

ECON 6140
Macroeconomics Notes

Gabe Sekeres

Spring 2025

Contents

1	Julieta Caunedo	2
1.1	Introduction	2
1.2	Growth Model	3
1.3	Dynamics in the One-Sector Growth Model	5
1.4	Equilibria	9
1.5	Overlapping Generations Model	13
1.6	Heterogeneity	18
1.7	Continuous Time Growth	25
1.8	Endogenous Growth Theory	29
2	Kristoffer Nimark	38
2.1	Introduction	38
2.2	A Classical Monetary Model	40
2.3	Basic New Keynesian Model I	42
2.4	Basic New Keynesian Model II	45
2.5	Rational Expectations Models	47
2.6	Monetary Policy	50
2.7	Policy Tradeoffs	52
2.8	Sticky Wages	54
2.9	Unemployment	58
2.10	State Space Models (and the Kalman Filter)	61
2.11	Calibration and Moment Matching	67
2.12	Likelihood-Based Estimation	68

1 Julieta Caunedo

1.1 Introduction

Administration Homeworks are due at the beginning of the lecture, and no late homework will be accepted. However, only $N - 1$ will be evaluated. Please feel free to work in groups, but turn in your own write-up. The midterm (like last semester, the final for this portion of the course) will be on Tuesday, March 11. Julieta will look for a room so we can take a full two hours. This course will occasionally refer to the notes by [Dirk Krueger](#), as well as the same textbooks as last semester, [Recursive Methods in Macroeconomic Dynamics](#) (SLP) and [Recursive Macroeconomic Theory](#) (LS).

This first week will be somewhat review, but the pace will pick up quickly. You need to do the homeworks. Homeworks will be heavy, but that's because the way you study is by doing them. Remember: at this point in your career, the grades are completely useless. They function as a progress marker, but in and of themselves they don't matter at all. They don't make you a good economist.

Motivation Why are people in the US today much richer than they were in 1800? Why are Germany and France much richer than Argentina and Kenya? Does growth generate inequality? What is the role of frictions in hindering growth?

Macroeconomists aim at answering these questions building *quantitative* models – models that can be contested with empirical facts. “The weight of evidence for an extraordinary claim must be proportioned to its strangeness” – Laplace, 1812. We will think about abstractions, yes, but *useful* (read: empirically testable) abstractions.

The main questions in this course are:

1. Why are some countries richer than others?
2. Why do some countries grow faster than others?
3. What is the effect of wealth and consumption inequality for cross-country differences in income per capita?
4. What is the role of financial frictions?
5. What is the role of education for long-run growth?
6. How do firms innovate, and what's their impact on economic growth?
7. Why do some firms operate older technologies while better ones are available?

Broadly, we will consider (in order):

1. One-sector growth model (Ramsey-Cass-Koopmans)
 - (a) Computation
 - (b) Extensions to multiple sectors
2. Competitive equilibrium
 - (a) Heterogeneity and aggregation
 - (b) The income fluctuations problem, incomplete markets
3. Overlapping generations
 - (a) Recursive representation

- (b) Dynamic inefficiencies
- 4. Long-run growth
 - (a) Human capital
 - (b) AK model
 - (c) Endogenous growth: externalities and innovation

Macroeconomics, more than anything, is a course on the *long-term*. It's not necessarily about monetary policy, or trade, or anything specific, but more the study of behavior over a long time horizon. It's the only field able to study topics like climate change and demographic change for that reason.

1.2 Growth Model

“Who does what, when” – Sargent, on models. Consider: Time, Preferences, Technology.

Model. *One-sector Growth Model (redux)*

- Time: Discrete, infinite horizon
- Preferences: representative dynasty preferences

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

where $\beta \in (0, 1)$ and u strictly concave, increasing, and differentiable

- Technology:

$$c_t + x_t + g_t \leq f(k_t)$$

$$k_{t+1} \leq x_t + (1 - \delta)k_t$$

where x_t is investment, g_t is government spending, and $f(k_t)$ is output. $f(k_t)$ is concave, strictly increasing, and $f(0) = 0$. We also have the following two conditions: $\lim_{k \rightarrow 0} f'(k_t) > \frac{1}{\beta} - (1 - \delta)$, and $\lim_{k \rightarrow \infty} f'(k_t) < \frac{1}{\beta} - (1 - \delta)$

Question. Why is this called a one-sector growth model? It looks like there are two sectors!

Because the marginal rate of transformation between consumption and next-period capital (investment) is one-to-one.

Assumption 1.1. *All quantities must be non-negative.*

We have the following maximization problem, given a sequence of $\{g_t\}$ and a k_0 :

$$\max_{\{c_t, x_t, k_{t+1}\}} \sum_{t=0}^{\infty} \beta^t u(c_t) \text{ subject to}$$

$$c_t + x_t + g_t \leq f(k_t)$$

$$k_{t+1} \leq x_t + (1 - \delta)k_t$$

$$(c_t, x_t, k_{t+1}) \geq (0, 0, 0)$$

Definition. An *allocation* is a set of sequences $\{c_t, x_t, k_{t+1}\}$. An allocation is *feasible* if it satisfies

$$c_t + x_t + g_t \leq f(k_t)$$

$$k_{t+1} \leq x_t + (1 - \delta)k_t$$

$$(c_t, x_t, k_{t+1}) \geq (0, 0, 0)$$

How will we solve this? We consider it as a finite horizon problem, and then derive (and apply) a transversality condition. We will solve the finite problem by an application of the KKT conditions, where the Lagrangian is

$$\mathcal{L}(c, x, k, \lambda, \theta, \gamma) = \sum_{t=0}^T \beta^t \{u(c_t) + \lambda_t[f(k_t) - (c_t + x_t + g_t)] + \theta_t[x_t + (1 - \delta)k_t - k_{t+1}] + \dots\}$$

where \dots are the non-negativity conditions with γ multipliers (trivially will hold).

The first-order necessary conditions are:

$$\begin{aligned} c_t : u'(c_t) - \lambda_t + \gamma_{1t} &= 0 & t = 0, 1, \dots, T \\ x_t : -\lambda_t + \theta_t + \gamma_{2t} &= 0 & t = 0, 1, \dots, T \\ k_{t+1} : -\theta_t + \gamma_{3t} + \beta\lambda_{t+1}f'(k_{t+1}) + \beta(1 - \delta)\theta_{t+1} &= 0 & t = 0, 1, \dots, T \\ k_{T+1} : -\theta_T + \gamma_{3T} &= 0 \end{aligned}$$

as well as complimentary slackness conditions. The region we think about the finite horizon is that this last period is different – you’re going to die, so there is no payoff for accumulating capital. We will use the complimentary slackness condition for time $t = T$ to construct the transversality condition. We need the agent to not want to accrue capital forever, for some payoff long in the future. The complimentary slackness condition is that $\beta^T k_{T+1} \gamma_{3T} = 0$. We can convert this to $\beta^T \theta_T k_{T+1}$, and the transversality condition will hold as long as $\lim_{T \rightarrow \infty} \beta^T \theta_T k_{T+1} = 0$.

Remark. Most importantly!! Never repeat this approach outside of this course.

Our first order necessary conditions are now:

$$\begin{aligned} c_t : u'(c_t) &= \lambda_t \\ x_t : \theta_t &= \lambda_t \\ k_{t+1} : \beta\lambda_{t+1}[f'(k_{t+1}) + (1 - \delta)] &= \lambda_t \\ \text{TC} : \lim_{T \rightarrow \infty} \beta^T \theta_T k_{T+1} &= 0 \\ f : c_t + x_t + g_t &\leq f(k_t) \\ k_t : k_{t+1} &\leq x_t + (1 - \delta)k_t \end{aligned}$$

Rearranging, we get

$$\begin{aligned} k_{t+1} : \beta u'(c_{t+1})[f'(k_{t+1}) + (1 - \delta)] &= u'(c_t) \\ \text{TC} : \lim_{T \rightarrow \infty} \beta^T \theta_T k_{T+1} &= 0 \end{aligned}$$

From Micro 101, we can recognize that

$$\begin{aligned} \frac{u'(c_t)}{\beta u'(c_{t+1})} &= [f'(k_{t+1}) + (1 - \delta)] \\ \text{MRS} &= \text{MRT} \end{aligned}$$

Now suppose that $g_t = g \forall t$. We have the following:

Definition. A *steady state* is an allocation such that, for all t , $c_t = c$, $x_t = x$, and $k_{t+1} = k$. At the steady

state, our first order necessary conditions become

$$\begin{aligned} c_t : u'(c^*) &= \lambda^* \\ k_{t+1} : \beta[f'(k^*) + (1 - \delta)] &= 1 \\ f : c^* + x^* + g &= f(k^*) \\ k_t : \delta k^* &= x^* \end{aligned}$$

Simplifying, we have that

$$\begin{aligned} \beta[f'(k^*) + (1 - \delta)] &= 1 & (\text{Euler}) \\ c^* + \delta k^* + g &= f(k^*) & (\text{Feasibility}) \end{aligned}$$

Essentially, the level of capital is fully determined by the production technology. To know consumption, we just need the feasible consumption rate.

Question. Does this steady state necessarily exist? If so, is it unique?

Solution. Define $L(k) = \beta[f'(k) + (1 - \delta)]$. We know from the assumptions (way above) that $\lim_{k \rightarrow 0} L(k) > 1$ and $\lim_{k \rightarrow \infty} L(k) < 1$. So what else do we need? Just continuity! Which we have from our assumption that this was a constant returns to scale production function in capital and labor, so decreasing in labor, so f' is decreasing! Conclusion follows from the Intermediate Value Theorem – we have both existence and uniqueness.

Are we done yet? We need to check feasibility! Is $c^* > 0$? Suppose that $g = 0$. We need to check $f'(k) < \delta$ at the point \hat{k} such that $f(\hat{k}) = \delta \hat{k}$, which implies that $\hat{c} = 0$. However, for all $k < \hat{k}$, $f(k) > \delta k$. It remains to show that $k^* < \hat{k}$, which follows directly because:

$$f'(k^*) = \frac{1}{\beta} - (1 - \delta) > \delta$$

for $\beta \in (0, 1)$.

1.3 Dynamics in the One-Sector Growth Model

Suppose now that we have a technology level z so that output is $zf(k)$, which is analogous to total factor productivity. Our first order conditions are now

$$\begin{aligned} c_t : u'(c_t) &= \lambda_t \\ k_{t+1} : \beta \lambda_{t+1} [zf'(k_{t+1}) + (1 - \delta)] &= \lambda_t \\ \text{TC} : \lim_{t \rightarrow \infty} \beta^T \lambda_T k_{T+1} &= 0 \\ f : c_t + x_t + g_t &\leq zf(k_t) \\ k_t : k_{t+1} &\leq x_t + (1 - \delta)k_t \end{aligned}$$

Remark. One could describe the dynamics in terms of consumption, which is what most books do, or in terms of shadow value, which is what we will do.

We define $u'(c(\lambda)) \equiv \lambda$ and will assume $g_t = g$ sufficiently ‘small’¹ for now. Our key conditions are, as before

$$\begin{aligned} \beta \lambda_{t+1} [zf'(k_{t+1}) + (1 - \delta)] &= \lambda_t & (\text{Euler}) \\ c_t + k_{t+1} - (1 - \delta)k_t + g_t &\leq zf(k_t) & (\text{Feasibility}) \end{aligned}$$

¹Basically, that we will not run into the non-negativity constraint on consumption.

We can solve this using a phase diagram, which we saw last semester. We can't explicitly solve an *infinite* system of non-linear equations explicitly, but we can do it using programs – specifically, for this [QuantEcon](#) has a good codebase.

In this case, we will use finite difference methods – specifically, the shooting algorithm. We will approximate k and c (or λ , in our case) with N discrete points in the time dimension. Denote the distance between grid points as Δt . Let (k_n, c_n) be a point on the grid, and use the equations that characterize the optimum with k_0 given:

$$\begin{aligned} c_{n+1} &= \beta c_n [z f'(f(k_n) - (1 - \delta)k_n - c_n) + (1 - \delta)] \\ k_{n+1} &= f(k_n) - (1 - \delta)k_n - c_n \end{aligned}$$

Remark. Why is this linear in c ? We are explicitly assuming log utility here, that's how we convert it. Any CARA utility would have an exponent, but we assume log for simplicity.

Explicitly, we have:

Algorithm. *Shooting Method*

1. Guess c_0
2. Obtain (c_n, k_n) for $n = 1, \dots, N$ by running the equations above forward
3. If the sequence converges to (c^*, k^*) then you have the correct saddle path. If not, update c_0 and go back to 1

This will converge because the solution is unique – there is explicitly only one steady state, and since the problem is convex there's a unique path to that steady state.

Now take the sequential planner's problem, which is

$$\max_{\{c_t, k_{t+1}\}} \sum_{t=0}^{\infty} \beta^t u(c_t)$$

subject to

$$\begin{aligned} c_t + k_{t+1} &= f(k_t) + (1 - \delta)k_t \\ c_t, k_{t+1} &\geq 0 \text{ and } k_0 > 0 \text{ given} \end{aligned}$$

We have the value function

$$V(k_0) = \max_{\{(c_t, k_{t+1}) \in \Gamma(k_t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(c_t)$$

where

$$\Gamma(k_t) = \{(c_t, k_{t+1}) : c_t \geq 0, k_{t+1} \geq 0, c_t + k_{t+1} \leq f(k_t) + (1 - \delta)k_t\}$$

Our recursive formulation is

$$V(k_0) = \max_{(c_0, k_1) \in \Gamma(k_0)} u(c_0) + \beta V(k_1)$$

In general,

$$V(k) = \max_{(c, k') \in \Gamma(k)} u(c) + \beta V(k')$$

where k is the state variable (or set of state variables). As long as u is differentiable, the solution to this problem satisfies

$$k' = f(k) + (1 - \delta)k - c \quad \text{and} \quad \frac{\partial u(c)}{\partial c} = \beta \frac{\partial V(k')}{\partial k}$$

From the envelope condition, we can recover the Euler equation:

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

As we saw last semester, the solution to this problem is a policy function $g : K \rightarrow K$ such that $k' = g(k)$ and a value function $V^* : K \rightarrow \mathbb{R}_+$. We will approximate g and V^* numerically, with methods that differ depending on the characteristics of the problem. In general, our algorithm will be either value function iteration, or finding policy functions from the Euler equations. Both of these *should* give us the same solutions.

Numerical Solutions. We have a set of non-linear equations to be solved over the state space K (these are our Euler equations):

$$\begin{aligned} \frac{\partial u(c)}{\partial c} &= \beta \frac{\partial u(c')}{\partial c} \left[z \frac{\partial f(k')}{\partial k} + (1 - \delta) \right] \\ \implies \frac{\partial u(f(k) + (1 - \delta)k - k')}{\partial c} &= \beta u'(f(k') + (1 - \delta)k' - k'') \left[\frac{\partial f(k')}{\partial k} + (1 - \delta) \right] \end{aligned}$$

We will discretize K into N nodes, and solve the system of equations: (and then make sure $k = k^*$ belongs to the set K)

$$\frac{\partial u(f(k) + (1 - \delta)k - g(k))}{\partial c} = \beta u'(f(g(k)) + (1 - \delta)g(k) - g(g(k))) \left[\frac{\partial f(g(k))}{\partial k} + (1 - \delta) \right]$$

Algorithm. *Value Function Iteration*

1. Discretize the state space K
2. Make a guess for the value function $V_0(k)$
3. Use the optimization algorithm to solve for $k'(k)$
4. Compute²

$$V_1(k) = u(f(k) + k(1 - \delta) - k'(k)) + \beta V_0(k'(k))$$

5. Check the distance between $V_0(k)$ and $V_1(k)$. If larger than the tolerance, update $V_0(k) = V_1(k)$ and go back to 3. Otherwise, stop!

Remark. With full depreciation ($\delta = 1$) and log utility, you can (and we did!) do this by hand!

Example. Suppose we want to know the answer to questions of the following form:

1. What is the role of the capital output ratio...
2. What is the role of productivity...

for differences in output per capita across countries?

We will choose functional forms for technology and preferences where parameters have clear economic interpretations, *i.e.*

$$u(c_t) = \frac{c_t^{1-\sigma} - 1}{1 - \sigma} \quad \text{and} \quad f(k) = zk_t^\alpha$$

We will set $z = 1$ in the US for now, and have four free parameters: σ , δ , β , and α . The literature suggests two approaches: estimation and calibration.

²There are various methods to approximate $V_0(k')$ for ‘off-grid’ points.

Question. Why not always estimate? The model is an abstraction in which we've deliberately abstracted away from some features. Standard formal statistical procedures for estimation use criteria that may not make economic sense. This was a big discussion in the 80s and 90s.

Alternative: choose the aspects of the data that your model was designed to capture.

The key idea of calibration is that choosing parameters boils down to choosing moments in the data to match. You can think of calibration as an exactly identified generalized method of moments (GMM) estimation.

This model is designed to explain the capital accumulation process. The key statistics that we care about are k/y , x/y , and r . Think of the modern US as fluctuations around the steady state (*i.e.* take averages to calculate the steady state). If we take one period as one year, we get that $k/y \approx 2.5$, $x/y \approx 0.2$, and $r \approx 0.04$.

Remark. Note that we have three moments, but four parameters! That's because σ does not affect the steady state, so it cannot be identified from the steady state. Estimates in the data have that $\sigma \in [1, 2.5]$. We will use $\sigma = 1$, corresponds to log utility.

Our steady state conditions give us that

$$\frac{x}{y} = \delta \frac{k}{y} \implies \delta \approx 0.08 \quad ; \quad r = \alpha \frac{y}{k} \implies \alpha \approx 0.1 \quad ; \quad \beta[r + (1 - \delta)] = 1 \implies \beta \approx 1$$

How much of the observed differences in income per capita are accounted for by differences in z , or by differences in k/y ? Recall that $Y_t = zK_t^\alpha H_t^{1-\alpha}$. Taking logs, we get

$$\ln(Y_t) = \underbrace{\alpha}_{\beta_k} \ln(K_t) + \underbrace{(1-\alpha)}_{\beta_H} \ln(H_t) + \underbrace{\ln(A_t)}_{\varepsilon}$$

Remark. We could estimate this as a regression directly... but orthogonality between TFP and the covariates is extremely unlikely. This condition will not be satisfied. Julieta gave a [full lecture](#) on this, if you are interested.

An important take-away from the one sector growth model: differences in TFP induce differences in K , but $\frac{K}{Y}$ is independent from TFP in steady state! We have from the Euler equation that

$$1 = \text{discount} \left[\underbrace{\alpha \frac{Y}{K}}_{MPK} + (1 - \text{depreciation}) \right]$$

So output per worker is

$$\frac{Y}{L} = Z \left(\frac{K}{Y} \right)^{\frac{\alpha}{1-\alpha}} \frac{H}{L} \quad \text{where } Z = A^{\frac{1}{1-\alpha}}$$

in logs:

$$\ln \left(\frac{Y}{L} \right) = \ln(Z) + \frac{\alpha}{1-\alpha} \ln \left(\frac{K}{Y} \right) + \ln \left(\frac{H}{L} \right)$$

If you want to play with these data to see the relationships yourself, Julieta has a [Github repo](#) with some broad growth data and code.

