) - wN$$

where I have used the fact that $F_N = w$ by profit maximization. Now we have:

$$\frac{\partial\left[q(t)K(t)\right]}{\partial t} = K(r+\delta)q - F(K,N) - wN + K\phi\left(\frac{I}{K}\right) + I - \delta K - \delta K\phi'\left(\frac{I}{K}\right)$$

Recall that the first order condition for I implies that $q = 1 + \phi'\left(\frac{I}{K}\right)$ so that the terms with δK drop out. This gives us:

$$\frac{\partial\left[q(t)K(t)\right]}{\partial t} = Krq - F(K,N) - wN + K\phi\left(\frac{I}{K}\right) + I$$

multiply through by e^{-rt} to get:

$$e^{-rt}\left\{\frac{\partial\left[q(t)K(t)\right]}{\partial t} - rqK\right\} = e^{-rt}\left[-F(K,N) - wN + K\phi\left(\frac{I}{K}\right) + I\right]dt$$

Notice that the derivative of $q(t)K(t)e^{-rt}$ w.r.t. t is $e^{-rt}\frac{\partial[q(t)K(t)]}{\partial t} - e^{-rt}rqK$ which is the left hand side. Now integrate with respect to time to get:

$$\int_0^\infty e^{-rt}\left\{\frac{\partial\left[q(t)K(t)\right]}{\partial t} - rqK\right\}dt = \int_0^\infty e^{-rt}\left[-F(K,N) - wN + K\phi\left(\frac{I}{K}\right) + I\right]dt$$

the right hand side is simply $-PV(0)$. Thus:

$$q(t)\,K(t)\,e^{-rt}\big|_0^\infty = -PV(0)$$

or:

$$\lim_{T\to\infty} e^{-rT} q(T)\,K(T) - q(0)\,K(0) = -PV(0)$$

using the transversality condition we have:

$$q(0)\,K(0) = PV(0)$$

(and in general $q(t)\,K(t) = PV(t)$) so that:

$$q(t) = \frac{PV(t)}{K(t)} = Q_t^{\text{Tobin}}$$

7. Nominal Rigidity and Real Rigidity

7.1. Mankiw (1985)

Firm profit function $\Pi(p)$. The firm's opportunity cost of not adjusting to their optimal price p^* from current price p is:

$$L(p) = \Pi(p^*) - \Pi(p)$$

The firm will adjust prices if $L(p) > z$ (z is the "menu cost"). Approximating this in the neighborhood of p^* gives:

$$L(p) \approx -\frac{1}{2}\Pi''(p^*)(p-p^*)^2$$

This implies that the firm is willing to tolerate a discrepancy $x = (p - p^*)$ of:

$$x = \sqrt{\frac{2z}{-\Pi''}}$$

without adjusting. (note $\Pi'' < 0$).

Notice that the firm is willing to tolerate a first-order difference in its price even if the menu cost is only 2nd order in magnitude.

7.2. Ball and Romer (1990)

Many firms each with profit functions $\Pi\left(\frac{M}{P}, \frac{p_i}{P}\right)$ (here P is the aggregate price, M is the total money supply, and p_i is firm $i's$ nominal price). By assumption $\Pi_1 > 0, \Pi_{22} < 0$, and $\Pi_{12} > 0$.

The optimal choice of p_i requires that $\Pi_2 = 0$. Let $\rho\left(\frac{M}{P}\right)$ be a function giving the optimal $\frac{p_i}{P}$ for any level of aggregate demand $\frac{M}{P}$. Ball and Romer say that the firm faces real rigidity if the derivative $\rho' = \partial\left(\frac{p_i}{P}\right)/\partial\left(\frac{M}{P}\right)$ is small. That is, if changes in aggregate demand do not cause you to change your desired price by too much then there is real rigidity. Differentiating the first order condition $\Pi_2 = 0$ with respect to $\frac{M}{P}$ it is easy to show that:

$$\rho' = \frac{\Pi_{12}}{-\Pi_{22}} > 0$$

If this is small then there is real rigidity in this system.

Assume that $\Pi(1,1)$ is an equilibrium (i.e. the optimal choice of p_i is 1 if $M=1$ and $P=1$). Suppose that all firms set $P=1$ because they expect $M=1$ but that in fact M turns out to differ from one by a small amount.

We look for symmetric Nash equilibria. If no one else adjusts, the opportunity cost to firm i of not adjusting is:

$$\begin{aligned} L(M) &= \Pi\left(M, \frac{p_i^*}{P}\right) - \Pi(M,1) \\ &= \Pi(m, \rho(m)) - \Pi(m,1) \end{aligned}$$

(here $\frac{p_i^*}{P}$ is the optimal choice of p_i^* if firm i chose to adjust). Taking a second order approximation around the steady state gives:

$$\begin{aligned} L(M) &\approx [\Pi_1(m,\rho(m)) + \Pi_2(m,\rho(m))\rho' - \Pi_1(m)](M-1) \\ &\quad + \frac{1}{2}\left[\Pi_{11} + \Pi_{12}\rho' + \Pi_{21}\rho' + \Pi_{22}(\rho')^2 + \Pi_2\rho'' - \Pi_{11}\right](M-1)^2 \\ &= \frac{1}{2}\left[2\Pi_{12}\rho' + \Pi_{22}(\rho')^2\right](M-1)^2 \end{aligned}$$

Let $x = M-1$ then, this firm would adjust (even if no one else adjusted) if:

$$\left[\Pi_{12}\rho' + \frac{1}{2}\Pi_{22}(\rho')^2\right]x^2 \approx L(M) > z$$

Recall that $\rho' = \frac{\Pi_{12}}{-\Pi_{22}}$ this implies that the critical value for x satisfies:

$$\left[-\frac{(\Pi_{12})^2}{\Pi_{22}} + \frac{1}{2}\frac{(\Pi_{12})^2}{\Pi_{22}}\right]x^2 = -\frac{1}{2}\frac{(\Pi_{12})^2}{\Pi_{22}}x^2 = -\frac{1}{2}\rho'\Pi_{12}x^2 = z$$

so that:

$$x^* = \sqrt{\frac{2z}{\Pi_{12}\rho'}}$$

As ρ' gets smaller and smaller, the range of inaction equilibria grows.