Remark. In accounting exercises, we very often get that TFP explains most of the difference in output per capita. That's really unsatisfying – it's basically a residual. However, macroeconomists are constantly looking under the hood of TFP to find the actual paths for that difference.

1.4 Equilibria

1.4.1 Competitive Equilibria

Question. How many markets are open?

In the most simple case, all markets are open, but this is unwieldy and not so realistic – recall that agents live forever in many models. Instead, we will look at a (recursive) market structure.

Model. *Competitive Equilibrium* Firms rent inputs in spot markets, and specifically do not own capital. They solve

$$\max_{\{k_t, n_t\}} c_t + p_{k_t} x_t - r_t k_t - w_t n_t$$

subject to

$$c_t + x_t \leq F(k_t, n_t) \quad t = 0, 1, \dots$$

and a bunch of non-negativity constraints.

Assumption 1.2. *F is strictly increasing in each argument, concave, twice continuously differentiable, and homogeneous of degree 1.*

This means we have constant returns to scale for production, meaning that the size of the firm does not matter. In an interior solution, $p_{k_t} = 1$, and $r_t = F_k(k_t, n_t)$ and $w_t = F_n(k_t, n_t)$. Let $f(k) = F(k, 1)$. Equilibrium factor prices satisfy $r_t = f'(k_t)$ and $w_t = f(k_t) - k_t f'(k_t)$. Firms choose the level of employment.

Households solve the problem

$$\max_{\{c_t, x_t, k_{t+1}, b_{t+1}\}} \sum_{t=0}^{\infty} \beta^t u(c_t)$$

subject to

$$\begin{aligned} c_t + p_{k_t} x_t + b_{t+1} &\leq w_t + r_t k_t + R_t b_t & t = 0, 1, \dots \\ k_{t+1} &\leq x_t + k_t(1 - \delta) & t = 0, 1, \dots \\ 0 &= \lim_{T \rightarrow \infty} \beta^T u(c_T) b_{T+1} \\ 0 &\leq c_t, x_t, k_{t+1} & t = 0, 1, \dots \end{aligned}$$

Why do we need a no-ponzi-scheme condition on b_t ? We don't want agents carrying debt forward forever, in order to maximize consumption.

Remark. This is *not* a transversality condition – the transversality condition is an optimality condition, this is a restriction on the problem structure itself. As we can see from real life – [Ponzi](#), [Madoff](#), etc. – this can happen, and would be optimal for maximizing consumption. We could alternatively impose $b_t \geq \underline{b}$ for some $\underline{b} \in (-\infty, 0)$.

The Lagrangian of this problem is:

$$\begin{aligned} \mathcal{L} = \sum_{t=0}^{\infty} \beta^t \Big\{ & u(c_t) + \\ & + \lambda_t [w_t + r_t k_t + R_t b_t - c_t - p_{k_t} x_t - b_{t+1}] \\ & + \theta_t [(1 - \delta)k_t + x_t - k_{t+1}] \\ & + \gamma_{1t} c_t + \gamma_{2t} x_t + \gamma_{3t} k_{t+1} \Big\} \end{aligned}$$

The first order conditions are similar to before, but have some new conditions:

$$\begin{aligned}
c_t : u'(c_t) &= \lambda_t \\
x_t : \lambda_t &= \theta_t \\
k_{t+1} : \theta_t &= \beta[\theta_{t+1}(1 - \delta) + \lambda_{t+1}r_{t+1}] \\
b_{t+1} : \lambda_t &= \beta R_{t+1}\lambda_{t+1} \\
TV C_k : \lim_{T \rightarrow \infty} \beta^T \lambda_T k_{T+1} &= 0 \\
TV C_b : \lim_{T \rightarrow \infty} \beta^T \lambda_T b_{T+1} &= 0
\end{aligned}$$

We end up with one main Euler equation:

$$u'(c_t) = \beta u'(c_{t+1})[1 - \delta + r_{t+1}]$$

We can also use the FOC on the choice of capital to obtain

$$u'(c_t) = \beta u'(c_{t+1})[1 - \delta + f'(k_{t+1})]$$

The Euler equation for bonds is

$$u'(c_t) = \beta u'(c_{t+1})R_{t+1}$$

Note that one of these two (capital and bonds) Euler equations implies the other!

Definition. A *recursive competitive equilibrium (RCE)* is a collection of price sequences $\{w_t^*, r_t^*, p_{k_t}^*, R_t^*\}$, an allocation $\{c_t^*, x_t^*, k_{t+1}^*\}$, and a sequence of bond holdings $\{b_{t+1}\}$ such that

1. Given prices, the allocation and sequence of bonds is utility maximizing.
2. Given prices, the allocation is profit maximizing
3. Markets clear
4. $b_0^* = b_0 = 0$, $k_0^* = k_0 > 0$ is given.

Proposition 1.1. Any budget feasible allocation must satisfy

$$\sum_{t=0}^{\infty} q_t c_t = \sum_{t=0}^{\infty} q_t w_t + q_0[1 - \delta + r_0]k_0 + q_0 R_0 b_0$$

where $q_t/q_0 = \prod_{j=1}^t R_j^{-1}$ for $t \geq 1$.

Remark. q_t/q_0 is the price of consumption at time t , in terms of consumption at time 0. These are the prices in an Arrow-Debreu economy with time zero trading.

Proof. Our budget constraint at time $t = 0$ is

$$b_1 = w_0 + r_0 k_0 - c_0 - x_0$$

while at time $t = 1$ it is

$$c_1 + x_1 + b_2 = w_1 + r_1 k_1 + R_1 b_1$$

We can combine these, and get

$$c_0 + x_0 + R_1^{-1}(c_1 + x_1) = w_0 + r_0 k_0 + R_1^{-1}(w_1 + r_1 k_1) + R_0 b_0 - R_1^{-1} b_2$$

Repeating, we get that for any finite T ,

$$\sum_{t=0}^T q_t(c_t + x_t) = \sum_{t=0}^T q_t(w_t + r_t k_t) + q_0 R_0 b_0 - q_T b_{T+1}$$

No Ponzi implies that we can take T to ∞ , and use the Euler equation for bonds to get

$$\sum_{t=0}^{\infty} q_t(c_t + x_t) = \sum_{t=0}^{\infty} q_t(w_t + r_t k_t) + q_0 R_0 b_0$$

We want to show that

$$\sum_{t=0}^{\infty} q_t(r_t k_t - x_t) = \sum_{t=0}^{\infty} q_t(r_t k_t - k_{t+1} + (1 - \delta)k_t) = q_0[1 - \delta + r_0]k_0$$

We can rearrange the terms of the second sum to read

$$\lim_{T \rightarrow \infty} q_0[1 - \delta + r_0]k_0 + k_1[-q_0 + q_1(1 - \delta + r_1)] + \cdots + k_T[-q_{T-1} + q_T(1 - \delta + r_T)]$$

And conclusion follows from the Transversality Condition □

Proposition 1.2. *In an equilibrium, $b_t^* = 0$ for all t*

Proof. From the household budget constraint,

$$w_t^* + r_t^* k_t^* + R_t^* b_t^* = c_t^* + x_t^* + b_{t+1}^*$$

and using the equilibrium conditions for w_t, r_t we have that

$$f(k_t^*) - k_t^* f'(k_t^*) + R_t^* b_t^* = c_t^* + x_t^* + b_{t+1}^*$$

And conclusion follows from noting that market clearing implies that $f(k_t^*) = c_t^* + x_t^*$, so the budget constraint for the private sector becomes

$$R_t^* + b_t^* = b_{t+1}^*$$

and conclusion follows from $b_0^* = 0$. □

1.4.2 Taxes

Question. How do deficits affect the economy?

Theorem 1.1. Ricardian Equivalence

1. *Budget policy does not matter – for any given sequence $\{g_t\}$ all non-distortionary tax structures that raise the appropriate level of revenue are associated to the same real allocation*
2. *Timing of the tax collection is irrelevant, in the sense that prices and allocations are independent of the timing. This result is often described as the Ricardian proposition*

Proof. Let $g = \{g_t\}$ be a sequence of expenditure. Define $\phi = \{\tau_t\}$ the sequence of lump-sum taxes to finance g and any initial debt. Consider an alternative sequence of taxes $\hat{\phi}$ such that $\hat{\tau}_0 < \tau_0$. Then

$$\sum_{t=0}^{\infty} q_t c_t = \sum_{t=0}^{\infty} q_t (w_t - \tau_t) + q_0[1 - \delta + r_0]k_0 + q_0 R_0 b_0$$

and assume that $R_t b_t + g_t = \tau_t + b_{t+1}$. This implies that

$$\sum_{t=0}^{\infty} q_t (\tau_t - g_t) = q_0 R_0 b_0 - q_T b_{T+1}$$

Impose a No-Ponzi condition, so $\lim_{T \rightarrow \infty} q_T b_{T+1} = 0$. Then $\sum_{t=0}^{\infty} q_t \tau_t = \sum_{t=0}^{\infty} q_t \hat{\tau}_t$, and

$$\sum_{t=0}^{\infty} q_t c_t = \sum_{t=0}^{\infty} q_t (w_t - g_t) + q_0 [1 - \delta + r_0] k_0$$

Guess that prices did not change after the tax change. Given that the budget constraint of the household is the same, it chose the same investment and consumption sequences, and firms chose the same allocations to maximize profits as before. Market clearing requires that

$$c_t + k_{t+1} = f(k_t) + (1 - \delta)k_t - g_t$$

which is the same since g_t did not change! □

Question. Why does the higher deficit not generate shifts in the interest rate?

For $\hat{\tau}_0 < \tau_0$,

$$b_1 + \tau_0 = g_0 + R_0 b_0 = \hat{b}_1 + \hat{\tau}_0 \implies \hat{b}_1 - b_1 = \hat{\tau}_0 - \tau_0$$

So debt goes up! However, households understand that lower taxes today will be compensated with higher taxes tomorrow. Households increase saving, at exactly the same rate as the government will increase taxes in the future. Since everyone is perfectly forward-looking, they do not change their habits at all.

Model. *Distortionary Taxes* We have a representative household, which solves

$$\max_{\{c_t, n_t, x_t, k_{t+1}, b_{t+1}\}} \sum_{t=0}^{\infty} \beta^t u(c_t, 1 - n_t)$$

subject to, for $t = 0, 1, 2, \dots$

$$\begin{aligned} (1 + \tau_t^c) c_t + (1 + \tau_t^x) x_t + b_{t+1} &\leq (1 + \tau_t^n) w_t n_t + r_t k_t - \tau_t^k (r_t - \delta_t) k_t + (1 + (1 - \tau_t^b r_t^b) b_t) \\ k_{t+1} &\leq x_t + k_t (1 - \delta) \\ 0 &= \lim_{T \rightarrow \infty} \beta^T u'(c_T) b_{T+1} \\ 0 &\leq c_t, x_t, k_{t+1} \end{aligned}$$

The representative firm's problem is

$$\max_{k_t, n_t} F(k_t, n_t) - r_t k_t - w_t n_t$$

The first order conditions for the household are

$$\begin{aligned} u'_c(c_t, 1 - n_t) &= \lambda_t (1 + \tau_t^c) & (c_t) \\ (1 + \tau_t^x) \lambda_t &= \beta \lambda_{t+1} [(1 - \delta_k)(1 + \tau_{t+1}^x) + (1 - \tau_{t+1}^k) r_{t+1} + \tau_{t+1}^k \delta_k] & (k_{t+1}) \\ \lambda_t &= \beta \lambda_{t+1} [1 + (1 - \tau_{t+1}^b) r_{t+1}^b] & (b_{t+1}) \\ u'_n(c_t, 1 - n_t) &= \lambda_t (1 - \tau_t^w) w_t & (n_t) \\ 0 &= \lim_{T \rightarrow \infty} \beta^T \lambda_T k_{T+1} & (\text{TVC}) \\ 0 &= \lim_{T \rightarrow \infty} \beta^T \lambda_T b_{T+1} & (\text{No Ponzi}) \end{aligned}$$

And the first order conditions for the firm are

$$F_k(k_t, n_t) = r_t \quad ; \quad F_n(k_t, n_t) = w_t$$

So in steady state, we have

$$\rho + \tau^x(\rho + \delta_k) = (1 - \tau^k)(F_k(k, n) - \delta_k) \quad (k)$$

$$1 = \beta[1 + \rho] \quad (b)$$

$$u'_n(c, 1 - n) = u'_c(c, 1 - n) \frac{1 - \tau^w}{1 + \tau^c} F_n(k, n) \quad (n)$$

$$F(k, n) = c + g + \delta_k k \quad (\text{Feasibility})$$

where $\rho \equiv (1 - \tau^b)r^b$. Thus, we have that taxes on bonds just change r^b . Similarly, consumption and income taxes are equivalent, and investment and capital taxes are equivalent. This leads to two extremely famous results:

Theorem 1.2. First Welfare Theorem *If $[(w^*, r^*, p_k^*, R^*), (c^*, x^*, k^*), b^*]$ is an interior competitive equilibrium, then (c^*, x^*, k^*) solves the planner's problem.*

Proof. Since $b^* = 0$ in any equilibrium, the household budget constraint is

$$f(k^*) = c^* + x^* \quad ; \quad (k^*)' = (1 - \delta)k^* + x^*$$

where $(k^*)'$ is a vector of capital stock with first element k_1^* . Also, the optimality conditions for the firm and the household together imply that

$$\begin{aligned} u'(c_t) &= \beta u'(c_{t+1})[(1 - \delta) + f'(k_{t+1})] \\ 0 &= \lim_{T \rightarrow \infty} \beta^T u'(c_T) k_{T+1} \end{aligned}$$

□

Theorem 1.3. Second Welfare Theorem *If (c^*, x^*, k^*) is an interior solution to the planner's problem, then there exist prices (w^*, r^*, p_k^*, R^*) and a sequence of bond holdings b^* such that $[(w^*, r^*, p_k^*, R^*), (c^*, x^*, k^*), b^*]$ is a competitive equilibrium of an economy with a representative agent who has initial wealth $b_0 = 0$ and initial capital holdings $k_0 > 0$ equal to the initial endowment of capital in the economy.*

Proof. Let $b^* = 0$. Since the solution to the planner's problem is feasible, we need to define prices and ensure that the optimality conditions (including transversality conditions) are met. Define

$$\begin{aligned} w^* &= f(k^*) - f'(k^*)k^* \\ r^* &= f'(k^*) \\ R^* &= (1 - \delta) + f'(k^*) \end{aligned}$$

At these prices, the optimality conditions of the firms and households are satisfied. Finally, we check the transversality conditions. The condition on bonds is trivially met since $b^* = 0$. If we set $\lambda_t = u'(c_t)$, the transversality condition of the consumer problem in capital is the same as the one the planner faces. This will be a competitive equilibrium. □

1.5 Overlapping Generations Model

There are three models you should know to read macroeconomic papers: growth models (like we've been doing), overlapping generations models, and partial equilibrium models. This will be a primer on OLG, which is the second major workhorse of modern macroeconomics. Why do we use this model? Well, individuals

don't actually live forever. In what sense is a model where people live for finite time equivalent to infinite lifecycles model? We want models (that macroeconomists call altruistic bequest motive models) where agents have interesting life-cycles: born, education, labor income, plan for retirement, partner, have children, retire, and die. This integration of micro data into macro models is very modern.

Model. *Overlapping Generations (OLG)* Time is discrete ($t = 1, 2, \dots$) and the economy lasts forever. There is a single non-storable consumption good in each period. A new generation is born each period, indexed by the year born. People live for two periods and then die.³ What happens to the population?

Generation t is endowed with the consumption good in periods 1 and 2 of life, (e_t^t, e_{t+1}^t) . Generation t 's consumption in those periods is (c_t^t, c_{t+1}^t) . At each point in time t , there are two generations alive: an old generation with endowment and consumption (e_t^{t-1}, c_t^{t-1}) and a young generation with endowment and consumption (e_t^t, c_t^t) . At time 0, there is one old generation (e_1^0, c_1^0) . We also have exponential population growth where $L_0 = 1$ and $L_t = (1 + n)^t L_0$.

We assume that the only endowment is labor (time) – one unit supplied inelastically when young in return for w_t (i.e. you retire in year 2). We have constant returns to supply (CRS) technology for production $Y_t = F(K_t, L_t)$ and competitive factor markets. We will assume that $\delta = 1$, define $k \equiv \frac{K}{L}$, $f(k) \equiv F(k, 1)$, and the gross return on savings and wage rates

$$1 + r_t = R_t = f'(k_t) \quad \text{and} \quad w_t = f(k_t) - k_t f'(k_t)$$

Each generation solves the problem

$$\max_{c_t^t, c_{t+1}^t, s_t} u(c_t^t) + \beta u(c_{t+1}^t)$$

subject to

$$c_t^t + s_t \leq w_t \quad \text{and} \quad c_{t+1}^t \leq R_{t+1} s_t$$

So old agents rent their savings to firms as capital, and as long as u is strictly increasing with Inada, the constraints hold with equality. We need no non-negativity constraints – why?

The Euler equation is

$$u'(c_t^t) = \beta R_{t+1} u'(c_{t+1}^t)$$

and since the individual problem is concave, this suffices. We obtain a savings function $s : \mathbb{R}^2 \rightarrow \mathbb{R}$ with $s_t = s(w_t, R_{t+1})$, where s is increasing in w and increasing or decreasing in R . Aggregate savings are $S_t = L_t s_t$, and with full depreciation capital stock is $K_{t+1} = L_t s(w_t, R_{t+1})$.

Definition. An *OLG competitive equilibrium* is a sequence of aggregate capital stocks, individual consumption, and factor prices $\{K_t, (c_t^t, c_{t+1}^t), R_t, w_t\}_{t=0}^\infty$ such that the factor price sequence satisfies the first two conditions, individual consumption decisions are given by the Euler equation, and the savings function, and the aggregate capital stocks follow the law of motion.

Steady state is defined as normal, holding $k \equiv \frac{K}{L}$ constant, and the equilibrium characterization requires normalizing by the size of the population $L_{t+1} = (1 + n)L_t$. We have that

$$k_{t+1} = \frac{K_{t+1}}{L_{t+1}} = \frac{L_t}{L_{t+1}} s(w_t, R_{t+1})$$

and combining with the above, we get

$$k_{t+1} = \frac{s(f(k_t) - k_t f'(k_t), f'(k_{t+1}))}{n + 1}$$

Steady state is a fixed point of this expression. Note that since s can take any form, in principle multiple steady states (as well as complicated dynamics) are possible. If we assume CRRA utility and Cobb-Douglas

³Alternatively, death can be stochastic. See Blanchard 1985.

production:

$$U_t = \frac{(c_t^t)^{1-\sigma} - 1}{1-\sigma} + \beta \frac{(c_{t+1}^t)^{1-\sigma} - 1}{1-\sigma} \quad ; \quad f(k) = k^\alpha$$

for $\theta > 0$, $\alpha, \beta \in (0, 1)$. The key outcome is the Euler equation

$$\frac{c_{t+1}^t}{c_t^t} = (\beta R_{t+1})^{\frac{1}{\theta}}$$

Rewritten in terms of savings rate, we have

$$s_t^\theta \beta R_{t+1}^{1-\theta} = (w_t - s_t)^{-\theta} \implies s_t = \frac{w_t}{\phi_{t+1}} \quad \text{where } \phi_{t+1} \equiv \left[1 + \beta^{-\frac{1}{\theta}} R_{t+1}^{-\frac{1-\theta}{\theta}}\right] > 1 \implies s_t < w_t$$

We have comparative statics for savings with respect to wages

$$s_w = \frac{\partial s}{\partial w} = \frac{1}{\phi_{t+1}} \in (0, 1)$$

and with respect to capital

$$s_R = \frac{\partial s}{\partial R} = \frac{1-\theta}{\theta} (\beta R_{t+1})^{-\frac{1}{\theta}} \frac{s_t}{\phi_{t+1}}$$

where the sign depends on θ , the intertemporal elasticity of substitution. By combining the above, we can define a steady state implicitly by

$$k^* = \frac{f(k^*) - k^* f'(k^*)}{(1+n) \left[1 + \beta^{-\frac{1}{\theta}} f'(k^*)^{-\frac{1-\theta}{\theta}}\right]}$$

We could also solve for the interest rate $R^* = \alpha(k^*)^{\alpha-1}$, and get

$$(1+n) \left[1 + \beta^{-\frac{1}{\theta}} (R^*)^{-\frac{1-\theta}{\theta}}\right] = \frac{1-\alpha}{\alpha} R^*$$

and finally, dynamics are given by the difference equation

$$k_{t+1} = \frac{(1-\alpha)k_t^\alpha}{(1+n) \left[1 + \beta^{-\frac{1}{\theta}} (\alpha k_{t+1}^{\alpha-1})^{-\frac{1-\theta}{\theta}}\right]}$$

Proposition 1.3. *In the OLG model with generations that live for two periods, Cobb-Douglas technology, and CRRA preferences, there exists a unique and stable steady state for all $k(0) > 0$.*

Remark. In this specific well-behaved case, equilibrium dynamics \approx the Solow Model. Even with CRRA and Cobb-Douglas the model can get quite messy – the canonical model assumes log preferences.

The social planner solves

$$\sum_{t=0}^{\infty} \beta_s^t U_t \equiv \sum_{t=0}^{\infty} \beta_s^t (u(c_t^t) + \beta u(c_{t+1}^t))$$

subject to

$$F(K_t, L_t) = K_{t+1} + c_t^t L_t + c_t^{t-1} L_{t-1}$$

where β_s is the planner's discount factor across generations. Dividing by L_t , we get

$$f(k_t) = (1+n)k_{t+1} + c_t^t + \frac{c_t^{t-1}}{1+n}$$

We have the Lagrangian

$$\mathcal{L} = \sum_{t=0}^{\infty} \beta_s^t \left[(u(c_t^t) + \beta u(c_{t+1}^t)) + \lambda_t \left(f(k_t) - (1+n)k_{t+1} + c_t^t + \frac{c_{t+1}^{t-1}}{1+n} \right) \right]$$

The Euler equation comes from the first order necessary conditions:

$$u'(c_t^t) = \beta f'(k_{t+1}) u'(c_{t+1}^t)$$

Since $f'(k_{t+1}) = R_{t+1}$, the intertemporal consumption decisions are identical to the household, meaning that there is no distortion in consumption allocation over time when shifting to the social planner's problem. However, there may be differences across generations – the social planner may weight them differently. The steady state gives us

$$f'(k^*) - (1+n)k^* = \overbrace{c_1^* + \frac{c_2^*}{1+n}}^{c^*}$$

where c_1 and c_2 are consumption when young and old respectively. We maximize overall consumption when

$$\frac{\partial c^*}{\partial k^*} = f'(k^*) - (1+n) = 0 \implies \exists k^{\text{gold}} \text{ s.t. } f'(k^{\text{gold}}) = (1+n)$$

so k^{gold} is the steady state capital that maximizes consumption (*golden rule capital*). Importantly, since f is concave, if $k^* > k^{\text{gold}}$, then $f'(k^*) < 1+n \implies \frac{\partial c^*}{\partial k^*} < 0$, so lower savings would increase consumption for everyone. This is depicted in Figure 1.

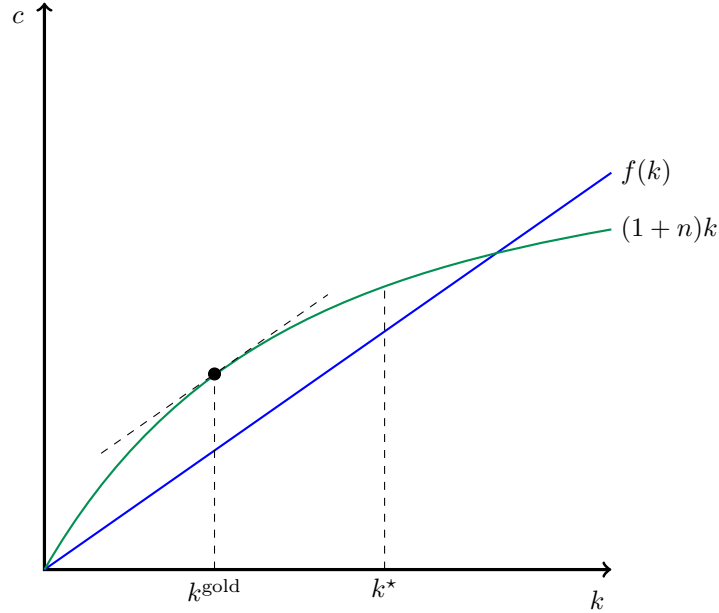


Figure 1: Golden Rule Capital

Definition. The economy is *dynamically inefficient* if it involves overaccumulation (*i.e.* if $k^* > k^{\text{gold}}$). An alternative to this condition is $R^* < 1+n \iff r^* < n$. Transversality in a standard one-sector growth model requires $r > g+n$, but we do not impose transversality in the OLG model, since agents live for two periods and solve finite problems.

Remark. Individuals born at time t face prices determined by the stock of capital chosen by the previous

generation. Pecuniary externality: actions of the previous generation affect those further on. These typically do not matter for welfare (because they are second order), but these affect an infinite stream of newborn agents. These pecuniary externalities can be exploited, and we will see this in the application.

Proposition 1.4. *In the baseline OLD, the competitive equilibrium is not necessarily Pareto optimal. Whenever $r^* < n$ the economy is dynamically inefficient. Hence, it is possible to reduce the capital stock in the steady state and increase consumption for all generations.*

Proof. Consider changing next period's capital stock so that $\Delta k < 0$, and then move towards the steady state. This will lead to lower savings in the first period, so $\Delta c_T = (1+n)\Delta k > 0$, meaning that everyone is happier. Further, since $k^* > k^{\text{gold}}$, for small Δk ,

$$\Delta c_t = -(f'(k^* - \Delta k) - (1+n))\Delta k \implies f'(k^* - \Delta k) - (1+n) < 0 \implies \Delta c_t > 0 \forall t > T$$

Thus, we've improved consumption in all future periods, showing a Pareto improvement. \square

Remark. We can think about two types of systems to deal with dynamic inefficiencies:

1. Fully funded (social security): Young make contributions to the social security system, which are paid back when they are old
2. Unfunded (pay-as-you-go): Transfers go directly from the young to the current old

Pay-as-you-go discourages savings, so it may lead to a Pareto improvement.

Example. Fully-Funded Social Security The household's problem is

$$\max_{c_t^t, c_{t+1}^t, s_t, d_t} u(c_t^t) + \beta u(c_{t+1}^t)$$

subject to

$$c_t^t + s_t + d_t \leq w_t \quad \text{and} \quad c_{t+1}^t \leq R_{t+1}(s_t + d_t)$$

The government raises d_t from the young, invests it in the capital stock, and pays it back when they are old. Market clearing for capital requires $s_t + d_t = (1+n)k_{t+1}$, but the household does not necessarily choose $s_t > 0$. If s_t is unconstrained, then given a (feasible) sequence $\{d_t\}_{t=0}^\infty$, the set of competitive equilibria without social security is the set of competitive equilibria with social security if $s_t > 0$. If you impose $s_t \geq 0$ (no borrowing), then a (feasible) sequence $\{d_t\}_{t=0}^\infty$ is a competitive equilibrium if the equilibrium savings is such that $s_t > 0$ for all t . This implies that there cannot be a Pareto improvement if we impose $s_t \geq 0$.

Example. Unfunded Social Security The household's problem is

$$\max_{c_t^t, c_{t+1}^t, s_t} u(c_t^t) + \beta u(c_{t+1}^t)$$

subject to

$$c_t^t + s_t + d_t \leq w_t \quad \text{and} \quad c_{t+1}^t \leq R_{t+1}s_t + (1+n)d_{t+1}$$

The government raises d_t from the young and distributed $(1+n)d_t$ to the old. The rate of return on social security is $1+n$ rather than R_{t+1} , and only s_t goes to capital accumulation.

Remark. Unfunded social security reduces capital accumulation! What effect will that have on growth and welfare? If the economy is dynamically inefficient, this may be good... however most poorer countries have capital accumulation that is too low rather than too high. However, social security transfers resources from future generations to the initial old generation. With no dynamic inefficiency, this will make some future generation worse off!

1.6 Heterogeneity

We will talk about heterogeneity and consumption distribution, Gorman aggregation, and variance of consumption. We will typically take income as exogenous, and consider differences in consumption as the objects of interest with respect to the heterogeneity.

Piketty talked in capital about how capital growth, when wages are stagnant, can lead to increases in inequality. Specifically, recent automation is a large concern here. This is a recent question – in older history, growth was associated with sharp declines in inequality, rather than increases. Weirdly, inequality in the 20th century seemed to decline until the 1960s, when inequality in the U.S. increased by massive amounts.

Question. Is the one-sector growth model consistent with some degree of heterogeneity across households?

Under some conditions, we can show that heterogeneity in initial wealth, effective labor (*i.e.* human capital), and (limited) differences in utility do not affect equilibrium. Why? Because the ‘averages’ remain the same across the economy, which is identical to the representative agent economy. The basic ideas go back to Gorman on aggregation.

Model. *Heterogeneity* We have N households, each characterized by a vector (θ_i, a_i, e_i) where a_i are assets and e_i endowments of labor, and individual utility

$$u_i(c) = \frac{(c + \theta_i)^\eta}{1 - \eta}$$

where $\eta > 0$ and θ_i can be either positive or negative. The household problem is

$$\max_{\{c_{i,t}, a_{i,t+1}\}} \sum_{t=0}^{\infty} \beta^t u_i(c_{i,t})$$

subject to

$$c_{i,t} + a_{i,t+1} \leq w_t e_i + R_t a_{i,t} \quad \forall t = 0, 1, \dots \quad \text{and} \quad \lim_{T \rightarrow \infty} \beta^T u'_i(c_{i,T}) a_{i,T+1} = u$$

with $a_{i,0}$ given. We also have a present value budget constraint, where

$$\sum_{t=0}^{\infty} q_t c_{i,t} \leq \sum_{t=0}^{\infty} q_t w_t e_i + q_0 a_{i,0} \quad \text{where} \quad q_t \equiv \prod_{j=1}^t R_j^{-1}$$

The representative firm solves the standard problem

$$\max_{c_t, x_t} c_t + p_{kt} x_t - w_t e_t - r_t k_t$$

subject to

$$c_t + x_t = F(k_t, e_t) \quad \text{and} \quad k_{t+1} = (1 - \delta)k_t + x_t$$

Let N be the number of households in the economy. For any variable z_{it} , let $z_t = N^{-1} \sum_{i \leq N} z_{i,t}$, meaning the population average. We have first and second moments (including covariance) as usual, and we define $\theta \equiv N^{-1} \sum_{i=1}^N \theta_i$, $a_t \equiv N^{-1} \sum_{i=1}^N a_{it}$, and $e \equiv N^{-1} \sum_{i=1}^N e_i$.

Definition. A competitive equilibrium is a collection of price sequences $\{q_t, w_t, r_t, R_t\}$, an allocation $\{x_t, k_t, \{c_{it}\}\}$ and a sequence of asset holdings $\{a_{it}\}$ such that

1. Given prices, the allocation and sequence of assets maximizes utility
2. Given the equilibrium prices, the allocation maximizes profits

3. The allocation is feasible: market clearing and aggregate law of motion for capital

4. $a_0 = k_0 > 0$ given

Proposition 1.5. *Average quantities corresponding to a competitive equilibrium also solve the planner's problem:*

$$\max_{\{c_t, x_t, k_{t+1}\}} \sum_{t=0}^{\infty} \beta^t \frac{(c_t + \theta)^{1-\eta}}{1-\eta}$$

subject to $c_t + x_t \leq F(k_t, e_t)$ and $k_{t+1} = (1 - \delta)k_t + x_t$.

Proof. Assume an interior solution. The first order condition for the planner's problem is just

$$(c_t + \theta)^{-\eta} = \beta(c_{t+1} + \theta)^{-\eta}[1 - \delta + F_k(k_{t+1}, e_{t+1})]$$

The Euler equation for family i is

$$(c_{it} + \theta_i)^{-\eta} = \beta(c_{it+1} + \theta_i)^{-\eta}R_{t+1}$$

but $R_{t+1} = [1 - \delta + F_k(k_{t+1}, e_{t+1})]$, which implies that

$$(c_{it} + \theta_i) = \beta^{-1/\eta}(c_{it+1} + \theta_i)[1 - \delta + F_k(k_{t+1}, e_{t+1})]^{-1/\eta}$$

Averaging each side, we get

$$N^{-1} \sum_{i=1}^N (c_{it} + \theta_i) = N^{-1} \sum_{i=1}^N \beta^{-1/\eta} (c_{it+1} + \theta_i) [1 - \delta + F_k(k_{t+1}, e_{t+1})]^{-1/\eta}$$

□

Question. What are the implications in this model for consumption mobility and cross-sectional dispersion of consumption?

If we write the FOC of the household with the present-value budget constraint, we get

$$(c_{it} + \theta_i)^{-\eta} = \lambda_i \frac{q_t}{\beta^t} \implies c_{it} + \theta_i = \left(\lambda_i \frac{q_t}{\beta^t} \right)^{-1/\eta}$$

We can use this expression in the budget constraint to solve for λ_i :

$$\lambda_i^{-1/\eta} \sum_{t=0}^{\infty} q_t \left(\frac{q_t}{\beta^t} \right)^{-1/\eta} = e_i \sum_{t=0}^{\infty} q_t w_t + \theta_i \sum_{t=0}^{\infty} q_t + q_0 a_{i0}$$

For any sequence z_t , let $v(z, q) = \sum_{t=0}^{\infty} q_t z_t$. Then we have

$$\lambda_i^{-1/\eta} v \left(\left(\frac{q_t}{\beta^t} \right)^{-1/\eta}, q \right) = e_i v(w, q) + \theta_i v(1, q) + q_0 a_{i0}$$

Let $\Phi_t = \frac{q_t}{\beta^t}$ and $\Phi_0 = 1$. Then we have

$$\lambda_i^{-1/\eta} = \frac{e_i v(w, q) + \theta_i v(1, q) + a_{i0}}{v(\Phi_t^{-1/\eta}, q)}$$

Using this multiplier in the FOC of the household, we get

$$c_{it} = \frac{e_i v(w, q) + \theta_i v(1, q) + a_{i0}}{v(\Phi_t^{-1/\eta}, q)} \Phi_t^{-1/\eta} - \theta_i$$

Note that consumption for each household is a linear function of e_i, θ_i, a_{i0} . Let M_i be such that

$$M_i \equiv m_e e_i + m_\theta \theta_i + m_a a_{i0}$$

where $m_e = \frac{v(w, q)}{v(\Phi_t^{-1/\eta}, q)}$, $m_\theta = \frac{v(1, q)}{v(\Phi_t^{-1/\eta}, q)}$, and $m_a = \frac{1}{v(\Phi_t^{-1/\eta}, q)}$. Also, let $M = m_e e + m_\theta \theta + m_a a_0$. Then we have that

$$c_{it} = M_i \Phi_t^{-1/\eta} - \theta_i \quad \text{and} \quad c_t = M \Phi_t^{-1/\eta} - \theta$$

Remark. Key features here: all individuals face the same prices, aggregate demand functions are independent of the distribution of the personal characteristics, and a necessary and sufficient condition is that the Engel curves be affine.

Household i 's relative consumption is $c_{it}^R = \frac{c_{it}}{c_t}$.

Proposition 1.6. *The long-run distribution of consumption is non-degenerate (i.e. it is not true that as $t \rightarrow \infty$, $c_{it}^R = 1$).*

Proof. We know that $c_t \rightarrow c^*$, and if $k_0 < k^*$ then c_t is monotonically increasing. Thus, $\Phi_t^{-1/\eta} \rightarrow (\Phi^*)^{-1/\eta} > 0$ for some Φ^* , and

$$c_i^{R^*} = \frac{M_i(\Phi^*)^{-1/\eta} - \theta_i}{M(\Phi^*)^{-1/\eta} - \theta} \neq 1$$

□

Question. When do we observe $c_i^{R^*} < 1$? Well,

$$c_i^{R^*} < 1 \iff M_i(\Phi^*)^{-1/\eta} - \theta_i < M(\Phi^*)^{-1/\eta} - \theta$$

which holds if

$$(\Phi^*)^{-1/\eta} [m_e(e_i - e) + m_a(a_{i0} - a_0)] < (\theta_i - \theta) \left(1 - (\Phi^*)^{-1/\eta} m_\theta\right)$$

Remark. Some comparative statics: higher e_i or a_{i0} leads to higher $c_i^{R^*}$; lower θ_i leads to lower $c_i^{R^*}$. We also have the (painful) second moments:

$$\begin{aligned} \sigma^2(c_i^{R^*}) = \frac{1}{[M(\Phi^*)^{-1/\eta} - \theta_i]^2} & \left\{ (\Phi^*)^{-2/\eta} [m_e^2 \sigma^2(e) + m_a^2 \sigma^2(a_0)] + \left((\Phi^*)^{-1/\eta} m_\theta - 1 \right)^2 \sigma^2(\theta) \right. \\ & \left. + 2(\Phi^*)^{-1/\eta} [m_e m_a \sigma(e, a_0) + (m_\theta - 1)(m_e \sigma(e, \theta) + m_a \sigma(a_0, \theta))] \right\} \end{aligned}$$

Notice that $\sigma^2(c_i^*) = \sigma^2(c_i^{R^*}) [M(\Phi^*)^{-1/\eta} - \theta_i]^2$, so even if we have two economies A and B with $\sigma_A^2(e) = \sigma_B^2(e)$ and $\sigma_A^2(a_0) = \sigma_B^2(a_0)$, we will have differences in consumption variance. A positive covariance between any two of the elements that make up a household's type increases long-run variance in consumption.

Question. What do you think of the statement 'Controlling for initial wealth inequality, all countries display the same amount of consumption inequality'?