What about if everyone else has chosen to adjust? When would firm i go along with the group and adjust? Here the opportunity cost to not adjusting is:

$$L = \Pi(1,1) - \Pi\left(1, \frac{1}{P}\right)$$

and since the new $P = M$ we have:

$$L = \Pi(1,1) - \Pi\left(1, \frac{1}{M}\right)$$

What is the range of inaction here? Again taking a second order approximation gives:

$$L \approx \Pi_2(M-1) - \frac{1}{2}\Pi_{22}(M-1)^2 = \frac{1}{2}\Pi_{22}(M-1)^2$$

If everyone else adjusts, then it is optimal for i to adjust if $L(M) > z$. Again letting $x = (M-1)$ we have the critical value for the adjustment equilibrium given as:

$$x^{**} = \sqrt{\frac{2z}{-\Pi_{22}}}$$

This is the same condition as in Mankiw. There since you are the only firm, you are "everyone". Note that

$$\frac{x^{**}}{x^*} = \frac{\sqrt{\frac{2z}{-\Pi_{22}}}}{\sqrt{\frac{2z}{\Pi_{12}\rho'}}} = \sqrt{\frac{\Pi_{12}\rho'}{-\Pi_{22}}} = \rho'$$

If $\rho' < 1$ (lots of real rigidity) then $x^{**} < x^*$ and there is the possibility for multiple equilibria. Note that high values of Π_{22} implies high real rigidity (ρ' small –> x^* big) but also implies that there is low nominal rigidity (x^{**} small) (a trade-off).

8. Basic Dynamic Sticky Price Models

8.1. A Model with Capital

This treatment closely follows Gilchrist [1999] and Gertler and Gilchrist [1999]. Here we augment the model we gave in class to allow for capital and investment. The introduction of capital has the (undesirable) implication that the real interest rate rises during economic expansions. There are a couple of reasons for this: (1) the close tie between the marginal product of capital and the real interest rate (i.e. without adjustment costs to capital $r_t = MP_{t+1}^K + (1-\delta)$), (2) the marginal product of capital rises when employment rises so to the extent that N stays above trend for several periods the MP^K and consequently the real interest rate, will be high; and finally (3) the anticipated inflation effect – the nominal interest rate is the real interest rate plus expected inflation :$(1+i_t) = (1+r_t)\frac{p_t}{p_{t+1}}$. All three of these factors together imply that (typically) both real and nominal interest rates rise following a monetary expansion.

The households in the model are standard. Consumers maximize:

$$\sum_{t=0}^{\infty} \beta^t \left[\frac{C_t^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} - \phi \frac{N_t^{1+\frac{1}{\eta}}}{1+\frac{1}{\eta}} \right]$$

subject to the nominal budget constraint:

$$P_t C_t + P_t I_t + M_t = W_t N_t + R_t K_t + M_{t-1} + T_t + \Pi_t$$

and the capital accumulation equation:

$$K_{t+1} = K_t(1-\delta) + I_t$$

T_t are transfers from the government (helicopter drops of money) and Π_t are profits returned to the consumer lump sum from firms.

The first order conditions can be reduced to an Euler equation and a labor supply curve. The Euler equation is:

$$\frac{1}{\beta}\left(\frac{C_{t+1}}{C_t}\right)^{\frac{1}{\sigma}} = \left(\frac{R_{t+1}}{P_{t+1}} + (1-\delta)\right)$$

note that the nominal rate of interest is:

$$\begin{aligned} 1+i_t &= \frac{1}{\beta}\left(\frac{C_{t+1}}{C_t}\right)^{\frac{1}{\sigma}} \frac{P_{t+1}}{P_t} = \frac{1}{\beta}\frac{MU(C_t)}{MU(C_{t+1})}\frac{P_{t+1}}{P_t} \\ &= (1+r_t)(1+\pi_t) \end{aligned}$$

Labor supply is standard:

$$\phi n_t^{\frac{1}{\eta}} c_t^{\frac{1}{\sigma}} = \frac{W_t}{P_t^c}$$

I assume that money demand is given by a simple quantity equation:

$$Mv = P_t Y_t$$

To get a more standard "LM" relationship we would have:

$$Mv(1 + i_t) + PY$$

with $v' > 0$ and $v\left(\frac{1}{\beta}\right) = 1$.

Production of final goods is competitive while the production of intermediate goods is monopolistically competitive.

8.2. Final Goods

The Final investment goods follow:

$$Y_t = \left[\int_0^1 y_t(z)^{\frac{\varepsilon-1}{\varepsilon}} dz\right]^{\frac{\varepsilon}{\varepsilon-1}}$$

The demand for intermediate goods is:

$$y_t(z) = \left[\frac{p_t(z)}{P_t}\right]^{-\varepsilon} Y_t$$

This implies the price index:

$$P_t = \left[\int_0^1 p_t(z)^{1-\varepsilon} dz\right]^{\frac{1}{1-\varepsilon}}$$

8.3. Intermediate Goods

Intermediate investment goods are each produced according to:

$$y_t(z) = A\left[k_t(z)\right]^{\alpha}\left[n_t(z)\right]^{1-\alpha}$$

and each wants to minimize nominal costs subject to producing a certain amount:

$$J = -W_t n_t(z) - R_t k_t(z) + MC_t\left[A\left[k_t(z)\right]^{\alpha}\left[n_t(z)\right]^{1-\alpha} - \bar{y}\right]$$

which gives rise to the following first order conditions:

$$\begin{aligned} -W_t + MC_t(1-\alpha)k_t(z)^{\alpha}n_t(z)^{-\alpha} &= 0 \\ -R_t + MC_t\alpha k_t(z)^{\alpha-1}n_t(z)^{1-\alpha} &= 0 \end{aligned}$$

The first equation implies that:

$$MC_t = W_t\frac{1}{1-\alpha}\left(\frac{k_t(z)}{n_t(z)}\right)^{-\alpha}$$

(which says that the marginal cost is related to the marginal product of labor and the wage.) while the second implies:

$$\frac{k_t^c(z)}{n_t^c(z)} = \left(\frac{R_t}{\alpha MC_t}\right)^{\frac{1}{\alpha-1}}$$

which gives marginal cost as:

$$MC_t = W_t^{1-\alpha}R_t^{\alpha}\left(\frac{1}{1-\alpha}\right)^{1-\alpha}\left(\frac{1}{\alpha}\right)^{\alpha}$$

Note that if these firms could set prices in every period then they would simply pick the price that maximizes profits:

$$(p_t^*(z) - MC_t)\left(\left[\frac{p_t^*(z)}{P_t}\right]^{-\varepsilon}Y_t\right) = \frac{Y_t}{P_t^{-\varepsilon}}\left(p_t^*(z) - MC_t\right)\left[p_t^*(z)\right]^{-\varepsilon}$$

This has the first order condition:

$$\begin{aligned} (1-\varepsilon)p^*(z)^{-\varepsilon} + \varepsilon MC_t p_t^*(z)^{-\varepsilon-1} &= 0 \\ (1-\varepsilon) + \varepsilon MC_t p_t^*(z)^{-1} &= 0 \end{aligned}$$

which implies that:

$$p^*(z) = \frac{\varepsilon}{\varepsilon-1}MC = \mu MC$$

Here $\mu > 1$ is the markup (Basu and Fernald estimate μ roughly at 1.1 (a 10% markup)).

8.4. Price Setting

If firms only reset prices infrequently then they must take into account the future ramifications of their actions now.

What do firms want to maximize? If they return an extra dollar of dividends at time $t+j$ the consumer values this according to:

$$\beta^t MU(C_{t+j}) \frac{1}{P_{t+j}}$$

This implies that the firm should choose $p_t^*(z)$ to maximize its expected profits:

$$\max_{p_t^*(z)} \left\{ \sum_{j=0}^{\infty} \theta^j E_t \left[\beta^j \frac{MU(C_{t+j})}{P_{t+j}} \left(p_t^*(z) - MC_{t+j} \right) \left(\left[\frac{p_t^*(z)}{P_{t+j}} \right]^{-\varepsilon} Y_{t+j} \right) \right] \right\} \tag{3}$$

Note that since $1 + i_t = \frac{1}{\beta} \frac{MU(C_t)}{MU(C_{t+1})} \frac{P_{t+1}}{P_t}$ we have:

$$\xi_{t,t+j} = \prod_{s=t}^{t+j-1} (1 + i_s) = \left(\frac{1}{\beta} \right)^j \frac{MU(C_t)}{MU(C_{t+j})} \frac{P_{t+j}}{P_t}$$

so that maximization of (3) is equivalent to maximizing the present discounted value of dividends:

$$\begin{aligned} & \frac{MU(C_t)}{P_t} \sum_{j=0}^{\infty} \theta^j E_t \left[\beta^j \frac{MU(C_{t+j})}{MU(C_t)} \frac{P_t}{P_{t+j}} \left(p_t^*(z) - MC_{t+j} \right) \left(\left[\frac{p_t^*(z)}{P_{t+j}} \right]^{-\varepsilon} Y_{t+j} \right) \right] \\ = & \frac{MU(C_t)}{P_t} \sum_{j=0}^{\infty} \theta^j E_t \left[\frac{\left(p_t^*(z) - MC_{t+j} \right)}{\xi_{t,t+j}} \left(\left[\frac{p_t^*(z)}{P_{t+j}} \right]^{-\varepsilon} Y_{t+j} \right) \right] \end{aligned}$$

Let's rewrite the objective function as revenue minus costs:

$$\Pi_t^e = [p_t^*(z)]^{1-\varepsilon} \sum_{j=0}^{\infty} \theta^j \beta^j E_t \left[\frac{MU(C_{t+j})}{P_{t+j}^{1-\varepsilon}} Y_{t+j} \right] - [p_t^*(z)]^{-\varepsilon} \sum_{j=0}^{\infty} \theta^j \beta^j E_t \left[\frac{MU(C_{t+j})}{P_{t+j}^{1-\varepsilon}} MC_{t+j} Y_{t+j} \right]$$

The first order condition for $p_t^*(z)$ is:

$$(1-\varepsilon) [p_t^*(z)]^{-\varepsilon} \sum_{j=0}^{\infty} \theta^j \beta^j E_t \left[\frac{MU(C_{t+j})}{P_{t+j}^{1-\varepsilon}} Y_{t+j} \right] + \varepsilon [p_t^*(z)]^{-\varepsilon-1} \sum_{j=0}^{\infty} \theta^j \beta^j E_t \left[\frac{MU(C_{t+j})}{P_{t+j}^{1-\varepsilon}} MC_{t+j} Y_{t+j} \right] =$$

solving for $p_t^*(z)$ gives:

$$p_t^*(z) = \left(\frac{\varepsilon}{\varepsilon-1}\right)\frac{\sum_{j=0}^{\infty}\theta^j\beta^j E_t\left[MU(C_{t+j})P_{t+j}^{\varepsilon-1}MC_{t+j}Y_{t+j}\right]}{\sum_{j=0}^{\infty}\theta^j\beta^j E_t\left[MU(C_{t+j})P_{t+j}^{\varepsilon-1}Y_{t+j}\right]}$$

as the optimal choice for prices in the current period. Note that since MC, P, MU, Y are all the same for every z they will all choose the same p^*.