Model. *Bewley Economies* We have that agents are *ex ante* identical, where the only source of heterogeneity is income risk. Agents cannot insure against these shocks due to incomplete markets. We have discrete time and no aggregate uncertainty, so the aggregate endowment \bar{e} is constant. We have a measure 1 of infinitely-

lived households, with preferences

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u(c_{it})$$

We have random endowments $e_{it} \in E$ where $|E| < \infty$, and transition matrix $\pi(e_{it+1}, e_{it})$ with stationary distribution $\Pi(e)$, where $\Pi(e)$ = the measure of households with endowment e . The households solve

$$\max_{c_{it}, a_{it+1}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u(c_{it})$$

subject to a budget constraint and a borrowing constraint:

$$c_{it} + a_{it+1} = e_{it} + (1+r)a_{it} \quad \text{and} \quad a_{it+1} \geq -b, \text{ where } b \in \mathbb{R}_+$$

In recursive form, we have that

$$v(a, e) = \max_c u(c) + \beta \sum_{e'} \pi(e', e) v(a', e')$$

subject to

$$c + a' = e + (1+r)a \quad \text{and} \quad a' \geq -b$$

Remark. We need to be careful here! The borrowing constraint may bind here.

This Bellman equation becomes

$$v(a, e) = \max_{a'} u(e + (1+r)a - a') + \beta \sum_{e'} \pi(e' | e) v(a', e') + \mu(a' + b)$$

Sufficient conditions for optimality are

$$\frac{\partial u(c)}{\partial c} = \beta \sum_{e'} \pi(e' | e) \frac{\partial v(a', e')}{\partial a'} + \mu \quad (a')$$

$$\frac{\partial v(a, e)}{\partial a} = \frac{\partial u(c)}{\partial c} (1+r) \quad (a)$$

$$\mu(a' + b) = 0 \quad (\text{CS})$$

The Euler equation is

$$\frac{\partial u(c)}{\partial c} \geq \beta \sum_{e'} \pi(e' | e) \frac{\partial u(c')}{\partial c'} (1+r)$$

with equality whenever $a' > -b$ (*i.e.* when $\mu = 0$). Solutions to the household problem are functions $v(a, e)$, $a'(a, e)$ that satisfy the sufficient conditions, for a given r .

Assume shocks are i.i.d. Then consumption and savings decisions depend on the current value of income $x \equiv e + (1+r)a$, and savings are increasing in the current value of income $\partial a' / \partial x > 0$. If x is sufficiently high, choose $a' > -b$ and the Euler equation will hold with equality. If not, choose $a' = -b$ and let the Euler equation be violated. Essentially, current consumption is lower than optimal whenever the constraint binds.

The aggregate state is the joint distribution of assets and endowments $\Phi(a, e)$. In a stationary recursive competitive equilibrium, aggregate quantities and prices are constant over time.

Definition. A *stationary equilibrium* is an allocation and prices such that

1. Households maximize utility per $v(a, e)$, $a'(a, e)$
2. Markets clear

3. $\Phi(a, e)$ is time invariant

For market clearing, we need aggregate consumption:

$$C = \int_a \int_e c(a, e) \Phi(da, de) = \int_0^1 e \Pi(de) = e$$

and aggregate assets:

$$\int_a \int_e a'(a, e) \Phi(da, de) = 0$$

(since the households borrow from each other). We define a transition function $Q((a, e), (A, E))$ which denotes the probability (or mass of households) in state (a, e) that transition to state $(a', e') \in (A, E)$ tomorrow – so

$$Q((a, e), (A, E)) = \begin{cases} \sum_{e' \in E} \pi(e' | e) & \text{if } a'(a, e) \in A \\ 0 & \text{otherwise} \end{cases}$$

Note that a' is determined today. Finally, we have the law of motion for Φ

$$\Phi'(A, E) = \int_e \int_a Q((a, e), (A, E)) \Phi(da, de)$$

Algorithm. To solve for the stationary equilibrium:

1. Given the interest rate, solve for the policy function of the household.
2. Given the policy function, iterate over the law of motion of the aggregate state until $\Phi' = \Phi$.
3. Using the stationary distribution, check market clearing.
4. If aggregate asset positions are positive, lower the interest rate and go back to 1.
5. If aggregate asset positions are negative, raise the interest rate and go back to 1.
6. Iterate until market clearing is satisfied.

Example. [Huggett \(1993\)](#) used six periods per year, a time discount $\beta = 0.66^{1/6} = 0.993$ per period, CRRA utility, a two-state Markov chain with $y_H = 1$, $y_L = 0.1$, and transition probabilities $\pi(H | H) = 0.925$, $\pi(L | L) = 0.5$, solved on a grid of borrowing constraints \bar{a} . The asset policy was basically linear, with optimal for the high types being almost exactly the 45° line, and the low types close to a linear downward shift from them. The consumption policy looked as usual, high types always above low types, with the difference being large when asset levels are low and small when asset levels are high. Finally, excess demand is clearly a decreasing, continuous function of the bond price, and so is an increasing and continuous function of the interest rate.

He additionally assumed complete risk-sharing so $c_{it} = C = Y$, bond price (interest rate) $q \equiv \frac{1}{1+r} = \beta$, and asset dynamics $a_{it+1} = (1+r)(a_{it} + e_{it} - Y)$. He found with low risk-aversion, the risk-free rates were

\bar{a}	r	q
-2	-7.1%	1.0124
-4	2.3%	0.9962
-6	3.4%	0.9944
-8	4.0%	0.9935

So a tighter borrowing constraint (higher \bar{a}) led to higher demands for savings ($\uparrow q, \downarrow r$), and looser borrowing constraint (lower \bar{a}) led to lower demands for savings ($\downarrow q, \uparrow r$), with convergence to complete market equilibrium. With high risk-aversion, the risk-free rates were

\bar{a}	r	q
-2	-23%	1.0448
-4	-2.6%	1.0045
-6	1.8%	0.9970
-8	3.7%	0.9940

Higher risk aversion α leads to higher demand for savings, meaning lower r for all \bar{a} .

Example. *Numerical Tools for Incomplete Markets* We want to solve the problem

$$V(a, \varepsilon) = \max_{a', \tilde{c} \geq 0} u(\tilde{c}) + \beta \mathbb{E}[V(a', \varepsilon')]$$

subject to

$$\begin{aligned}\tilde{c} + a' &= w\varepsilon + (1+r)a \\ a' &\geq \underline{a} \\ \varepsilon' &= \rho\varepsilon + \epsilon \\ \epsilon &\sim \text{i.i.d}\end{aligned}$$

where ϵ are errors from the autoregressive process. We will solve this using the *Collocation method*, where we say that $V(a, \varepsilon) \approx B(a, \varepsilon)c$ where B are basis functions and c are *collocation coefficients*. Our objective is to find \hat{f} that minimizes

$$\|f - \hat{f}\|_{\infty} \equiv \sup_{x \in D} |f(x) - \hat{f}(x)|$$

where a standard function interpolation would have that $\hat{f}(x, c)$ is a linear combination of polynomials with coefficients c , where we choose c to minimize $|f - \hat{f}|$ at a finite number of nodes. The key is our choice of polynomial and nodes.

Algorithm. *Collocation Method*

1. Choose basis functions $b_1(x), \dots, b_n(x)$

Remark. $b_n(x) = x^n$ typically a bad idea. Often better are Chebyshev orthogonal polynomials or Splines k -th order polynomials spliced together (popular choice for linear models).

2. Choose nodes $x = x_1, \dots, x_n$

Remark. Equidistant typically are bad. Chebyshev nodes, the roots of the n -th degree Chebyshev polynomial, are optimal.

The Chebyshev polynomial has a certain number of waves, depending on the basis. Equidistant nodes will lead to massive tail errors, as n increases. However, using Chebyshev nodes the approximation error decreases as n increases, and we can get a really good approximation of most smooth functions.

Splines basis functions are basically indicator functions, each has a peak at a certain point among the domain. A nice feature of these is that they are shape-preserving, and can be combined to span basically any function you want. Splines can be non-linear, but linear splines actually tend to do a good job.

Algorithm. *Collocation Method Computation*

1. Choose order n of polynomial
2. Choose $m \geq n$ nodes
3. Evaluate at nodes $B(x)$
4. Find c using $F = Bc$, then compute $c = B \setminus F$ (where F is also evaluated on x).

The nodes $B(x)$ are defined as

$$B(x) = \begin{bmatrix} b_1(x_1) & b_2(x_1) & \cdots & b_n(x_1) \\ b_1(x_2) & b_2(x_2) & \cdots & b_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ b_1(x_m) & b_2(x_m) & \cdots & b_n(x_m) \end{bmatrix}$$

If we are approximating a multidimensional function $f(x, y)$, we generate x and y nodes, choose basis functions b_i^x, b_i^y , and approximate

$$F = \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} b_i^x(x) b_j^y(y) c_{ij}$$

Finally, we find c with $c = B \setminus F$, where $B = B_m \otimes B_{m-1} \otimes \cdots \otimes B_1$.

Back to the incomplete markets model, we will discretize the shock process and solve it using the collocation method, where $V(a, \varepsilon) \approx B(a, \varepsilon)c$. Commonly, people use Tauchen or Rouwenhorst to do this discretization. A common assumption is that $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. In broad strokes, we will:

1. Choose the parameters and grid
2. Define the space, meaning B and B^E

Remark. Note that B^E can be computed using B and c :

$$\mathbb{E}[V(a, \varepsilon)] = B(a, \varepsilon)c_E = \sum_{\omega_i} \omega_i V(a, \rho\varepsilon + \varepsilon_i) = \sum_{\omega_i} \omega_i B(a, \rho\varepsilon + \varepsilon_i)c \implies c_E = \underbrace{B(a, \varepsilon)^{-1} \sum_{\omega_i} \omega_i B(a, \rho\varepsilon + \varepsilon_i)c}_{B^E}$$

3. Then we guess c_0 and solve the following linear equation until convergence:

$$B(a, \varepsilon)c_{j+1} = \max_{a' \geq 0} u(w\varepsilon + (1+r)a - a') + (a', \varepsilon)B^E c_j$$

Aggregate assets are $A = \sum_{\varepsilon} a \cdot \phi(a, \varepsilon)$, and we will assume $A = K$ in equilibrium, and find r such that markets clear. To aggregate we need to compute the *Transition Probability Matrix (TPM)* P using the policy functions. We have a complete ergodic distribution $\phi(a, \varepsilon)$ such that P is a fixed point: $\phi(a, \varepsilon)P = \phi(a, \varepsilon)$. We can solve it by computing the TPM and finding the eigenvectors.

Remark. A quick detour on Markov processes. A stochastic process is a sequence of random vectors, indexed by time. A stochastic process has the *Markov property* if for all t and $k \geq 1$,

$$\mathbb{P}(x_{t+1} \mid x_t, \dots, x_{t-k}) = \mathbb{P}(x_{t+1} \mid x_t)$$

Assuming the Markov property, we can characterize a process by a *Markov chain*, where π_0 is the initial distribution, P is the transition matrix where

$$P_{ij} = \mathbb{P}(x_{t+1} = e_j \mid x_t = e_i), \quad \sum_j P_{ij} = 1$$

and e_i is a vector of zeros except for 1 in entry i . The probability of being in state j in two periods is:

$$P_{ij}^2 = \mathbb{P}(x_{t+2} = e_j \mid x_t = e_i) = \sum_{h=1}^n \mathbb{P}(x_{t+2} = e_j \mid x_{t+1} = e_h) \mathbb{P}(x_{t+1} = e_h \mid x_t = e_i)$$

Therefore, unconditional probabilities are determined by

$$\pi'_1 = \mathbb{P}(x_1) = \pi'_0 P \quad \text{and} \quad \pi'_2 = \mathbb{P}(x_2) = \pi'_0 P^2$$

A *stationary distribution satisfies* $\pi' = \pi' P \implies \pi'(I - P) = 0 \implies (I - P')\pi = 0$. Does this distribution necessarily exist? Yes! P has at least one eigenvalue and at least one eigenvector π , which we can normalize such that $\sum_i \pi_i = 1$. Multiplicity is possible, however. For a given π_0 , it might be the case that

$$\lim_{t \rightarrow \infty} \pi_t = \pi_\infty$$

If this equation holds and $\pi_\infty \perp \pi_0$, then the process is asymptotically stationary with a unique distribution π_∞ , called the *asymptotic* (or *invariant*) distribution of P .

Remark. Some examples that are interesting to work through:⁴

1. Discrete and continuous choice problems
2. Heterogeneous firm models
3. Panel simulations:
 - (a) We can also simulate many households over time
 - (b) For example: Select a very long period of time and N agents, draw ε many times, use the policy function to compute the consumption and assets of each agent, for initial conditions to remain irrelevant, remove the first periods of the simulations, and the simulation can either be discrete or continuous.
 - (c) We can study micro behavior of consumption and income jointly, checking if the distribution converges to the one we compute by inverting the TPM, etc
 - (d) Further notes on how to do this in the slides.

1.7 Continuous Time Growth

Some facts about continuous time growth (assuming that production technology is Cobb-Douglas, where $Y_t = A_t M_t K_t^\alpha H_t^{1-\alpha}$ for H_t human capital, M_t is a measure of uncertainty, and A_t is the level of technology):

1. Mathematically, we have that

$$\frac{Y_t}{L_t} = Z_t \frac{H_t}{L_t} \left(\frac{K_t}{Y_t} \right)^{\frac{\alpha}{1-\alpha}} \quad \text{and} \quad Z_t = (A_t M_t)^{\frac{1}{1-\alpha}}$$

So all long-term growth in the Ramsay-Cass-Koopmans model is from either A or population L , which is exogenous!

2. From the one-sector growth model, we have that (i) $\frac{Y_t}{L_t}$ and $\frac{K_t}{L_t}$ grow over time at roughly constant equal rates (implication: $\frac{K_t}{Y_t}$ constant over time), (ii) that $\frac{L_t}{Y_t}$ is constant, and (iii) that $\frac{R_t K_t}{Y_t}$ and $\frac{W_t L_t}{Y_t}$ constant.
3. These predictions are consistent with the Kaldor facts!
4. They also are consistent with *some* catch-up dynamics, but transitions are too fast.

Model. *Ramsay-Cass-Koopmans (RCK)* We have a representative, infinitely-lived family (or dynasty), with population growth $n > 0$, labor force $L(t) = \exp(nt)$ where $L(0) = 1$ by assumption, constant returns

⁴CompEcon is very flexible and has some very handy computational tools!

technology $Y(t) = F(K(t), A(t)L(t))$, and capital dynamics:

$$\dot{K}(t) + \delta K(t) = F(K(t), A(t)L(t)) - C(t)$$

define $c_t = \frac{C(t)}{L(t)}$, $\zeta(t) = \frac{C(t)}{A(t)L(t)}$, and $\kappa(t) = \frac{K(t)}{A(t)L(t)}$ (this is in *efficiency units*). We have that

$$\frac{\dot{K}(t)}{A(t)L(t)} + \frac{\delta K(t)}{A(t)L(t)} = F\left(\frac{K(t)}{A(t)L(t)}, 1\right) - \frac{C(t)}{A(t)L(t)}$$

which simplifies to

$$\dot{\kappa}(t) = F(\kappa(t)) - \zeta(t) - (n + g + \delta)\kappa(t)$$

and thus

$$\dot{\kappa}(t) = \left(\frac{\dot{K}(t)}{A(t)L(t)} \right) = \frac{\dot{K}(t)}{A(t)L(t)} - n\kappa(t) - g\kappa(t)$$

Note the intuition here: we have that $\frac{1}{A(t)} \frac{\partial A(t)}{\partial t} = g$, and that $\frac{1}{L(t)} \frac{\partial L(t)}{\partial t} = n$. We can show this by using Ekaterina's trick of converting to logarithmic forms, taking derivatives, and showing that those partials are equivalent to the (more complicated) fractions.

Problem: F is decreasing in capital, meaning no endogenous long-run growth!

Solution: Constant returns to capital! See below.

Model. *The AK Model.* The simplest version of this assumes that $Y(t) = A(t)K(t)$, capital per capital is $\frac{K(t)}{L(t)}$, and feasibility is

$$\dot{k}(t) = Ak(t) - c(t) - \delta k(t)$$

or, with population growth,

$$\dot{k}(t) = Ak(t) - c(t) - (n + \delta)k(t)$$

The family values *per capita* consumption as

$$u(c) = \int_0^\infty \exp(-\rho t) U(c(t)) dt$$

where ρ is the time discount factor, and $U(c(t))$ is a CRRA function, so $U(c) = \frac{c^{1-\sigma}}{1-\sigma}$. We can show that an allocation (k, c) is Pareto optimal if and only if it solve the social planner's problem, which is

$$\max_{k, c \geq 0} \int_0^\infty \exp(-\rho t) U(c(t)) dt$$

subject to

$$\dot{k}(t) = Ak(t) - c(t) - (n + \delta)k(t) \quad \text{with} \quad k(0) = k_0$$

We can solve this problem using Pontryagin's Maximum Principle, with control variable c , state variable k , and co-state variable λ . The present-value Hamiltonian is

$$\mathcal{H}(t, k, c, \lambda) = \exp(-\rho t) U(c(t)) + \lambda(t) [f(k(t)) - c(t) - (n + \delta)k(t)]$$

which admits sufficient conditions

$$0 = \exp(-\rho t) U'(c(t)) - \lambda(t) \tag{c}$$

$$\dot{\lambda}(t) = -[f'(k(t)) - (n + \delta)] \lambda(t) \tag{k}$$

$$0 = \lim_{t \rightarrow \infty} \lambda(t) k(t) \tag{TVC}$$

plus the constraint

$$\dot{k}(t) = Ak(t) - c(t) - (n + \delta)k(t)$$

We can get rid of the co-state variable by differentiating

$$\dot{\lambda}(t) = \exp(-\rho t)U''(c(t))\dot{c}(t) - \rho \exp(-\rho t)U'(c(t)) \implies \frac{\dot{\lambda}(t)}{\lambda(t)} = \frac{U''(c(t))}{U'(c(t))}\dot{c}(t) - \rho$$

where we replace

$$- [f(k(t)) - (n + \delta + \rho)]c(t) = \frac{U''(c(t))}{U'(c(t))}c(t)\dot{c}(t)$$

and because U is CRRA,

$$\sigma \dot{c}(t) = [A - (n + \delta + \rho)]c(t)$$

Definition. A *balanced growth path (BGP)* is an allocation such that consumption, capital, and output grow at a constant (possibly different) growth rate

$$\frac{\dot{x}(t)}{x(t)} = \gamma_x > 0 \text{ for any variable } x$$

In this particular problem, the consumption dynamics are independent of capital.⁵

$$c(t) = \exp\left(\frac{1}{\sigma} [A - (n + \delta + \rho)] t\right) c(0)$$

So consumption grows at a constant rate from the beginning. Capital dynamics are

$$\frac{\dot{k}(t)}{k(t)} = A - \frac{c(t)}{k(t)} - (n + \delta)$$

Along the BGP, consumption and capital grow at the same rate, so the economy is on the BGP from time $t = 0$. We need an additional condition so that utility does not blow up, so U remains well-defined. Observe that:

$$\int_0^\infty \exp(-\rho t)U(c(t))dt = c(0)^{\frac{1-\sigma}{\sigma}} \int_0^\infty \exp(-\rho t) \frac{\exp\left(\frac{1-\sigma}{\sigma} [A - (n + \delta + \rho)] t\right) dt}{1 - \sigma}$$

So we need that

$$\frac{1 - \sigma}{\sigma} \left[A - (n + \delta) - \frac{\rho}{1 - \sigma} \right] < 0$$

Remark. The main implication is that no country can ever catch up. However, we see some of that in the data – how do we reconcile that? What would happen if we added labor? With increasing returns to scale, we don't have a well-defined equilibrium in the standard AK model. [Romer \(1986\)](#) argued for increasing returns to scale but the agents in the economy behave as if constant returns to scale.