What is the linearization of this:

$$p_t^* = \frac{A_t}{B_t}$$

with

$$A_t = \mu\sum_{j=0}^{\infty}\theta^j\beta^j E_t\left[MU(C_{t+j})P_{t+j}^{\varepsilon-1}MC_{t+j}Y_{t+j}\right]$$

and

$$B_t = \sum_{j=0}^{\infty}\theta^j\beta^j E_t\left[MU(C_{t+j})P_{t+j}^{\varepsilon-1}Y_{t+j}\right]$$

Then, $\tilde{p}_t^* = \tilde{A}_t - \tilde{B}_t$.

$$A_t = \mu MU(C_t)P_t^{\varepsilon-1}MC_tY_t + \theta\beta E_t\left[A_{t+1}\right]$$

so that:

$$\tilde{A}_t = \frac{\mu MU(C)P^{\varepsilon-1}(MC)Y}{A}\left\{\tilde{\lambda}_t + (\varepsilon-1)\tilde{P}_t + \widetilde{MC}_t + \tilde{Y}_t\right\} + \theta\beta E_t\left[\tilde{A}_{t+1}\right]$$

Similarly,

$$\tilde{B}_t = \frac{MU(C)P^{\varepsilon-1}Y}{B}\left\{\tilde{\lambda}_t + (\varepsilon-1)\tilde{P}_t + \tilde{Y}_t\right\} + \theta\beta E_t\left[\tilde{B}_{t+1}\right]$$

In steady state $A = \mu MU(C)P^{\varepsilon-1}(MC)Y\frac{1}{1-\theta\beta}$ and $B = MU(C)P^{\varepsilon-1}Y\frac{1}{1-\theta\beta}$ so that these equations become:

$$\begin{aligned}\tilde{A}_t &= [1-\theta\beta]\left\{\tilde{\lambda}_t + (\varepsilon-1)\tilde{P}_t + \widetilde{MC}_t + \tilde{Y}_t\right\} + \theta\beta E_t\left[\tilde{A}_{t+1}\right]\\ \tilde{B}_t &= [1-\theta\beta]\left\{\tilde{\lambda}_t + (\varepsilon-1)\tilde{P}_t + \tilde{Y}_t\right\} + \theta\beta E_t\left[\tilde{B}_{t+1}\right]\end{aligned}$$

and:

$$\tilde{p}_t^* = [1-\theta\beta]\,\widetilde{MC}_t + \theta\beta E_t\left[\tilde{p}_{t+1}^*\right]$$

If we denote real marginal cost as mc_t then:

$$\tilde{p}_t^* = (1-\theta\beta)\left[\widetilde{mc}_t + \tilde{P}_t\right] + \theta\beta p_{t+1}^*$$

The price index is

$$P_t = \left[\int_0^1 P_t(z)^{1-\varepsilon}dz\right]^{\frac{1}{1-\varepsilon}},$$

and since $(1-\theta)$ agents change prices and switch to the same price p_t^* the price index evolves according to:

$$P_t = \left[\theta P_{t-1}^{1-\varepsilon} + (1-\theta)\left(P_t^*\right)^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}}$$

the linear version of this is:

$$\tilde{P}_t = \theta\tilde{P}_{t-1} + (1-\theta)\tilde{p}_t^*$$

note that:

$$p_t^* = \frac{P_t - \theta P_{t-1}}{1-\theta}$$

then the price setting equation becomes:

$$\frac{P_t - \theta P_{t-1}}{1-\theta} = (1-\theta\beta)\left[\widetilde{mc}_t + \tilde{P}_t\right] + \theta\beta\left[\frac{P_{t+1} - \theta P_t}{1-\theta}\right]$$

Define inflation in period t as $\tilde{\pi}_t = \tilde{P}_t - \tilde{P}_{t-1}$ so that:

$$\begin{aligned}
P_t - \theta P_{t-1} &= (1-\theta)(1-\theta\beta)\left[\widetilde{mc}_t + \tilde{P}_t\right] + \theta\beta\left[P_{t+1} - \theta P_t\right] \\
(1-\theta)P_t + \theta P_t - \theta P_{t-1} &= (1-\theta)(1-\theta\beta)\left[\widetilde{mc}_t + \tilde{P}_t\right] + \theta\beta\left[P_{t+1} - P_t + (1-\theta)P_t\right]
\end{aligned}$$

$$(1-\theta)P_t + \theta\tilde{\pi}_t = (1-\theta)(1-\theta\beta)\left[\widetilde{mc}_t + \tilde{P}_t\right] + \theta\beta\left[\tilde{\pi}_{t+1} + (1-\theta)P_t\right]$$

gives us a dynamic Phillips Curve:

$$\tilde{\pi}_t = \gamma\widetilde{mc}_t + \beta E_t\left[\tilde{\pi}_{t+1}\right]$$

This is basically a short run AS relationship. Suppose $\tilde{\pi}_t > 0$ (inflation is above trend). By assumption, $\tilde{\pi}$ will return to its steady state value of zero so at some point $\tilde{\pi}_{t+1} < \tilde{\pi}_t$. When this occurs (and in fact prior to this) we will have to have $\widetilde{mc} > 0$. This is the same as saying that the markup is falling. As a result, output (and employment) must also expand.

Note that there is no long run trade-off for the monetary authority. Any steady state π is consistent with $\widetilde{mc} = 0$ in the long run (and thus a constant level of employment).

8.5. Other Topics:

8.5.1. Sticky Consumption Prices vs. Flexible Investment Prices

(This is based on Barsky, House, and Kimball [2002]). If non-durable consumption goods have sticky prices while durable goods (capital) have flexible prices then money will be approximately neutral with respect to output and employment *regardless* of how sticky or how large the non-durable consumption goods sector is.

To see this result consider the labor supply decision of the agent. This can be written as:

$$v'(N_t) = \left[P_t^c u'(c_t)\right] W_t$$

The term in brackets is the marginal utility of an additional dollar of income. Alternatively the additional income could be spent on some new investment. This would imply:

$$v'(N_t) = MU(K_t) \frac{W_t}{P_t^i}$$

where $MU(K_t)$ is the marginal utility from getting an additional unit of the investment good.

I claim that this shadow value ($MU(K_t)$) will stay very close to its steady state level. The intuition for this is that $MU(K_t)$ is the present discounted value of all of the future payoffs to this piece of capital.

$$MU(K_t) = \frac{R_{t+1}}{(1+i_t)} \frac{u'(c_{t+1})}{P_{t+1}} + \frac{(1-\delta) R_{t+2}}{(1+i_t)(1+i_{t+1})} \frac{u'(c_{t+2})}{P_{t+2}} + \ldots$$

$$\lambda_t(i) = \sum_{j=1}^{\infty} \frac{u'(c_{t+j})}{P_{t+j}} R_{t+j} (1-\delta)^{j-1} \left\{ \prod_{s=t}^{j-1} (1+i_s) \right\}$$

The steady state value of this is:

$$\begin{aligned}\lambda(i) &= \frac{R}{(1+i)}\frac{u'(c)}{P} + \frac{(1-\delta)\,R\,u'(c)}{(1+i)^2\quad P} + ... \\ &= \sum_{j=1}^{\infty}\left[\prod_{s=t}^{j-1}(1+i_s)^{j-1}\right]^{-1} R(1-\delta)^{j-1} \\ &= \frac{R}{P}\frac{u'(c)\beta}{1-\beta(1-\delta)}\end{aligned}$$

If the shock is resolved quickly (what Miles calls the "fast price adjustment approximation") then this expansion will be dominated by the future terms and will be close to the steady state value. Thus, if $\tilde{K}$ is roughly constant (if capital does not change much) and $MU(K)$ is roughly constant we have:

$$v'(N_t) = MU(K)\frac{W_t}{P_t^i}$$

Note also that since the investment sector is fully flexible $P_t^I = \mu MC_t = \mu W_t \frac{1}{MP_{N,t}}$ so that:

$$v'(N_t) = MU(K)\frac{MP_{N,t}}{\mu}$$

The marginal product of labor is just $(1-\alpha)\left(\frac{K_t}{N_t}\right)^{\alpha}$ and since $\tilde{K}$ is roughly zero we have:

$$v'(N_t) = MU(K)\frac{MP_{N,t}(N_t)}{\mu}$$

The important thing about this is that this condition only depends on N_t. As a result, the change in N_t from its steady state value will only result from large changes in K or large swings in $MU(K)$. If both of these are small then $\tilde{N}_t$ will be small and $\tilde{Y} = \alpha\tilde{K}_t + (1-\alpha)\tilde{N}_t \approx 0$.

8.5.2. Sticky Wages

This treatment follows a standard setup used by Christiano and Eichenbaum (who adapt from some other guy whose name I forget).

The consumer's problem is the same as above. Labor for the firms is an aggregate of labor "types". Assume that for any amount of labor the firm needs many types of workers. Specifically, if L_t is effective labor at time t we have:

$$L_t = \left[\int_0^1 l_{it}^{\frac{\psi-1}{\psi}} di \right]^{\frac{\psi}{\psi-1}}$$

This means that if the firm wants labor force L_t, the demand for type i is given by:

$$l_{it} = L_t \left(\frac{w_{it}}{W_t} \right)^{-\psi}$$

Wages for each type of labor are set by a monopolist in that type (similar to a union). We assume that the aggregate wage is:

$$W_t = \left[\int_0^1 w_{it}^{1-\psi} di \right]^{\frac{1}{1-\psi}}$$

The probability of adjusting a wage is $1-\theta_w$ and the probability of not adjusting (being stuck) is θ_w. An extra dollar in period $t+j$ is worth $\beta^j \frac{MU(C_{t+j})}{P_{t+j}}$ while working more in period $t+j$ costs the consumer the marginal utility of leisure $\beta^j MU(N_{t+j})$ thus the monopolists will try to maximize:

$$\max_{w_{it}^*} \left\{ E_t \left[\sum_{j=0}^{\infty} (\beta\theta)^j \left(\frac{MU(C_{t+j})}{P_{t+j}} w_{it}^* - MU(N_{t+j}) \right) L_{t+j} \left(\frac{w_{it}}{W_{t+j}} \right)^{-\psi} \right] \right\}$$

Note that the labor market clearing condition implies that $\tilde{N}_{t+j} \approx \tilde{L}_{t+j}$.

In a perfectly flexible wage setting environment the monopolist would maximize:

$$\left(\frac{MU(C_t)}{P_t} w_{it}^* + MU(N_t) \right) L_t \left(\frac{w_{it}}{W_t} \right)^{-\psi}$$

or more simply:

$$\frac{MU(C_t)}{P_t} (w_{it}^*)^{1-\psi} + MU(N_t) w_{it}^{-\psi}$$

which gives the first order condition:

$$(\psi - 1) \frac{MU(C_t)}{P_t} = \psi MU(N_t) w_{it}^{-1}$$

or

$$w^* = \frac{\psi}{(\psi - 1)} \frac{-MU(N_t)}{\frac{MU(C_t)}{P_t}}$$

which says that the real wage for this supplier is the competitive wage $\frac{-MU(N)}{MU(C)}$ plus a markup. Assuming the utility structure given above, the optimal choice of w^* will solve the following:

$$\tilde{w}_t^* = (1-\theta_w\beta)\left[\tilde{P}_t + \frac{1}{\eta}\tilde{N}_t + \frac{1}{\sigma}\tilde{C}_t\right] + \theta_w\beta E\left[\tilde{w}_{t+1}^*\right]$$

As before, the aggregate wage will evolve according to:

$$W_t = \left[\theta_w W_{t-1}^{1-\psi} + (1-\theta)\left(w_t^*\right)^{1-\psi}\right]^{\frac{1}{1-\psi}}$$

which is linearized as:

$$\tilde{W}_t = \theta_w\tilde{W}_{t-1} + (1-\theta_w)\tilde{w}_t^*$$

note that:

$$\tilde{w}^* = \frac{\tilde{W}_t - \theta_w\tilde{W}_{t-1}}{1-\theta_w}$$

Defining wage inflation as $\tilde{\pi}_t^w = \tilde{W}_t - \tilde{W}_{t-1}$ gives us a "Phillips Curve" for wage growth:

$$\tilde{\pi}_t^w = \gamma_w\left[\tilde{P}_t + \frac{1}{\eta}\tilde{N}_t + \frac{1}{\sigma}\tilde{C}_t - \tilde{W}_t\right] + \beta E\left[\tilde{\pi}_{t+1}^w\right]$$

with $\gamma_w = \frac{(1-\theta_w)(1-\theta_w\beta)}{\theta_w}$.