Question. Why do we consider discounting rather than no discount factor and maximize mean consumption in each period rather than have discounting?

Answer. We don't exactly have a Permanent Income Hypothesis here, as this is a production economy. We essentially, in most cases, have a single optimal path, and it's always worth it to get on the optimal path as quickly as possible. See [the Turnpike Theorem](#).

⁵Note not true in the standard RCK model, remember the sequential economy.

Model. *RCK with Productivity Growth* Recall that the optimal allocation satisfies

$$\begin{aligned}\dot{\zeta}(t) &= \frac{1}{\sigma} [f'(\kappa(t)) - (n + \sigma g + \delta + \rho)] \zeta(t) \\ \dot{\kappa}(t) &= f(\kappa(t)) - \zeta(t) - (n + g + \delta)\kappa(t) \\ 0 &= \lim_{t \rightarrow \infty} \exp(-\hat{\rho}t) U'(\zeta(t)) \kappa(t)\end{aligned}$$

where g is exogenous growth of technology, and $\kappa(t) = \frac{K(t)}{A(t)L(t)}$, $\zeta(t) = \frac{C(t)}{A(t)L(t)}$. We have steady state analysis $x^* = (\zeta^*, \kappa^*)$, and saddle paths. From the theory of linear approximation, we know that when the path is “nice” (this condition is quite important! think of it as a convex problem that is everywhere differentiable), the behavior around the steady state is well-approximated by the behavior of a linear system around the steady state. We use the first-order Taylor approximation

$$f(x) = f(x^*) + \nabla f(x^*) \cdot (x - x^*)$$

where $\nabla f(x^*)$ is the \mathbb{R}^n gradient to f at x^* . In our case, $x^* = (\zeta^*, \kappa^*)$ and we have g_1 and g_2 that solve

$$\begin{aligned}g_1(\zeta(t), \kappa(t)) &= \dot{\zeta}(t) = \frac{1}{\sigma} [f'(\kappa(t)) - (n + g + \delta + \rho)] \zeta(t) \\ g_2(\zeta(t), \kappa(t)) &= \dot{\kappa}(t) = f(\kappa(t)) - \zeta(t) - (n + g + \delta)\kappa(t)\end{aligned}$$

where $g_1(\zeta^*, \kappa^*) = g_2(\zeta^*, \kappa^*) = 0$. From our linear approximation, we have

$$\begin{bmatrix} \dot{\zeta}(t) \\ \dot{\kappa}(t) \end{bmatrix} \approx \begin{bmatrix} -\frac{1}{\sigma} [f'(\kappa(t)) - (n + g + \delta + \rho)] & -\frac{1}{\sigma} f''(\kappa(t)) \zeta(t) \\ -1 & f'(\kappa(t)) - (n + g + \delta) \end{bmatrix}_{(\zeta^*, \kappa^*)} \begin{bmatrix} \zeta(t) - \zeta^* \\ \kappa(t) - \kappa^* \end{bmatrix}$$

which becomes

$$\begin{bmatrix} \dot{\zeta}(t) \\ \dot{\kappa}(t) \end{bmatrix} \approx \begin{bmatrix} 0 & -\frac{1}{\sigma} f''(\kappa^*) \zeta^* \\ -1 & \rho \end{bmatrix}_{(\zeta^*, \kappa^*)} \begin{bmatrix} \zeta(t) - \zeta^* \\ \kappa(t) - \kappa^* \end{bmatrix}$$

This is a two-dimensional difference equation that can be solved analytically. We look at the eigenvalues of $\nabla f(\zeta^*, \kappa^*)$, λ , which satisfy

$$0 = \det(\nabla f(\zeta^*, \kappa^*) - \lambda I) = -\lambda(\rho - \lambda) + \frac{1}{\sigma} f''(\kappa^*) \zeta^*$$

This quadratic equation has two roots, one negative and one positive. Let l be the number of negative eigenvalues, and let m be the number of state variables of the problem. Stability comes directly from these: if $l = m$, we are *saddle-path stable*, we have a unique optimal trajectory (the negative eigenvalue governs the speed of convergence), if $l < m$, unstable, no convergence to the steady state, and if $l > m$, indeterminacy, multiple optimal trajectories.

The speed of convergence is

$$\kappa(t) - \kappa^* \approx e^{-|\lambda_1|t} (\kappa(0) - \kappa^*)$$

with half-life

$$\kappa(t_{1/2}) - \kappa^* \approx \frac{1}{2} (\kappa(0) - \kappa^*) \text{ hence } t_{1/2} = \frac{\ln(2)}{|\lambda_1|}$$

1.8 Endogenous Growth Theory

1.8.1 Externalities

The AK model generates long-term growth. Typically, we think about this as either spillovers or learning by doing. How do we endogenize growth?

Remark. First, what worked in the AK model? Constant returns in a reproducible factor (capital), and we left out the non-reproducible factor (labor). The dilemma is that we left out the non-reproducible factor. Paul Romer had an elegant solution.

Model. *Endogenous Growth Model* (from Romer (1986)) The economy is the same as the AK model setup but we have a continuum of firms that produce output with technology

$$y_i(t) = F(k_i(t), l_i(t)K(t))$$

where firms are indexed with $i \in [0, 1]$, k_i, l_i are capital and labor at firm i , $K(t) = \int k_i(t)di$ is aggregate capital, and $F(\cdot)$ has constant returns to scale with respect to k_i and l_i . This gives us an externality, where firms do not understand that requesting more capital will increase the aggregate capital, they just take that as given.⁶ The firm's problem is

$$\max_{k_i, l_i} F(k_i(t), l_i(t)K(t)) - w(t)l_i(t) - r(t)k_i(t)$$

where $K(t)$ is *exogenous to the firm*. Note that there are increasing returns to scale overall:

$$F(\theta k_i(t), \theta l_i(t) \underbrace{\theta K(t)}_{=\int \theta k_i(t)di}) = F(\theta k_i(t), \theta^2 l_i(t)K(t)) > \theta F(k_i(t), l_i(t)K(t))$$

for any $\theta > 1$. The competitive equilibrium exists in this economy because from the firm's perspective there are constant returns to scale. Will it be Pareto Optimal? Not in general! This is straightforward to see, in a world with externalities.

As before, the households are a size L of identical people, with no population growth.

Definition. A *competitive equilibrium* is a set of allocations $\{\hat{c}(t), \hat{a}(t)\}$ for the representative household, a set of allocations $\{\hat{k}_i(t), \hat{l}_i(t)\}$ for each firm i , a stream of aggregate capital stock $\{\hat{K}(t)\}$, and a stream of prices $\{\hat{w}(t), \hat{r}(t)\}$ such that:

1. Given prices, $\{\hat{c}(t), \hat{a}(t)\}$ solves

$$\max_{\{c(t), a(t)\}} \int \exp(-\rho t) \frac{c(t)^{1-\sigma}}{1-\sigma} dt$$

subject to

$$c(t) + \dot{a}(t) + a(t) = w(t) + (r(t) - \delta)a(t)$$

$$a(0) = k(0) \text{ given}$$

$$\lim_{t \rightarrow \infty} a(t) \exp\left(-\int_0^t (r(\tau) - \delta)d\tau\right) \geq 0$$

2. Given $\{\hat{w}(t), \hat{r}(t)\}$ and $\{\hat{K}(t)\}$, the path $\{\hat{k}_i(t), \hat{l}_i(t)\}$ maximizes firm profits in each period for each firm

⁶This is often called the *Big K / small k model*.

3. Feasibility: for all t , markets clear:

$$L\hat{c}(t) + \dot{\hat{K}}(t) + \hat{K}(t)\delta = \int_0^1 F(\hat{k}_i(t), \hat{l}_i(t)\hat{K}(t))di$$

$$\int_0^1 \hat{l}_i(t)di = L \quad \text{and} \quad \int_0^1 \hat{k}_i(t)di = L\hat{a}(t)$$

4. Rational Expectations:

$$\int_0^1 \hat{k}_i(t)di = \hat{K}(t)$$

The planner's problem is

$$\max_{c(t), K(t)} \int \exp(-\rho t) \frac{c(t)^{1-\sigma}}{1-\sigma} dt$$

subject to

$$Lc(t) + \dot{K}(t) - K(t)\delta = F(K(t), LK(t)) \text{ with } K(0) = Lk(0)$$

Note that the planner can deal in terms of total capital and total labor, because the firms have identical production functions, so when they choose capital and labor so that the marginal products of capital and labor respectively equal the wage and rental price, they all choose the same levels, so they can be aggregated. The planner can choose first the aggregates and next how to split them across the agents. Optimality requires that

$$\gamma_C^{SP}(t) = \frac{\dot{c}(t)}{c(t)} = \frac{1}{\sigma} \left[F_1(K(t), LK(t)) + F_2(K(t), LK(t))L - (\delta + \rho) \right]$$

Since F is homogeneous of degree 1 in capital, F' is homogeneous of degree 0, so we have that

$$F_1(K(t), LK(t)) + F_2(K(t), LK(t))L = F_1(1, L) + F_2(1, L)L$$

Thus, consumption growth is

$$\gamma_C^{SP}(t) = \frac{\dot{c}(t)}{c(t)} = \frac{1}{\sigma} \left[F_1(1, L) + F_2(1, L)L - (\delta + \rho) \right]$$

which is constant in time! From the aggregate resource constraint:

$$L \frac{c(t)}{K(t)} + \frac{\dot{K}(t)}{K(t)} + \delta = F(1, L)$$

so in the balanced growth path, $\gamma_C^{SP}(t) = \gamma_k^{SP}(t) = \gamma_K^{SP}(t)$. From the first order conditions of the household and the firm, we have that

$$\gamma_C^{CE}(t) = \frac{\dot{c}(t)}{c(t)} = \frac{1}{\sigma} \left[r(t) - (\delta + \rho) \right] \quad \text{and} \quad r(t) = F_1(k_i(t), l_i(t)K(t))$$

Since all firms are identical and choose the same allocations, $k_i(t) = k(t) = \int k_i(t)di = K(t)$ and $l_i(t) = L$, so $r(t) = F_1(1, L)$, so

$$\gamma_C^{CE}(t) = \frac{1}{\sigma} \left[F_1(1, L) - (\delta + \rho) \right] < \gamma_C^{SP}(t)$$

since $F_2(1, L) > 0$. So in the balanced growth path, $\gamma_C^{CE} = \gamma_k^{CE} = \gamma_K^{CE}$. How do we bring γ_K^{CE} to γ_K^{SP} ? With a subsidy per unit of capital invested, equal to $F_2(1, L)L$. That will make the cost of capital to the firms $r(t) - F_2(1, L)L$.

Remark. A prediction of this model is that larger economies grow more, where

$$\gamma_C^{CE}(t) = \frac{\dot{c}(t)}{c(t)} = \frac{1}{\sigma} [F_1(1, L) - (\delta + \rho)] < \gamma_C^{SP}(t)$$

since $\frac{\partial F(1, L)}{\partial L} > 0$. If L is defined as the labor force of a country (or population), the data across post-World War II countries show no evidence of this fact. Technically, this is because we have constant returns in K and increasing returns in K and L . We can avoid this by assuming that productivity depends on the ratio $\frac{K}{L}$ rather than the aggregate. An example of this is the following:

Model. *Externalities in Human Capital* (from Lucas, 1988) We have households with measure 1 and human capital h_i for $i \in (0, 1)$, with $h_i(0) = h_0$ and $k_i(0) = k_0$. The household chooses time to work $1 - s_i(t)$ and time to accumulate human capital $s_i(t)$, with budget constraint(s)

$$\begin{aligned} c_i(t) + \dot{a}_i(t) &= (r(t) - \delta)a_i(t) + (1 - s_i(t))h_i(t)w(t) \\ \dot{h}_i(t) &= \theta h_i(t)s_i(t) - \delta h_i(t) \end{aligned}$$

and production technology

$$Y(t) = AK(t)^\alpha L(t)^{1-\alpha} H(t)^\beta$$

Firms choose capital and labor and they take $H(t)$ as given. Since the total supply of labor is $(1 - s(t))H(t)$, this becomes

$$AK(t)^\alpha H(t)^{1-\alpha} H(t)^\beta (1 - s(t))^{1-\alpha} = AK(t)^\alpha H(t)^{\beta+(1-\alpha)} (1 - s(t))^{1-\alpha}$$

The planner's problem is

$$\max \int_0^\infty \exp(-\rho t) \frac{c(t)^{1-\sigma}}{1-\sigma} dt$$

subject to

$$\begin{aligned} C(t) + \dot{K}(t) + \delta K(t) &= AK(t)^\alpha H(t)^{\beta+(1-\alpha)} (1 - s(t))^{1-\alpha} \\ \dot{H}(t) &= \theta H(t)s(t) - \delta H(t) \end{aligned}$$

with $H(0)$ and $S(0)$ given, $s(t) \in [0, 1]$, and we have used the fact that

$$\int (1 - s_i(t))h_i(t)di = L(t) \quad ; \quad \int a_i(t)di = K(t) \quad ; \quad \int c_i(t)di = C(t)$$

Firms solve the simple

$$\max_{L(t), K(t)} Y(t) - w(t)L(t) - r(t)K(t)$$

where

$$r(t) = \alpha \frac{Y(t)}{K(t)} \quad \text{and} \quad w(t) = (1 - \alpha) \frac{Y(t)}{(1 - s)H(t)}$$

Households solve the Hamiltonian

$$\begin{aligned} \mathcal{H}(\cdot) &= \exp(-\rho t) \frac{c_i(t)^{1-\sigma}}{1-\sigma} + \lambda(t) [(r(t) - \delta)a_i(t) + (1 - s_i(t))h_i(t)w(t) - c_i(t)] \\ &\quad + \mu(t) [\theta h_i(t)s_i(t) - \delta h_i(t)] \end{aligned}$$

Again, we have that $\gamma_C^{SP} > \gamma_C^{CE}$.

1.8.2 Innovation

In the economies we've studied so far, we've been on the balanced growth path from the beginning. We can think of transition dynamics by relaxing that, and think of growth through variety innovation.

Model. *Transition Dynamics* (from Jones & Manuelli, 1990) We have production technology

$$Y = F(K, L) = AK + BK^\alpha L^{1-\alpha}$$

so that $y = Ak + Bk^\alpha$ and $\lim_{t \rightarrow \infty} f'(k) = A$. We have growth rates

$$\frac{\dot{k}}{k} = \frac{f(k)}{k} - \frac{c}{k} - (n + \delta) \quad \text{and} \quad \frac{\dot{c}}{c} = \frac{1}{\theta} (A + B\alpha k^{\alpha-1} - (\rho + \delta))$$

where θ is the elasticity of substitution in consumption. In the balanced growth path, we have growth rates

$$\gamma^* = \frac{1}{\theta} (A - (\rho + \delta))$$

The problem is that we have no steady state. We can rewrite the variables in stationary terms (*i.e.* *detrend* them). Which variables do we detrend with respect to? Here, capital. Generally it depends on the problem. We have

$$z = \frac{f(k)}{k} \quad \text{and} \quad \chi = \frac{c}{k}$$

By doing a bunch of (annoying) algebra, we can rewrite the dynamic system as

$$\begin{aligned} \dot{z} &= -(1 - \alpha)(z - A)(z - \chi - n - \delta) \\ \dot{\chi} &= \chi \left((\chi - \varphi) - \frac{\theta - \alpha}{\theta} (z - A) \right) \end{aligned}$$

where $\varphi \equiv (A - \delta)\frac{\theta-1}{\theta} + \frac{\rho}{\theta} - n$. We can now draw a phase diagram in (χ, z) space!

Writing Recursively. (*Hamilton-Jacobi-Bellman Equations*) We have

$$V(k_0) = \max_{c(t)} \int_0^\infty \exp(-\rho t) U(c(t)) dt$$

subject to

$$\dot{k}(t) = F(k(t)) - \delta k(t) - c(t)$$

for $t \geq 0$, with $k(0) = k_0$ given. Our state x is $k(t)$, and the control u is $c(t)$. Let $h(x, u) = U(u)$ and $g(x, u) = F(x) - \delta x - u$. The value function of the generic optimal control problem solves the Hamilton-Jacobi-Bellman equation

$$\rho V(x) = \max_u h(x, u) + V'(x)g(x, u)$$

With multiple states, $V'(x)$ is a vector of dimension m . This implies that

$$\rho V(k) = \max_c U(c) + V'(k) [F(k) - \delta k - c] \iff U'(c) = V'(k)$$

We define a discount factor $\beta(\Delta) = \exp(-\rho\Delta)$, where Δ is the length of a period. We multiply all of our flows by Δ , and keep the stocks the same. The Bellman equation is

$$V(k_t) = \max_{c_t} \Delta U(c_t) + \exp(-\rho\Delta) V(k_{t+\Delta})$$

subject to

$$k_{t+\Delta} = \Delta[F(k_t) - \delta k_t - c_t] + k_t$$

For small Δ , $\exp(-\rho\Delta) \sim (1 - \rho\Delta)$, so

$$\rho\Delta V(k_t) = \max_{c_t} \Delta U(c_t) + (1 - \rho\Delta)(V(k_{t+\Delta}) - V(k_t))$$

and so

$$\rho V(k_t) = \max_{c_t} U(c_t) + (1 - \rho\Delta) \left(\frac{V(k_{t+\Delta}) - V(k_t)}{k_{t+\Delta} - k_t} \frac{k_{t+\Delta} - k_t}{\Delta} \right)$$

and taking the limit as $\Delta \rightarrow 0$, we get

$$\rho V(k_t) = \max_{c_t} U(c_t) + V'(k_t) \dot{k}_t$$

Recall that we have had the two general forms:

$$\underbrace{\mathcal{H}(x, u, \lambda) \equiv h(x, u) + \lambda g(x, u)}_{\text{Hamiltonian}} \quad ; \quad \underbrace{\rho V(x) = \max_u h(x, u) + V'(x)g(x, u)}_{\text{Bellman}}$$

so we can directly see that $\lambda(t) = V'(x(t))$, the co-state variable is the shadow value of the future. Thus:

$$\rho V(x) = \max_{u \in U} \mathcal{H}(x, u, V'(x))$$

we get the reason for the ‘Hamilton’ in the name of this subsection!