8.6. Kimball (1995)

Kimball's [1995] model is like the DNK model with a general output aggregator and with general cost curves. As before, the model satisfies a quantity equation:

$$Mv = PY$$

in every period.

8.6.1. Price Setting:

Kimball distinguishes between "desired price" and "reset price". The desired price ($p^{\#}$) is the price a firm would charge if it were able to costlessly set its price without concern for inflexibilities.

The aggregate price level evolves according to the Calvo aggregator:

$$\tilde{P}_t = \theta\tilde{P}_{t-1} + (1-\theta)\tilde{P}_t^*$$

and the firms seek to maximize:

$$\max_{p_t^*(z)} \left\{ \sum_{j=0}^{\infty} \theta^j E_t \left[\beta^j MU(C_{t+j}) \Pi \left(\frac{p^*(z)}{P_{t+j}}, ... \right) \right] \right\}$$

where Π is the real profit at date $t+j$. The first order condition for this maximization problem is:

$$\sum_{j=0}^{\infty} \theta^j E_t \left[\beta^j MU(C_{t+j}) \Pi' \left(\frac{p^*(z)}{P_{t+j}}, ... \right) \frac{1}{P_{t+j}} \right] = 0$$

The desired price $p_{t+j}^{\#}$ satisfies $\Pi' \left(\frac{p_{t+j}^{\#}}{P_{t+j}}, ... \right) = 0$ this is approximately:

$$0 = \Pi' \left(\frac{p_{t+j}^{\#}}{P_{t+j}}, ... \right) \approx \Pi'(1,.) + \Pi''(1,.) \left[\frac{p_{t+j}^{\#}}{P_{t+j}} - 1 \right]$$

Note that the derivative at the reset price is close to this ...

$$\Pi' \left(\frac{p^*(z)}{P_{t+j}}, ... \right) \approx \Pi'(1,.) + \Pi''(1,.) \left[\frac{p^*(z)}{P_{t+j}} - 1 \right]$$

so that

$$\Pi' \left(\frac{p^*(z)}{P_{t+j}}, ... \right) = \Pi''(1,.) \left[\frac{p^*(z) - p_{t+j}^{\#}}{P_{t+j}} \right]$$

and the first order condition is:

$$\sum_{j=0}^{\infty} \theta^j E_t \left[\beta^j MU(C_{t+j}) \frac{1}{P_{t+j}} \Pi''(1,.) \left(\frac{p^*(z) - p_{t+j}^{\#}}{P_{t+j}} \right) \right] = 0$$

which implies that:

$$p_t^*(z) = p_t^* = \frac{E_t \left[\sum_{j=0}^{\infty} (\theta\beta)^j MU(C_{t+j}) \Pi''(1,.) \frac{p_{t+j}^{\#}}{P_{t+j}^2} \right]}{E_t \left[\sum_{j=0}^{\infty} (\theta\beta)^j MU(C_{t+j}) \Pi''(1,.) \frac{1}{P_{t+j}^2} \right]}$$

This says that the optimal reset price is an (appropriately chosen) average of the desired prices $p_{t+j}^{\#}$. We linearize as before to get:

$$\tilde{p}_t^* = [1 - \theta\beta] \tilde{p}_t^{\#} + \theta\beta E_t \left[\tilde{p}_{t+1}^* \right]$$

8.6.2. Demand and Cost Functions.

Kimball (1995) allows for a variable elasticity of demand and for a more general marginal cost function. In particular, the desired markup is no longer simply $\frac{\varepsilon}{\varepsilon-1}$ rather it solves:

$$\mu\left(\frac{y_t^{\#}(z)}{Y_t}\right)=\frac{\varepsilon\left(y_t^{\#}(z)/Y_t\right)}{\varepsilon(y_t^{\#}(z)/Y_t)-1}$$

so that by assumption it depends on the market share of the i^{th} firm. Locally, (in the neighborhood of the steady state) we still have the demand curve:

$$y_t(z)=Y_t\left(\frac{p_t(z)}{P_t}\right)^{-\varepsilon^*}$$

where ε^* is the elasticity of demand at the steady state.

Marginal cost is also more general. Specifically, real marginal cost is given by the function:

$$mc_t(z)=\Phi(y_t(z),Y_t,...)$$

so that marginal cost depends on the firms individual output as well as aggregate output. Note that we assume that marginal cost rises with a balanced expansion of all firms. Note that:

$$\widetilde{mc}_t(z)=\frac{y\Phi_y}{\Phi}\tilde{y}_t(z)+\frac{Y\Phi_Y}{\Phi}\tilde{Y}_t$$

and if $\tilde{y}_t(z)=\tilde{Y}_t$ we want marginal cost to rise. Thus we require:

$$=\frac{y\Phi_y+Y\Phi_Y}{\Phi}>0$$

Note that the desired price $p_t^{\#}(z)/P_t$ will solve:

$$\left(p_t^{\#}(z)/P_t\right)=\mu\left(y_t^{\#}(z)/Y_t\right)\Phi(y_t^{\#}(z),Y_t,...)$$

and at full employment (the flexible price level of output Y^f) we will have:

$$1=\mu(1)\Phi(Y^f,Y^f)$$

Dividing these two equations gives us:

$$\left(p_t^{\#}(z)/P_t\right)=\frac{\mu\left(y_t^{\#}(z)/Y_t\right)\Phi(y_t^{\#}(z),Y_t,...)}{\mu(1)\Phi(Y^f,Y^f)}=f(y_t^{\#}(z),Y_t,Y_t^f,...)$$

and linearizing this around the steady state gives:

$$p_t^{\#}(z) - P_t = \left[\frac{\mu'}{\mu} + \frac{y\Phi_y}{\Phi}\right] y_t^{\#} + \left[\frac{Y\Phi_Y}{\Phi} - \frac{\mu'}{\mu}\right] \tilde{Y}_t - \left[\frac{y\Phi_y + Y\Phi_Y}{\Phi}\right] \tilde{Y}_t^f$$

$$p_t^{\#}(z) - P_t = \left[\frac{\mu'}{\mu} + \frac{y\Phi_y}{\Phi}\right] \left(y_t^{\#}(z) - \tilde{Y}_t\right) + \quad \left(\tilde{Y}_t - \tilde{Y}_t^f\right)$$

Since the demand curve solves:

$$y_t^{\#}(z) = Y_t \left(\frac{p_t^{\#}(z)}{P_t}\right)^{-\varepsilon}$$

$$y_t^{\#}(z) - Y_t = -\varepsilon^* \left(p_t^{\#}(z) - P_t\right)$$

so that:

$$-\frac{1}{\varepsilon^*}\left(y_t^{\#}(z) - Y_t\right) = \left[\frac{\mu'}{\mu} + \frac{y\Phi_y}{\Phi}\right] \left(y_t^{\#}(z) - \tilde{Y}_t\right) + \quad \left(\tilde{Y}_t - \tilde{Y}_t^f\right)$$

$$0 = \omega\left(y_t^{\#}(z) - \tilde{Y}_t\right) + \quad \left(\tilde{Y}_t - \tilde{Y}_t^f\right)$$

where $\omega = \left[\frac{\mu'}{\mu} + \frac{y\Phi_y}{\Phi} + \frac{1}{\varepsilon^*}\right]$ (note in Kimball he writes $\omega(y/Y)$ so that

$$\omega\left(y/Y\right) = \frac{(y/Y)\,\mu'(y/Y)}{\mu\left(y/Y\right)} + \frac{y\Phi_y}{\Phi} + \frac{1}{\varepsilon\left(y/Y\right)}$$

solving for desired output as a function of actual output gives:

$$y_t^{\#}(z) - \tilde{Y}_t = -\frac{}{\omega}\left(\tilde{Y}_t - \tilde{Y}_t^f\right)$$

This with the firms demand curve gives:

$$\left(p_t^{\#}(z) - P_t\right) = \frac{}{\varepsilon^*\omega}\left(\tilde{Y}_t - \tilde{Y}_t^f\right)$$

Using Kimball's terminology, the "inflationary price gap" $p_t^{\#} - P_t$ is equal to $\frac{}{\varepsilon^*\omega}$ times the "inflationary output gap $\tilde{Y}_t - \tilde{Y}_t^f$.

8.6.3. New Phillips Curve:

Recall that the optimal reset price is:

$$\tilde{p}_t^* = [1 - \theta\beta]\, \tilde{p}_t^{\#} + \theta\beta E_t \left[\tilde{p}_{t+1}^*\right]$$

and substitute in for $p^{\#}$ to get:

$$\tilde{p}_t^* = [1 - \theta\beta] \left(P_t + \frac{}{\varepsilon^* \omega} \left[\tilde{Y}_t - \tilde{Y}_t^f\right] \right) + \theta\beta E_t \left[\tilde{p}_{t+1}^*\right]$$

As a result, the forward looking Phillips Curve is:

$$\tilde{\pi}_t = \gamma \frac{}{\varepsilon^* \omega} \left[\tilde{Y}_t - \tilde{Y}_t^f\right] + \beta E_t \left[\tilde{\pi}_{t+1}\right]$$

The output gap will "cause" inflation. Note that as $\frac{}{\varepsilon^* \omega}$ becomes smaller and smaller, inflation responds less and less to output (alternatively: even small amounts of inflation will necessitate very large changes in GDP). The term $\frac{}{\varepsilon^* \omega}$ is exactly the "real rigidity" concept used in Ball and Romer.

Low values of or high values of ω will generate low $\frac{}{\varepsilon^* \omega}$ and thus high degrees of real rigidity. is the change in marginal costs due to a balanced expansion in output. If wages, or the real rental price of capital rise sharply then will be large and there won't be much real rigidity. If wages don't rise much, or if there are productive (or trading) externalities in aggregate output then will be lower and there will be real rigidity.

ω is given by :

$$\left[\frac{\mu'}{\mu} + \frac{y\Phi_y}{\Phi} + \frac{1}{\varepsilon^*}\right]$$

Near the steady state, the slope of the demand curve is $\frac{1}{\varepsilon}$ (i.e., in logs we would have $\ln p_t = \frac{1}{\varepsilon} \ln Y_t - \frac{1}{\varepsilon} \ln y_t + \ln P_t$). The slope of the marginal revenue curve is steeper and is given by $\frac{\mu'}{\mu} + \frac{1}{\varepsilon^*}$. Thus if $\frac{\mu'}{\mu} + \frac{1}{\varepsilon^*}$ is large, MR falls very fast. At the firm level marginal costs rise at the rate $\frac{y\Phi_y}{\Phi}$. So if MC rises sharply and MR falls quickly, the individual firm will not have incentive to change its price (output).

So flat aggregate MC, steep individual MC and steep individual MR will all contribute to real rigidity. (A good environment for price rigidity: productivity spillovers & efficiency wages contribute to flat aggregate MC (low); convex / kinked demand curves cause MR to drop fast; Kimball suggests having workers tied to firms (e.g. due to firm specific skills) will cause firms to have sharp MC curves).