Model. *Endogenous Technological Change* (from [Romer, 1990](#)) We can think about expanding input varieties, where a greater variety of inputs increases the ‘division of labor’ which raises the productivity of final goods firms. We have competitive markets for final goods, monopolistic competition for intermediate goods, and competitive markets for R&D. We have final goods, produced by:

$$Y(t) = L(t)^{1-\alpha} \left(\int_0^{A(t)} x_i(t)^{1-\mu} di \right)^{\frac{\alpha}{1-\mu}}$$

where $1/\mu$ is the elasticity of substitution. We are interested in $\mu \in (0, 1)$, meaning that there is *some* substitution. We also have intermediate goods, produced by

$$x_i(t) = a l_i(t)$$

and R&D, with constant returns to scale:

$$\dot{A}(t) = bX(t)$$

where $X(t)$ are final goods devoted to R&D. We normalize the population to 1, and the planner’s problem is

$$\max \int_0^\infty \exp(-\rho t) \frac{c(t)^{1-\sigma}}{1-\sigma} dt$$

subject to

$$\begin{aligned}
c(t) + X(t) &= L(t)^{1-\alpha} \left(\int_0^{A(t)} x_i(t)^{1-\mu} di \right)^{\frac{\alpha}{1-\mu}} \\
x_i(t) &= al_i(t) \\
\dot{A}(t) &= bX(t) \\
L(t) + \int_0^{A(t)} l_i(t) di &= 1
\end{aligned}$$

Given that there is imperfect substitution, the optimal planner strategy is

$$x_i(t) = x(t) \quad \text{and} \quad l_i(t) = l(t)$$

Hence,

$$c(t) + X(t) = L(t)^{1-\alpha} (A(t)(al(t))^{1-\mu})^{\frac{\alpha}{1-\mu}}$$

Feasibility requires that $l(t) = \frac{1-L(t)}{A(t)}$, and equilibrium aggregate output is

$$Y(t) = a^\alpha L(t)^{1-\alpha} (1-L(t))^\alpha A(t)^{\frac{\alpha\mu}{1-\mu}}$$

Output maximization implies that $L(t) = 1 - \alpha$. The planner's problem is

$$\max \int_0^\infty e^{-\rho t} \frac{c(t)^{1-\sigma}}{1-\sigma} dt \quad \text{s.t.} \quad c(t) + \frac{\dot{A}(t)}{b} = CA(t)^{\frac{\alpha\mu}{1-\mu}}$$

where $C = a^\alpha (1-\alpha)^{1-\alpha} \alpha^\alpha$. If $\frac{\alpha\mu}{1-\mu} \in (0, 1)$, we have Ramsey-Cass-Koopmans. Romer requires that $\alpha\mu = 1 - \mu$, so we have an AK model where

$$\frac{\dot{c}(t)}{c(t)} = \frac{1}{\sigma} [bC - \rho]$$

so we are on the balanced growth path from the beginning.

Alternatively, we could solve the decentralized economy. When a new idea is introduced, the R&D sector charges $k(t)$, where since the sector is competitive and has CRS technology, so $\pi_{R\&D} = 0$. We will assume that $k(t)$ is the present value of the profits of the monopolistic firm, and let $p_i(t)$ be the price of intermediate goods, $p(t)$ be the price of final goods, and $w(t)$ the cost of labor. The final goods firms maximize profits, so we attain the first order conditions

$$\begin{aligned}
w(t) &= (1-\alpha)p(t)L(t)^{-\alpha} \left(\int_0^{A(t)} x_i(t)^{1-\mu} di \right)^{\frac{\alpha}{1-\mu}} & (L(t)) \\
p_i(t) &= \alpha p(t)L(t)^{1-\alpha} \left(\int_0^{A(t)} x_i(t)^{1-\mu} di \right)^{\frac{\alpha}{1-\mu}-1} x_i(t)^{-\mu} & (x_i(t))
\end{aligned}$$

which simplify to

$$w(t) = (1-\alpha)p(t) \frac{Y(t)}{L(t)} \quad \text{and} \quad x_i(t)^\mu = \alpha \frac{p(t)}{p_i(t)} \frac{Y(t)}{\int_0^{A(t)} x_i(t)^{1-\mu} di}$$

Since markets are competitive, $\pi = 0$, and we can use the zero profit condition to get that

$$x_i^*(t) = \alpha^{\frac{1}{\mu}} \left(\frac{p(t)}{p_i(t)} \right)^{\frac{1}{\mu}} Y(t)^{\frac{\mu+\alpha-1}{\alpha\mu}} L(t)^{\frac{(1-\mu)(1-\alpha)}{\alpha\mu}}$$

The intermediate goods producer, who has already paid the fixed cost, solves

$$\max_{p_i(t)} p_i(t) x_i(t) - w(t) \frac{x_i(t)}{a}$$

Because of monopolistic competition, x_i is a function of $p_i(t)$. Optimality yields that $p_i(t) = \frac{w(t)}{a(1-\mu)}$, which is a constant markup over marginal cost.

In equilibrium, all firms are identical so $p_i(t) = \hat{p}(t)$. We could show that $L^{CE}(t) = \frac{1-\alpha}{1-\alpha\mu} > 1-\alpha$ so there is more labor in the final goods sector than in the planner's allocation. Intuitively, $x_i(t)$ is relatively expensive, so firms switch to labor rather than innovation – there is less investment in ideas. The growth rates are

$$\frac{\dot{c}^{CE}(t)}{c^{CE}(t)} = \frac{1}{\sigma} \left(b\alpha^\alpha \alpha\mu \frac{1-\alpha}{1-\alpha\mu} - \rho \right) < \frac{\dot{c}(t)}{c(t)}$$

Remark. Variety models deal with horizontal innovation, but typically innovations both (i) improve the quality of the good, and (ii) lower the cost of production. Endogenous growth, in the [Schumpeterian](#) economies involves price competition, the replacement of old vintages, and business stealing effects (entrants).

Model. *Endogenous Vertical Innovation* (from [Aghion & Howitt, 1992](#); [Howitt & Aghion, 1998](#), and [Grossman & Helpman, 1991](#)) We have a representative household with CRRA preferences, constant population, and inelastic labor supply L . The resource constraint is that $C(t) + X(t) + Z(t) = Y(t)$ for R&D $Z(t)$, investment $X(t)$, and consumption $C(t)$. The final good is produced:

$$Y(t) = \frac{1}{1-\beta} L(t)^\beta \left(\int_0^1 q(v, t) x(v, t | q)^{1-\beta} dv \right)$$

where $x(v, t | q)$ is the quantity of machines of vintage v of quality $q(v, t)$ at time t . The source of growth here is [quality improvements](#). We have a [quality ladder](#) for each machine type, where

$$q(v, t) = \lambda^{n(v, t)} q(v, 0)$$

where $\lambda > 1$ and $n(v, t)$ is the number of innovations so far. At any point in time, only one quality of any machine v is used – *i.e.* creative destruction, where the invention of a higher-quality machine ‘destroys’ (replaces) an older machine. Once a machine of quality $q(v, t)$ is invented, any quantity can be produced at cost $\psi q(v, t)$.

Remark. Innovation is driven by entrants. This is sometimes called [Arrow's Replacement Effect](#), where incumbents have little incentive to innovate as that would destroy their own profits.

Innovation requires investment, where $Z(v, t)$ units of the final good are used for research in line v with quality $q(v, t)$. The rate of innovation is $z(v, t | q) = \eta \frac{Z(v, t)}{q(v, t)}$, and we have (i) free entry into research, and (ii) perpetual patents for innovators.

Definition. An [allocation](#) in this economy is a time path for (i) consumption levels, aggregate spending on machines, and aggregate R&D spending $\{C(t), X(t), Z(t)\}_{t=0}^\infty$; (ii) machine qualities $\{q(v, t)\}_{t=0}^\infty$ for $v \in [0, 1]$; (iii) prices and quantities of each machine and the net present value of profits from each machine: $\{p^x(v, t | q), x(v, t | q), V(v, t | q)\}_{t=0}^\infty$ for $v \in [0, 1]$; and (iv) interest rates and wages $\{r(t), w(t)\}_{t=0}^\infty$.

The final good producer solves

$$\max_{L, x(v)} \frac{1}{1-\beta} L(t)^\beta \left(\int_0^1 q(v, t) x(v, t | q)^{1-\beta} dv \right) - w(t) L(t) - \int_0^1 p^x(v, t) x(v, t | q) dv$$

The optimal machine demand is

$$x(v, t | q) = \left(\frac{q(v, t)}{p^x(v, t | q)} \right)^{\frac{1}{\beta}} L$$

There are typically two regimes: (i) drastic innovation, where firms charge monopoly prices, and (ii) limit prices. For now, we assume (i), attained when λ is sufficiently large, such that $\lambda \geq \left(\frac{1}{1-\beta} \right)^{(1-\beta)/\beta}$. We normalize $\psi = 1 - \beta$, and we have that

$$\begin{aligned} \pi(v, t) &= \max_x p^x(v, t | q) x(v, t | q) - \psi q(v, t) x(v, t | q) \\ &= \max_x q(v, t) L^\beta x(v, t | q)^{1-\beta} - \psi q(v, t) x(v, t | q) \end{aligned}$$

then for a profit maximizing monopoly, we have that

$$x(v, t | q) = L \quad ; \quad p^x(v, t | q) = q(v, t) \quad ; \quad \pi(v, t) = \beta q(v, t) L$$

Total output will be

$$Y(t) = \frac{1}{1-\beta} Q(t) L \quad \text{where} \quad Q(t) = \int_0^1 q(v, t) dv$$

and aggregate spending in machines (and the implied equilibrium wage rate) is therefore

$$\int_0^1 p^x(v, t | q) x(v, t | q) dv = Q(t) L \quad \text{which implies that} \quad w(t) = \frac{\beta}{1-\beta} Q(t)$$

The monopolist with vintage v and quality $q(v, t)$ has a value function defined by

$$r(t) V(v, t | q) - \dot{V}(v, t | q) = \pi(v, t | q) - z(v, t | q) V(v, t | q)$$

where $z(v, t | q)$ is the arrival rate of innovations to vintage v . The creative destruction implies that when an innovation occurs, the monopolist loses its monopoly and is replaced by a higher quality producer; and from there the innovation has zero value. This implies that $z(v, t | q)$ is the rate of replacement of incumbents in vintage v . Since we have free entry, entrants have

$$\eta V(v, t | q) \leq \frac{q(v, t)}{\lambda} \quad \text{with equality if } z(v, t | q) > 0$$

The consumer's maximization problem admits the Euler equation

$$\frac{\dot{C}(t)}{C(t)} = \frac{1}{\theta} (r(t) - \rho)$$

with the transversality condition

$$\lim_{t \rightarrow \infty} \exp \left(- \int_0^t r(s) ds \right) \int_0^1 V(v, t | q) dv = 0$$

for all q , where $V(v, t)$ is nonstochastic.

Definition. An *equilibrium* is an allocation such that (i) the aggregate feasibility constraints for goods and

machines and the transversality condition are satisfied; (ii) the value of the firm and the average quality satisfy optimality of the monopolist and machine demands, and free entry; (iii) prices and quantities of machines are as described in the monopolist's problem; and (iv) the interest rate and wages are consistent with the Euler equation and feasibility in labor markets.

Given that consumption grows at a constant rate on the balanced growth path, feasibility implies that output also grows at a constant rate, and from the Euler equation, the interest rate is constant. If there is positive growth there must be research in at least one sector. Linearity of the value of the firm and innovation costs in quality together imply that free entry holds for all varieties. If free entry holds in all periods, then $\dot{V}(v, t | q) = 0$ and R&D for each machine type has the same productivity $z(v, t) = z(t) = z^*$. Then, the firm value is

$$V(v, y | q) = \frac{\beta q(v, t)L}{r^* + z^*}$$

so we have an effective discount of $r^* + z^*$, and by free entry and the Euler equation respectively, we have:

$$r^* + z^* = \eta\beta\lambda L \quad ; \quad g^* = \frac{r^* - \rho}{\theta} \implies r^* = g^* \cdot \theta + \rho$$

From the definition of output, we have $\frac{\dot{Y}(t)}{Y(t)} = \frac{\dot{Q}(t)}{Q(t)}$, and dynamics

$$Q(t + \Delta t) = \lambda Q(t)z(t)\Delta t + (1 - z(t)\Delta t)Q(t) + o(\Delta t)$$

Note that the measure of the varieties that experience more than one innovation is second order in t , so $\frac{o(\Delta t)}{\Delta t} \rightarrow 0$. Thus, we have that

$$\dot{Q}(t) = (\lambda - 1)z(t)Q(t) \quad ; \quad g^* = (\lambda - 1)z^*$$

So the equilibrium growth rate is

$$g^* = \frac{\eta\lambda\beta L - \rho}{\theta + (\lambda - 1)^{-1}}$$

2 Kristoffer Nimark

2.1 Introduction

This course will cover New Keynesian business cycle theory with applications to monetary policy, and will introduce tools for applied macroeconomic research. Grades will be 40% homework (four problem sets) and 60% a comprehensive final exam. The texts are mainly [Jordi Gali's *Monetary Policy, Inflation, and the Business Cycle*](#), with some use of [James Hamilton's *Time Series Analysis*](#).

Macroeconomists think about (i) growth, (ii) business cycles, (iii) fiscal and monetary policy, (iv) inequality and heterogeneity, (v) efficiency, (vi) dynamic decisions, (vii) information and uncertainty, and (viii) general equilibrium. This section of the course will think about business cycles, monetary policy, efficiency, dynamic decisions, and general equilibrium models. All macro models are simplified versions of reality, and most simplify using aggregation, rationality, equilibrium, and mathematics. We need simple models so that we can understand economic mechanisms, and a good researcher finds the *right* simple model for a given question.

Our theoretical framework here will be the New Keynesian business cycle model, where we have a representative household that works and consumes goods, firms that produce heterogeneous goods and have some market power, and a monetary policy authority that sets nominal interest rates. These are the three equations that characterize the base of all New Keynesian models. We will use this model to learn to manipulate a (linearized) business cycle model, to study stabilization policy and welfare, and as a vehicle to learn tools and strategies for relating models to data.

To call something a business cycle, we need (i) comovement across macro aggregates (so simultaneous increases / decreases in GDP, employment, consumption, and investment); and (ii) comovement across broad sectors of the economy (so simultaneous increases / decreases in construction, manufacturing, services, etc.). Are business cycles actually *cyclical*? Probably not.

The Plan:

- Lecture [2.1](#) is today.
- Lecture [2.2](#) will introduce a classical monetary model with firms and production.
- Lecture [2.3](#) will introduce a basic New Keynesian model with imperfectly competitive markets – monopolistic competition and CES demand – and will describe how inefficiencies arise.
- Lecture [2.4](#) will continue the basic New Keynesian model, with sticky prices (Calvo pricing and the New Keynesian Phillips Curve, and monetary non-neutrality), and will show that sticky prices imply that monetary policies have real effects.
- Lecture [2.5](#) will cover solution methods for rational expectations problems, including method of undetermined coefficients, iterative projection methods, and stable-unstable decoupling. We will solve by imposing model-consistent expectations.
- Lecture [2.6](#) will cover fluctuations, welfare, and efficiency, meaning micro-founded welfare criteria and level, composition, and production efficiency, which gives us a simple but coherent framework to study what policy should achieve.
- Lecture [2.7](#) will talk about policy trade-offs, where we have multi-dimensional policy objectives but only a single instrument.
- Lecture [2.8](#) will introduce sticky wages, and talk about monetary policy design with sticky wages.
- Lecture [2.9](#) will introduce unemployment and monetary policy when there is unemployment – how unemployment will change welfare incentives.

- Lecture 2.10 will take the model to the data, where we write the linear models in state space form, estimating latent states using the Kalman filter.
- Lecture 2.11 will talk about calibration and matching moments, largely thinking about how we choose parameter values for a model.
- Lecture 2.12 will conclude with likelihood-based estimation, meaning numerical optimization with the models as likelihood functions, thinking of choosing parameters as a form of maximum likelihood estimation.

We will finish today by doing something tedious but useful: linear difference equations.

First Order Scalar Difference Equations. We will write

$$x_t = c + \rho \cdot x_{t-1} + \varepsilon_t$$

where ε_t is a white noise shock process with mean 0 and variance σ^2 . We can recursively write these equations as a function of the initial state x_0 and the sequence of shocks, so

$$\begin{aligned} x_1 &= c + \rho \cdot x_0 + \varepsilon_1 \\ x_2 &= c + \rho \cdot (c + \rho \cdot x_0 + \varepsilon_1) + \varepsilon_2 \\ &= c(1 + \rho) + \rho^2 \cdot x_0 + \rho \cdot \varepsilon_1 + \varepsilon_2 \\ \Rightarrow x_t &= c + \rho \cdot c + \dots + \rho^{t-1} \cdot c + \rho^t \cdot x_0 + \rho^{t-1} \cdot \varepsilon_1 + \dots + \rho \cdot \varepsilon_{t-1} + \varepsilon_t \\ &= c \frac{1 - \rho^t}{1 - \rho} + \rho^t \cdot x_0 + \sum_{s=1}^t \rho^{t-s} \cdot \varepsilon_s \end{aligned}$$

So when $|\rho| < 1$, the initial state becomes unimportant for sufficiently long time horizons, since $\lim_{t \rightarrow \infty} \rho^t = 0$.

Impulse Response Functions. The impulse response function describes the effect on x_{t+s} of a unit change in ε_t . From the expression for x_t above, we get that

$$\frac{\partial x_{t+s}}{\partial \varepsilon_t} = \rho^s$$

Higher Order Difference Equations. A p^{th} order difference equation of the form

$$x_t = c + \rho_1 \cdot x_{t-1} + \rho_2 \cdot x_{t-2} + \dots + \rho_p \cdot x_{t-p} + \varepsilon_t$$

It is often easier to rewrite this into an equivalent vector-valued difference equation, where we define the vectors ξ_t , \mathbf{c} , and \mathbf{v}_t , and the matrix F as

$$\xi_t = \mathbf{c} + F \cdot \xi_{t-1} + \mathbf{v}_t$$

where

$$\xi_t = \begin{bmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_{t-p+1} \end{bmatrix} \quad ; \quad \mathbf{c} = \begin{bmatrix} c \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad ; \quad F = \begin{bmatrix} \rho_1 & \rho_2 & \dots & \rho_p \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \quad ; \quad \mathbf{v}_t = \begin{bmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

This equation is now the same as the p^{th} order difference equation as above, and we can rewrite it recursively as before, where

$$\xi_t = (I - F^{t-s})(I - F)^{-1} \cdot \mathbf{c} + F^t \cdot \xi_0 + \sum_{s=0}^t F^{t-s} \cdot \mathbf{v}_s$$

As long as $\max |\text{eig}(F)| < 1$, we can again ignore ξ_0 for sufficiently large t . The impulse response function is now given by

$$\frac{\partial x_{t+s}}{\partial \varepsilon_t} = e_1' F^s e_1$$

where $e_1 = [1 \quad 0 \quad \dots \quad 0]'$

Variances. The variance of a process x_t is defined as $\text{Var}(x_t) = \mathbb{E}(x_t - \mu)^2$, where $\mu = \mathbb{E}x_t = \frac{c}{1-\rho}$. We thus have that

$$\mathbb{E}(x_t - \mu)^2 = \sigma^2(\rho^2 + \rho^4 + \dots) = \frac{\sigma^2}{1 - \rho^2}$$

Alternatively, we could define $\sigma_x^2 = \text{Var}(x_t)$, and get that taking the variance of the most simple form of x_t , we have

$$\sigma_x^2 = \rho^2 \sigma_x^2 + \sigma^2 \implies \sigma_x^2 = \frac{\sigma^2}{1 - \rho^2}$$

For vector-valued processes ξ_t , we cannot solve directly for the covariance, but defining $\text{cov}(\xi_t, \xi_t) = \Sigma_\xi$ and using that ξ_t and \mathbf{v}_t are uncorrelated, we get that

$$\Sigma_\xi = F \Sigma_\xi F' + \mathbb{E}(\mathbf{v}_t \mathbf{v}_t')$$

which has a closed-form solution as long as the largest eigenvalue of F is less than 1, and can be found using function iteration.