8.7. Appendix:

8.7.1. Output.

Note that output, in the basic New Keynesian model, is given by:

$$Y_t = \left[\int_0^1 y_t(z)^{\frac{\varepsilon-1}{\varepsilon}} dz\right]^{\frac{\varepsilon}{\varepsilon-1}}$$

since some of the $y(z)$'s are different Y_t will not simply be Cobb-Douglas. To see this note:

$$Y_t = \left[\int_0^1 \left\{A_t \left[k_t(z)\right]^{\alpha} \left[n_t(z)\right]^{1-\alpha}\right\}^{\frac{\varepsilon-1}{\varepsilon}} dz\right]^{\frac{\varepsilon}{\varepsilon-1}}$$

since the firms all minimize costs they all have the same capital-labor ratio:

$$Y_t = A_t \left[\frac{K_t}{N_t}\right]^{\alpha} \left[\int_0^1 n_t(z)^{\frac{\varepsilon-1}{\varepsilon}} dz\right]^{\frac{\varepsilon}{\varepsilon-1}} \neq A_t \left[\frac{K_t}{N_t}\right]^{\alpha} N_t$$

since $\left[\int_0^1 n_t(z)^{\frac{\varepsilon-1}{\varepsilon}} dz\right]^{\frac{\varepsilon}{\varepsilon-1}} \neq \int_0^1 n_t(z)dz = N_t$. The demand for intermediate z is given by:

$$A_t \left[k_t(z)\right]^{\alpha} \left[n_t(z)\right]^{1-\alpha} = y(z) = Y_t \left(\frac{p_t(z)}{P_t}\right)^{-\varepsilon}$$

Since the capital / labor ratios are constant across producers we can write this as:

$$A_t \left[\frac{K_t}{N_t}\right]^{\alpha} n_t(z) = y(z) = Y_t \left(\frac{p_t(z)}{P_t}\right)^{-\varepsilon}$$

Integrating with respect to z gives:

$$A_t \left[\frac{K_t}{N_t}\right]^{\alpha} \int_0^1 n_t(z)dz = Y_t P_t^{\varepsilon} \int_0^1 p_t(z)^{-\varepsilon} dz$$

$$A_t K_t^{\alpha} N_t^{1-\alpha} = Y_t P_t^{\varepsilon} \int_0^1 p_t(z)^{-\varepsilon} dz$$

or:

$$Y_t = A_t K_t^{\alpha} N_t^{1-\alpha} X_t$$

where X_t is given by:

$$X_t = \frac{P_t^{-\varepsilon}}{Z_t} = \frac{P_t^{-\varepsilon}}{\int_0^1 p_t(j)^{-\varepsilon} d(j)}$$

$$\tilde{X}_t = -\varepsilon \tilde{P}_t - \tilde{Z}_t$$

The definition of the optimal price p^* and the linearizations of P and Z are given by:

$$\tilde{p}_t^* = (1-\beta\theta)\left[MC_t\right] + E_t\left[\beta\theta p_{t+1}^*\right]$$
$$\tilde{P}_t = \theta \tilde{P}_{t-1} + (1-\theta)\tilde{p}_t^*$$
$$\tilde{Z}_t = \theta \tilde{Z}_{t-1} - \varepsilon(1-\theta)\tilde{p}_t^*$$

These terms drop:

$$\begin{aligned}
\tilde{X}_t &= -\varepsilon \tilde{P}_t - \tilde{Z}_t = -\varepsilon\left[\theta \tilde{P}_{t-1} + (1-\theta)\tilde{p}_t^*\right] - \left[\theta \tilde{Z}_{t-1} - \varepsilon(1-\theta)\tilde{p}_t^*\right] \\
&= -\varepsilon\theta \tilde{P}_{t-1} - \varepsilon(1-\theta)\tilde{p}_t^* - \theta \tilde{Z}_{t-1} + \varepsilon(1-\theta)\tilde{p}_t^* \\
&= -\varepsilon\theta \tilde{P}_{t-1} - \theta \tilde{Z}_{t-1} = \theta \tilde{X}_{t-1}
\end{aligned}$$

Thus X is (dynamically) independent of the other equations in the system. This implies that $\tilde{X}_t = 0$ in the system. Thus production can be approximated by a Cobb-Douglas relationship,

$$Y_t = \tilde{A}_t + \alpha \tilde{K}_t + (1-\alpha)\tilde{N}_t$$

8.7.2. The relation between L and N?

A similar problem applies to the sticky wage model. In that model, total labor is given by:

$$L_t = \left[\int_0^1 l_{it}^{\frac{\psi-1}{\psi}} di\right]^{\frac{\psi}{\psi-1}}$$

while N_t is simply the sum of the individual labor supplies:

$$N_t = \int_0^1 l_{it} di$$

Note: consider the demand function for one type of labor:

$$l_{it} = L_t \left(\frac{w_{it}}{W_t}\right)^{-\psi}$$

and integrate with respect to i. This gives:

$$N_t = L_t W_t^{\psi} \int_0^1 w_{it}^{-\psi} di$$

so that

$$L_t = N_t \frac{W_t^{-\psi}}{\int_0^1 w_{it}^{-\psi} di} = N_t X_t$$

As before with output, the first order changes in X_t are negligible so we can simply write:

$$\tilde{L}_t = \tilde{N}_t$$

To see this note:

$$\tilde{X}_t = -\psi \tilde{W}_t - \tilde{Z}_t$$

where $Z_t = \int_0^1 w_{it}^{-\psi} di$. Alternatively we can write:

$$Z_t = \left[\theta_w Z_{t-1} + (1 - \theta_w) (w_t^*)^{-\psi} \right]$$

The definition of the optimal reset wage $\tilde{w}_t^*$ and the linearizations of W and Z are given by:

$$\tilde{w}_t^* = (1 - \theta_w \beta) \left[\tilde{P}_t + \frac{1}{\eta} \tilde{N}_t + \frac{1}{\sigma} \tilde{C}_t \right] + \theta_w \beta E \left[\tilde{w}_{t+1}^* \right]$$

$$\tilde{W}_t = \theta_w \tilde{W}_{t-1} + (1 - \theta_w) \tilde{w}_t^*$$
$$\tilde{Z}_t = \theta_w \tilde{Z}_{t-1} - \psi (1 - \theta_w) \tilde{w}_t^*$$

Then, as with output, terms other than X will drop. So again, X is dynamically independent of the other equations in the system. Consequently we can write $\tilde{L}_t = \tilde{N}_t$.

As with the mixing problem with the intermediate goods, because the "Calvo mechanism" draws agents at random, the inefficiency caused by having a "wrong mix" of labor types is to a first order approximation negligible.