2.2 A Classical Monetary Model

Remark. The plan is to present a simple ‘real’ economy, derive equilibrium, and introduce monetary policy.

Model. *Classical Monetary Model*

Assumption 2.1. *We have two types of agents: a representative household, and firms. We have perfect competition in goods and labor markets, flexible prices and wages, no capital accumulation, and no fiscal sector. The economy is closed.*

The representative household makes two decisions: how much labor to supply, and how much income to consume versus save. We also assume, as usual, that the representative household owns the firms. They solve:

$$\max \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t U(C_t, N_t) \right]$$

subject to, for all t ,

$$P_t C_t + Q_t B_t \leq B_{t-1} + W_t N_t + D_t$$

and the solvency constraint

$$\lim_{T \rightarrow \infty} \mathbb{E}_t [\Lambda_{t,T} (B_T / P_T)] \geq 0$$

where $\Lambda_{t,T} \equiv \beta^{T-t} U_{c,T} / U_{c,t}$ is the stochastic discount factor. We have two optimality conditions:

$$\begin{aligned} -U_{n,t} &= \frac{W_t}{P_t} U_{c,t} \\ U_{c,t} &= \frac{\beta}{Q_t} \mathbb{E}_t \left[U_{c,t+1} \frac{P_t}{P_{t+1}} \right] \end{aligned}$$

The intuition of the first is that the additional disutility of working more today must be exactly offset by the utility of the consumption that the additional income could buy, meaning that the labor supply decision responds to the relative price of leisure against consumption. The intuition of the second is that the additional utility of consuming more today must be equal to the additional utility of consuming more tomorrow, controlling for everything else – *i.e.*, people smooth.

We further assume that households have CRRA utility functions that are *separable* in consumption and labor:

$$U(C_t, N_t) = \begin{cases} \frac{C_t^{1-\sigma}-1}{1-\sigma} - \frac{N_t^{1+\varphi}}{1+\varphi} & \text{for } \sigma \neq 1 \\ \log C_t - \frac{N_t^{1+\varphi}}{1+\varphi} & \text{for } \sigma = 1 \end{cases}$$

When we have these explicit forms, the optimality conditions become:

$$\begin{aligned} N_t^\varphi &= C_t^{-\sigma} \frac{W_t}{P_t} \\ Q_t &= \beta \mathbb{E}_t \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} \frac{P_t}{P_{t+1}} \right] \end{aligned}$$

These admit the labor supply decision

$$w_t - p_t = \sigma c_t + \varphi n_t$$

and the consumption Euler equation

$$c_t = \mathbb{E}[c_{t+1}] - \frac{1}{\sigma} (i_t - \mathbb{E}_t[\pi_{t+1}] - \rho)$$

where $\pi_t \equiv p_t - p_{t-1}$, $i_t \equiv -\log Q_t$, and $\rho \equiv -\log \beta$. In steady state, with zero growth, we have that $i = \pi + \rho$, so the implied real interest rate r is $r \equiv i - \pi = \rho$, which is the (log of the) inverse of the discount rate.

Firms hire labor from households to produce a uniform good using the technology

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

where $a_t \equiv \log A_t$ follows an exogenous process

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

The firm's profit is, as always, the difference between revenue and cost. Maximizing profits subject to the above while taking the price and wage as given results in the optimality condition

$$\frac{W_t}{P_t} = (1 - \alpha) A_t N_t^{-\alpha}$$

that equates the marginal product of labor with the real marginal cost. In log-linear terms,

$$w_t - p_t = a_t - \alpha n_t + \log(1 - \alpha)$$

Definition. *Equilibrium* requires market clearing in: (i) goods, $y_t = c_t$; (ii) household labor supply $w_t - p_t = \sigma c_t + \varphi n_t$; (iii) firm labor demand $w_t - p_t = a_t - \alpha n_t + \log(1 - \alpha)$; (iv) aggregate output $y_t = a_t + (1 - \alpha)n_t$; and (v) asset market clearing $r_t = i_t - \mathbb{E}[\pi_{t+1}] = \rho + \sigma \mathbb{E}[\Delta c_{t+1}]$.

We attain the implied equilibrium values for real variables as functions of productivity

$$\begin{aligned} n_t &= \psi_{na} a_t + \psi_n \\ y_t &= \psi_{ya} a_t + \psi_y \\ r_t &= \rho - \sigma \psi_{ya} (1 - \rho_a) a_t \\ \omega_t &\equiv w_t - p_t = a_t - \alpha n_t + \log(1 - \alpha) = \psi_{\omega a} a_t + \psi_{\omega} \end{aligned}$$

where

$$\begin{aligned} \psi_{na} &\equiv \frac{1 - \sigma}{\sigma(1 - \alpha) + \varphi + \alpha} ; \psi_n \equiv \frac{\log(1 - \alpha)}{\sigma(1 - \alpha) + \varphi + \alpha} ; \psi_{ya} \equiv \frac{1 + \varphi}{\sigma(1 - \alpha) + \varphi + \alpha} ; \psi_y \equiv (1 - \alpha)\psi_n \\ \psi_{\omega a} &\equiv \frac{\sigma + \varphi}{\sigma(1 - \alpha) + \varphi + \alpha} ; \psi_{\omega} \equiv \frac{(\sigma(1 - \alpha) + \varphi) \log(1 - \alpha)}{\sigma(1 - \alpha) + \varphi + \alpha} \end{aligned}$$

We also have *nominal neutrality*, where real variables are determined independently of monetary policy, that *optimal monetary policy is undetermined*, meaning that inflation does not affect welfare, and that *price level is undetermined*, meaning that a rule for money supply or the nominal interest rate is needed.

Definition. We also have a simple *interest rate rule*, where nominal interest rate increases with inflation, so

$$i_t = \rho + \pi + \phi_{\pi}(\pi_t - \pi) + v_t$$

where $\phi_{\pi} \geq 0$. We can combine with the *Fischer equation* $r_t = i_t - \mathbb{E}_t[\pi_{t+1}]$ to get

$$\phi_t \hat{\pi}_t = \mathbb{E}_t[\hat{\pi}_{t+1}] + \hat{r}_t - v_t$$

where $\hat{r}_t \equiv r_t - \rho$ and $\hat{\pi}_t \equiv \pi_t - \pi$. If $\phi_{\pi} > 1$, then

$$\hat{\pi}_t = \sum_{k=0}^{\infty} \phi_{\pi}^{-(k+1)} \mathbb{E}_t[\hat{r}_{t+k} - v_{t+k}] = -\frac{\sigma(1 - \rho_a)\psi_{ya}}{\phi_{\pi} - \rho_a} a_t - \frac{1}{\phi_{\pi}} v_t$$

Remark. In summation, the representative household trades off (i) leisure against consumption; and (ii) consumption today against consumption tomorrow. Firms hire labor until marginal cost equals marginal revenue, and monetary policy does not affect real variables. We should know how to set up and solve the model and how parameters affect how the economy responds to exogenous variables.

2.3 Basic New Keynesian Model I

Today, we will talk about *Dixit-Stiglitz (CES) Demand Systems* (from Dixit & Stiglitz, 1977), the basic *New Keynesian Business Cycle Model*, and some sources of inefficiency.

Remark. Empirical evidence shows that monetary policy shocks have persistent effects on real variables, prices that adjust slowly, and a liquidity effect. These stylized facts are in conflict with the predictions of the classical monetary models. Micro evidence suggests that there may be significant price and wage rigidities. For the stylized facts, see [Leeper, Sims, and Zha \(1996\)](#), and for the micro evidence, see [Alvarez et al. \(2006\)](#).

Model. *Basic New Keynesian Model (Outline)* We have a goods market and a labor market. In the goods market, households consume a basket of goods, and firms produce different consumption goods. Prices have a fixed probability of resetting their prices as in [Calvo \(1983\)](#). In the labor market, firms hire labor and households supply labor. Households also optimally invest in a one-period riskless bond.

Definition. The *New Keynesian Phillips Curve* is

$$\pi_t = \beta \mathbb{E}_t[\pi_{t+1}] + \kappa \tilde{y}_t$$

The *Dynamic IS Equation* is

$$\tilde{y}_t = \mathbb{E}_t[\tilde{y}_{t+1}] - \frac{1}{\sigma} (i_t - \mathbb{E}_t[\pi_{t+1}] - r_t^n)$$

The *Monetary Policy Rule* is

$$i_t = \rho + \phi_\pi \pi_t + \phi_y \tilde{y}_t + v_t$$

Model. *Dixit-Stiglitz Demand Systems* Consider the utility function $U(C) = \frac{C^{1-\sigma}}{1-\sigma}$, where C is a *CES aggregator* over a continuum of goods,

$$C \equiv \left(\int_0^1 C_i^{\frac{\varepsilon-1}{\varepsilon}} \partial i \right)^{\frac{\varepsilon}{\varepsilon-1}} \quad \text{for } \varepsilon > 1, i \in (0, 1)$$

For finite ε , goods are imperfect substitutes and firms therefore have some market power over the pricing of goods. We want to maximize C subject to the budget constraint

$$\int_0^1 P_i C_i \partial i \leq R$$

We can set up the Lagrangian

$$\max_{C_i} \left(\int_0^1 C_i^{\frac{\varepsilon-1}{\varepsilon}} \partial i \right)^{\frac{\varepsilon}{\varepsilon-1}} - \lambda \left(\int_0^1 P_i C_i \partial i - R \right)$$

and taking first order conditions, we get

$$\frac{\varepsilon}{\varepsilon-1} \frac{\varepsilon-1}{\varepsilon} \left(\int_0^1 C_i^{\frac{\varepsilon-1}{\varepsilon}} \partial i \right)^{\frac{\varepsilon}{\varepsilon-1}-1} C_i^{\frac{\varepsilon-1}{\varepsilon}-1} = \lambda P_i$$

Using the fact that $C = \left(\int_0^1 C_i^{\frac{\varepsilon-1}{\varepsilon}} \partial i \right)^{\frac{\varepsilon}{\varepsilon-1}}$, we can rewrite this as

$$C^{\frac{1}{\varepsilon}} C_i^{-\frac{1}{\varepsilon}} = \lambda P_i \implies C - i = (\lambda P_i)^{-\varepsilon} C$$

It remains to eliminate λ . If we multiply both sides by C_i and integrate over i , we get that

$$C^{\frac{1}{\varepsilon}+1-\frac{1}{\varepsilon}} = \lambda \int_0^1 P_i C_i \partial i$$

and from the definition of $P \equiv C^{-1} \int_0^1 P_i C_i \partial i$ implies that $p = \lambda^{-1}$, so

$$C_i = \left(\frac{P_i}{P} \right)^{-\varepsilon} C$$

Question. Given this demand, what is the optimal price to charge?

Remark. Consider firm i facing the above demand curve. The maximization problem can be written as

$$\max_{P_i} P_i Y_i - C(Y_i)$$

where $\mathcal{C}(Y_i)$ is nominal cost as a function of output $Y_i = C_i$. We get the first order condition

$$Y_i + P_i \frac{\partial Y_i}{\partial P_i} - \frac{\partial \mathcal{C}}{\partial Y_i} \frac{\partial Y_i}{\partial P_i} = 0$$

Defining the nominal marginal cost as $\psi_i = \frac{\partial \mathcal{C}}{\partial Y_i}$, we can simplify this to

$$1 + \frac{P_i}{Y_i} \frac{\partial Y_i}{\partial P_i} - \psi_i \frac{1}{Y_i} \frac{\partial Y_i}{\partial P_i} = 0 \implies P_i = \frac{\varepsilon}{\varepsilon - 1} \psi_i$$

using the definition of elasticities. This gives the optimal price as the markup $\mathcal{M} = \frac{\varepsilon}{\varepsilon - 1} > 1$ times nominal marginal cost.

Model. *The Basic New Keynesian Model* The representative household solves

$$\max \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t U(C_t, N_t; Z_t) \right]$$

where

$$U(C_t, N_t; Z_t) = \left(\frac{C_t^{1-\sigma} - 1}{1-\sigma} - \frac{N_t^{1+\varphi}}{1+\varphi} \right) Z_t$$

and

$$C_t = \left(\int_0^1 C_t(i)^{1-\frac{1}{\varepsilon}} \partial i \right)^{\frac{\varepsilon}{\varepsilon-1}}$$

subject to

$$\int_0^1 P_t(i) C_t(i) \partial i + Q_t B_t \leq B_{t-1} + W_t N_t + D_t$$

for all t . We will allocate expenditure across different goods as

$$C_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\varepsilon} C_t \implies c_t(i) = -\varepsilon(p_t(i) - p_t) + c_t$$

labor supply will meet

$$-\frac{U_{n,t}}{U_{c,t}} = \frac{W_t}{P_t} \implies w_t - p_t = \sigma c_t + \varphi n_t$$

and intertemporal consumption

$$Q_t = \beta \mathbb{E}_t \left[\frac{U_{c,t+1}}{U_{c,t}} \frac{P_t}{P_{t+1}} \frac{Z_{t+1}}{Z_t} \right] \implies c_t = \mathbb{E}_t[c_{t+1}] - \frac{1}{\sigma} (i_t - \mathbb{E}_t[\pi_{t+1}] - \rho) + \frac{1}{\sigma} (1 - \rho_z) z_t$$

where $i_t = -\log Q_t$ and $\rho = -\log \beta$. We have exogenous demand shocks $z_t = \rho_z z_{t-1} + \varepsilon_t^z$.

Finally, we have a continuum of firms indexed by $i \in (0, 1)$, where each firm produces a differentiated good with identical technology $Y_t(i) = A_t N_t(i)^{1-\alpha}$, where $a_t = \rho_a a_{t-1} + \varepsilon_t^a$. The probability of getting to reset the price in any period is $1 - \theta$, with the implied average price duration $\frac{1}{1-\theta}$.

Remark. We will think about two sources of inefficiency: (i) decreasing marginal utility of individual goods (love of variety); and (ii) decreasing returns to scale in production. For the first, consider the CES aggregator with only two goods:

$$C_t \equiv \left(\frac{1}{2} \sum_{i=1}^2 C_t(i)^{1-\frac{1}{\varepsilon}} \partial i \right)^{\frac{\varepsilon}{\varepsilon-1}}$$

Clearly, the consumer would prefer $\{C_t(1), C_t(2)\}$ to $\{2 \cdot C_t(1)\}$. For the second, consider the production

function with common productivity A_t :

$$Y_t(i) = A_t N_t(i)^{1-\alpha}$$

Clearly, the bundle $\{Y_t(1), Y_t(2)\}$ takes less labor inputs to produce than $\{2 \cdot Y_t(1)\}$.

Remark. In summary, we introduced monopolistic competition with household preferences, through a CES demand system where optimal price is a fixed markup over marginal cost, and explored two potential sources of inefficiency. From this part, we should know (i) how to derive demand for individual goods; (ii) how to derive the CES price index; (iii) how to find the optimal price P_i ; and (iv) the economic forces that determine efficiency in allocation of labor.

2.4 Basic New Keynesian Model II

If a fraction $(1 - \theta)$ of firms set price P_t^* , the aggregate price level will follow

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

which can be rearranged to

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

Log-linearization around the zero-inflation steady state gives

$$\pi_t = (1 - \theta)(p_t^* - p_{t-1}) \equiv p_t = \theta p_{t-1} + (1 - \theta)p_t^*$$

Using the stochastic discount factor of the households, firms maximize expected discounted profits

$$\max_{P_t^*} \sum_{k=0}^{\infty} \theta^k \mathbb{E}_t \left\{ \Lambda_{t,t+k} (1/P_{t+k}) (P_t^* \cdot Y_{t+k|t} - C_{t+k}(Y_{t+k|t})) \right\}$$

subject to

$$Y_{t+k|t} = \left(\frac{P_t^*}{P_{t+k}} \right)^{-\varepsilon} Y_{t+k} \quad \forall k = 0, 1, \dots$$

Note that the firm can only affect the rightmost term in the maximization problem, not the given prices or stochastic discount factor. The first order condition of the price-setting problem is

$$\sum_{k=0}^{\infty} \theta^k \mathbb{E}_t \left\{ \Lambda_{t,t+k} Y_{t+k|t} (1/P_{t+k}) (P_t^* - \mathcal{M} \Psi_{t+k|t}) \right\} = 0$$

where $\Psi_{t+k|t} = \mathcal{C}'_{t+k}(Y_{t+k|t})$ is the nominal marginal cost of a firm producing a good in period $t + k$ with price P_t^* , and $\mathcal{M} = \frac{\varepsilon}{\varepsilon-1}$ is the desired markup. Note that if $\theta = 0$, this equation simplifies to $P_t^* = \mathcal{M} \Psi_{t|t}$. The linearized optimal price setting condition around a zero-inflation steady state is given by

$$p_t^* = \mu + (1 - \beta\theta) \sum_{k=0}^{\infty} (\beta\theta)^k \mathbb{E}_t \{\psi_{t+k|t}\}$$

where $\mu = \log \mathcal{M}$ and $\psi = \log \Psi$. Essentially, firms are looking to set the price that results in the optimal markup, weighted to the discount factor and the probability of the price still being in place k periods in the future. The nominal marginal cost is the wage divided by the marginal productivity of labor, so

$$\psi_{t+k|t} = w_{t+k} - (a_{t+k} - \alpha \cdot n_{t+k|t} + \log(1 - \alpha))$$

We can define the component of marginal cost that is constant across firms (so not dependent on the price set today) as

$$\psi_{t+k} = w_{t+k} - (a_{t+k} - \alpha \cdot n_{t+k} + \log(1 - \alpha))$$

Thus, we have that

$$\psi_{t+k|t} = \psi_{t+k} + \alpha(n_{t+k|t} - n_{t+k}) = \psi_{t+k} + \frac{\alpha\varepsilon}{1-\alpha}(p_t^* - p_{t+k})$$

Define the marginal cost as $\widehat{mc}_t = \psi_t - p_t + \mu$, and substituting into the optimal price setting equation we get

$$p_t^* = (1 - \beta\theta) \sum_{k=0}^{\infty} (\beta\theta)^k \mathbb{E}_t \left\{ \widehat{mc}_{t+k} - \frac{\alpha\varepsilon}{1-\alpha}(p_t^* - p_{t+k}) + p_{t+k} \right\}$$

Simplifying, this becomes

$$p_t^* = (1 - \beta\theta) \sum_{k=0}^{\infty} (\beta\theta)^k \mathbb{E}_t \left\{ \frac{1-\alpha}{1-\alpha+\alpha\varepsilon} \widehat{mc}_{t+k} + p_{t+k} \right\}$$

Writing recursively, we get

$$p_t^* = \beta\theta \mathbb{E}_t p_{t+1}^* + (1 - \beta\theta) \left(\frac{1-\alpha}{1-\alpha+\alpha\varepsilon} \widehat{mc}_t + p_t \right)$$

Subtracting p_{t-1} from both sides and adding and subtracting $\beta\theta p_t$, and using the fact that $p_t^* - p_{t-1} = (1 - \theta)^{-1} \pi_t$, we get

$$(1 - \theta)^{-1} \pi_t = (1 - \theta)^{-1} \beta\theta \mathbb{E}_t \{\pi_{t+1}\} + (1 - \beta\theta) \left(\frac{1-\alpha}{1-\alpha+\alpha\varepsilon} \widehat{mc}_t \right) + \pi_t$$

so we get that

$$\pi_t = \beta\mathbb{E}_t \{\pi_{t+1}\} + \frac{(1-\theta)(1-\beta\theta)}{\theta} \frac{1-\alpha}{1-\alpha+\alpha\varepsilon} \widehat{mc}_t$$

The real marginal cost is the real wage divided by the marginal productivity of labor, so

$$mc_t = \psi_t - p_t = \left(\sigma + \frac{\varphi + \alpha}{1 - \alpha} \right) y_t - \left(\frac{1 + \varphi}{1 - \alpha} \right) a_t - \log(1 - \alpha)$$

So real marginal cost is a function of both output and productivity. If we have flexible prices (where $\theta = 0$), the markup is constant and equal to negative marginal cost:

$$mc = -\mu = \left(\sigma + \frac{\varphi + \alpha}{1 - \alpha} \right) y_t^n - \left(\frac{1 + \varphi}{1 - \alpha} \right) a_t - \log(1 - \alpha)$$

where y_t^n is the natural level of output which we can solve for:

$$y_t^n = \frac{1 + \varphi}{\sigma(1 - \alpha) + \varphi + \alpha} a_t = \frac{(1 - \alpha)(\mu - \log(1 - \alpha))}{\sigma(1 - \alpha) + \varphi + \alpha}$$

The deviation of marginal cost from the steady state is equal to the output gap:

$$\widehat{mc}_t = \left(\sigma + \frac{\varphi + \alpha}{1 - \alpha} \right) (y_t - y_t^n)$$

If we combine this with the (previously assumed) goods and labor market clearing conditions, we can finally derive the *New Keynesian Phillips Curve*

$$\pi_t = \beta \mathbb{E}_t \{\pi_{t+1}\} + \kappa \tilde{y}_t$$

where $\tilde{y}_t = y_t - y_t^n$ and $\kappa = \left(\sigma + \frac{\varphi + \alpha}{1 - \alpha}\right) \frac{(1 - \theta)(1 - \beta\theta)}{\theta} \frac{1 - \alpha}{1 - \alpha + \alpha\epsilon}$. We also have the *Dynamic IS equation*

$$\tilde{y}_t = \mathbb{E}_t \{\tilde{y}_{t+1}\} - \frac{1}{\sigma} (i_t + \mathbb{E}_t \{\pi_{t+1}\} - r_t^n)$$

where r_t^n is the *natural rate of interest*, given by $r_t^n = \rho - \sigma(1 - \rho_a)\psi_{ya}a_t + (1 - \rho_z)z_t$. We can finally close the model by writing a Taylor-type rule for the nominal interest rate, where

$$i_t = \rho - \phi_\pi \pi_t + \phi_y \hat{y}_t + v_t$$

where $\hat{y}_t = y_t - y$ and $v_t = \rho_v v_{t-1} + \varepsilon_t^v$. We can use the method of undetermined coefficients to solve for inflation and the output gap, which we'll be able to write as a function of the composite shocks. Next time, we will talk more about the dynamics of the solved model.

2.5 Rational Expectations Models

Remark. There are three different ways of solving linear rational expectations models that we will discuss:

1. The method of undetermined coefficients, which can be very quick when feasible and illustrates the fixed point nature of the rational expectations solution
2. Decouple the stable and unstable dynamics of the model and set the unstable part to zero
3. Replacing expectations with linear projections onto observable variables

We will discuss methods (1) and (2) today.

As a vehicle to demonstrate the different solution methods, we will use the basic New Keynesian model:

$$\begin{aligned} \pi_t &= \beta \mathbb{E}_t \{\pi_{t+1}\} + \kappa \tilde{y}_t \\ \tilde{y}_t &= \mathbb{E}_t \{\tilde{y}_{t+1}\} - \frac{1}{\sigma} (i_t - \mathbb{E}_t \{\pi_{t+1}\} - r_t^n) \\ i_t &= \rho + \phi_\pi \pi_t + \phi_y \hat{y}_t + v_t \\ r_t^n &= \rho - \sigma(1 - \rho_a)\psi_{ya}a_t + (1 - \rho_z)z_t \end{aligned}$$

where $\pi_t, \tilde{y}_t, \hat{y}_t, i_t, r_t^n$ are inflation, output gap, output deviation from steady state, nominal interest rate, and the natural rate of interest respectively.

Example. *Method of Undetermined Coefficients*

Remark. This method is quick when feasible and illustrates well the fixed point nature of rational expectations equilibria but is difficult to implement in large models.

We start by substituting the interest rate into the IS equation to get

$$\tilde{y}_t = \mathbb{E}_t \{\tilde{y}_{t+1}\} - \frac{1}{\sigma} (\rho + \phi_\pi \pi_t + \phi_y \hat{y}_t + v_t - \mathbb{E}_t \{\pi_{t+1}\} - r_t^n)$$

and define the composite shock u_t as

$$u_t \equiv r_t^n - \phi_y \hat{y}_t^n - v_t = -\psi_{ya}(\phi_y + \sigma(1 - \rho_a))a_t + (1 - \rho_z)z_t - v_t$$

and use the fact that

$$\hat{y}_t = \tilde{y}_t + \hat{y}_t^n \quad ; \quad \hat{y}_t^n = \psi_y a_t$$

to simplify the IS equation to

$$\tilde{y}_t = \mathbb{E}_t\{\tilde{y}_{t+1}\} - \frac{1}{\sigma} (\phi_\pi \pi_t + \phi_y \tilde{y}_t - \mathbb{E}_t\{\pi_{t+1}\} + u_t)$$

By assuming that the composite shock is an AR(1) process with persistence parameter ρ_u , so $u_t = \rho_u u_{t-1} + \eta_t$, we can write all endogenous variables as functions only of u_t . We conjecture that the model can be put in the form

$$\pi_t = \psi_\pi u_t \quad ; \quad \tilde{y}_t = \psi_y u_t$$

so that

$$\mathbb{E}_t\{\pi_{t+1}\} = \psi_\pi \rho_u u_t \quad ; \quad \mathbb{E}_t\{\tilde{y}_{t+1}\} = \psi_y \rho_u u_t$$

Solving the model implies finding the coefficients ψ_π and ψ_y .

We can substitute the conjectured solution into the structural Phillips curve and IS equation

$$\begin{aligned} \psi_\pi u_t &= \beta \psi_\pi \rho_u u_t + \kappa \psi_y u_t \\ \psi_y u_t &= \psi_y \rho_u u_t - \frac{1}{\sigma} (\phi_\pi \psi_\pi u_t + \phi_y \psi_y u_t - \psi_\pi \rho_u u_t + u_t) \end{aligned}$$

Equating coefficients on the left and right hand sides, we get

$$\begin{aligned} 0 &= \psi_\pi - \beta \psi_\pi \rho_u - \kappa \psi_y \\ -\frac{1}{\sigma} &= \psi_y - \psi_y \rho_u + \frac{1}{\sigma} (\phi_\pi \psi_\pi + \phi_y \psi_y - \psi_\pi \rho_u) \end{aligned}$$

which is a system of linear equations in ψ_π and ψ_y :

$$\begin{bmatrix} (1 - \beta \rho_u) & -\kappa \\ (\frac{1}{\sigma} \phi_\pi - \frac{1}{\sigma} \rho_u) & (1 - \rho_u + \frac{1}{\sigma} \phi_y) \end{bmatrix} \cdot \begin{bmatrix} \psi_\pi \\ \psi_y \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{1}{\sigma} \end{bmatrix}$$

We can multiply both sides by the inverse of the coefficient matrix, so

$$\begin{bmatrix} \psi_\pi \\ \psi_y \end{bmatrix} = \Lambda \begin{bmatrix} -\kappa \\ (\beta \rho_u - 1) \end{bmatrix}$$

where Λ is the determinant of the coefficient matrix.

Example. *Stable/Unstable Decoupling* Originally due to [Blanchard & Kahn \(1980\)](#), and has been developed significantly since then. See for example [Söderlind \(1999\)](#). This method has several advantages: it's fast, provides conditions for when a solution exists, and provided conditions for when a solution is unique.

We begin by putting the model in matrix form:

$$A_0 \begin{bmatrix} x_{t+1}^1 \\ \mathbb{E}_t\{x_{t+1}^2\} \end{bmatrix} = A_1 \begin{bmatrix} x_t^1 \\ x_t^2 \end{bmatrix} + C_1 u_{t+1}$$

where x_t^1 is a vector of the control variables and x_t^2 is a vector of the jump variables. The respective matrices are:

$$x_t^1 = [a_t \quad z_t \quad u_t]^T \quad ; \quad x_t^2 = [\pi_t \quad \tilde{y}_t] \quad ; \quad C_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$A_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \beta & 0 \\ 0 & 0 & 0 & \frac{1}{\sigma} & 1 \end{bmatrix} ; \quad A_1 = \begin{bmatrix} \rho_a & 0 & 0 & 0 & 0 \\ 0 & \rho_z & 0 & 0 & 0 \\ 0 & 0 & \rho_u & 0 & 0 \\ 0 & 0 & 0 & 1 & -\kappa \\ \frac{1}{\sigma}(\phi_y\psi_{ya} + (1-\rho_a)\psi_{ya}) & -\frac{1}{\sigma}(1-\rho_z) & \frac{1}{\sigma} & \frac{1}{\sigma}\phi_\pi & (1 + \frac{1}{\sigma}\phi_y) \end{bmatrix}$$

With some linear algebra, we convert this to

$$\begin{bmatrix} x_{t+1}^1 \\ \mathbb{E}_t\{x_{t+1}^2\} \end{bmatrix} = A \begin{bmatrix} x_t^1 \\ x_t^2 \end{bmatrix} + Cu_{t+1}$$

where $A = A_0^{-1}A_1$ and $C = A_0^{-1}C_1$. For the model to have a unique stable solution, the number of stable eigenvalues of A must be equal to the number of exogenous (control) variables.

We will use a Schur decomposition to get $A = ZTZ^H$ where T is upper block triangular (*i.e.* $T = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix}$) and Z is a unitary matrix so that $Z^H Z = Z Z^H = I$, where the H exponent denotes the [Hermitian operator](#). For any square matrix W , $W^{-1}AW$ is called a [similarity transformation](#) of A . Similarity transformations do not change the eigenvalues of a matrix, and it is a property of the Schur decomposition that it is always possible to choose Z and T so that the unstable eigenvalues of A are shared with T_{22} (the lower right block of T).

We define the auxiliary variables

$$\begin{bmatrix} \gamma_t \\ \delta_t \end{bmatrix} = Z^H \begin{bmatrix} x_t^1 \\ x_t^2 \end{bmatrix}$$

and we can rewrite the system as

$$Z^H \begin{bmatrix} x_{t+1}^1 \\ \mathbb{E}_t\{x_{t+1}^2\} \end{bmatrix} = Z^H Z T Z^H \begin{bmatrix} x_t^1 \\ x_t^2 \end{bmatrix} \equiv \mathbb{E} \begin{bmatrix} \gamma_{t+1} \\ \delta_{t+1} \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} \begin{bmatrix} \gamma_t \\ \delta_t \end{bmatrix}$$

since $Z^H Z = I$.

For this system to be stable, the auxiliary variables associated with the unstable roots in T_{22} must be zero for all t . Imposing $\delta_t = 0 \forall t$ reduces the relevant state dynamics to $\gamma_t = T_{11}\gamma_{t-1}$, and to get back the original variables we simply use that

$$\begin{bmatrix} x_t^1 \\ x_t^2 \end{bmatrix} = \begin{bmatrix} Z_{11} \\ Z_{21} \end{bmatrix} \gamma_t \equiv \begin{bmatrix} x_t^1 \\ x_t^2 \end{bmatrix} = \begin{bmatrix} Z_{11} \\ Z_{21} \end{bmatrix} Z_{11}^{-1} x_t^1$$

which is the solution to the model. The solved model is of the form

$$\begin{aligned} x_t^1 &= Mx_{t-1}^1 + Cu_t \\ x_t^2 &= Gx_t^1 \end{aligned}$$

where

$$M = Z_{11}T_{11}Z_{11}^{-1} = \begin{bmatrix} \rho_a & 0 & 0 \\ 0 & \rho_z & 0 \\ 0 & 0 & \rho_u \end{bmatrix}$$

and $G = Z_{21}Z_{11}^{-1}$.

Remark. There are slides on how to implement this in Matlab in Kris' files. I don't use Matlab so that material is omitted here.

2.6 Monetary Policy

Recall that we have two sources of inefficiency: market power of individual firms in the goods market, and nominal rigidities. The efficient allocation solves

$$\max U(C_t, N_t; Z_t) \text{ s.t. } C_t(i) = A_t N_t(i)^{1-\alpha} \quad \text{and} \quad N_t = \int_0^1 N_t(i) \partial i$$

which admits optimality conditions

$$C_t(i) = C_t \quad ; \quad N_t(i) = N_t \quad ; \quad -\frac{U_{n,t}}{U_{c,t}} = MPN_t = (1-\alpha)A_t N_t^{-\alpha}$$

Under flexible prices the optimal price is marginal cost times a fixed markup $P_t = \mathcal{M} \frac{W_t}{MPN_t}$ for a fixed markup $\mathcal{M} = \frac{\varepsilon}{\varepsilon-1} > 1$ so

$$-\frac{U_{n,t}}{U_{c,t}} = \frac{W_t}{P_t} = \frac{MPN_t}{\mathcal{M}} < MPN_t$$

We could restore efficiency with an employment subsidy $\tau = \frac{1}{\varepsilon}$, so $\mathcal{M}(1-\tau) = 1$, since $P_t = \mathcal{M} \frac{W_t(1-\tau)}{MPN_t}$. With a constant employment subsidy that implies an efficient level of output under flexible prices, variation in mark-ups resulting from sticky prices are inefficient:

$$\mathcal{M}_t \equiv \frac{P_t}{(1-\tau)(W_t/MPN_t)} = \frac{P_t \mathcal{M}}{W_t/MPN_t} \implies -\frac{U_{n,t}}{U_{c,t}} = \frac{W_t}{P_t} = MPN_t \frac{\mathcal{M}}{\mathcal{M}_t} \neq MPN_t$$

Efficiency requires that average markup is the desired markup for all t .

We also have composition effects, which are relative price distortions resulting from staggered price setting, which could be either $C_t(i) \neq C_t(j)$ if $P_t(i) \neq P_t(j)$, decreasing marginal utility of individual goods, or decreasing marginal productivity of labor in a given firm. Since utility of, and production technology for making, all goods are symmetric across all firms, optimal policy requires that prices and quantities are equalized.

Model. For monetary policy, we assume the optimal (constant) employment subsidy as above, no inherited price distortions, meaning that $P_{-1}(i) = P_{-1}$ for all $i \in [0, 1]$, and only demand and productivity shocks (meaning no shocks that make the flex-price equilibrium inefficient).

The optimal monetary in this case replicated the flexible price equilibrium allocation – the central bank should stabilize marginal costs at a level consistent with firms' desired markup *at given existing prices*. Thus, no firm has an incentive to adjust its price, and equilibrium output and unemployment match their natural levels. Equilibrium under the optimal policy then implies that $y_t = y_t^n$, $\tilde{y}_t = 0$, $\pi_t = 0$, and $i_t = r_t^n$ for all t . This result is called *the divine coincidence*.

Inserting the exogenous interest rate rule $i_t = r_t^n$ into the non-policy block gives

$$\begin{aligned} \pi_t &= \beta \mathbb{E}_t \{\pi_{t+1}\} + \kappa \tilde{y}_t \\ \tilde{y}_t &= -\frac{1}{\sigma} (\mathbb{E}_t \{\pi_{t+1}\}) + \mathbb{E}_t \{\tilde{y}_{t+1}\} \end{aligned}$$

where $r_t^n = \rho - \sigma(1-\rho_a)\psi_{ya}a_t + (1-\rho_z)z_t$. Equilibrium dynamics with an exogenous optimal rule can be represented as

$$\begin{bmatrix} \tilde{y}_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{\sigma} \\ \kappa & \beta + \frac{\kappa}{\sigma} \end{bmatrix} \cdot \begin{bmatrix} \mathbb{E}_t \{\tilde{y}_{t+1}\} \\ \mathbb{E}_t \{\pi_{t+1}\} \end{bmatrix}$$

A shortcoming here: the solution $\tilde{y}_t = \pi_t = 0$ is not unique, as one eigenvalue of the coefficient matrix is strictly greater than one, meaning we have indeterminacy. We could resolve this indeterminacy by specifying

a rule that responds to inflation and output, of the form

$$i_t = r_t^n + \phi_\pi \pi_t + \phi_y \tilde{y}_t$$

Equilibrium dynamics become

$$\begin{bmatrix} \tilde{y}_t \\ \pi_t \end{bmatrix} = \frac{1}{\sigma + \phi_y + \kappa \phi_\pi} \begin{bmatrix} \sigma & 1 - \beta \phi_\pi \\ \sigma \kappa & \kappa + \beta(\sigma + \phi_y) \end{bmatrix} \begin{bmatrix} \mathbb{E}_t\{\tilde{y}_{t+1}\} \\ \mathbb{E}_t\{\pi_{t+1}\} \end{bmatrix}$$

Why does this work? An existence and uniqueness condition was provided by [Bullard & Mitra \(2002\)](#), of the form

$$\kappa(\phi_\pi - 1) + (1 - \beta)\phi_y > 0$$

and a Taylor-principle interpretation was provided by [Woodford \(2002\)](#):

$$\lim_{k \rightarrow \infty} \frac{\partial i_{t+k}}{\partial \pi_t} = \phi_\pi + \phi_y \lim_{k \rightarrow \infty} \frac{\partial \tilde{y}_{t+k}}{\partial \pi_t} = \phi_\pi + \frac{\phi_y(1 - \beta)}{\kappa}$$

Remark. Optimal rules assume that the natural rate of interest is observable in real time. This requires the policymaker to know (i) the true model, (ii) the true parameter values, and (iii) realized shocks, all of which are unrealistic. An operational alternative is a *simple rule*, which depends only on observable variables, and does not require knowledge of the true parameter values if they approximate the optimal rule across different models.

We can approximate the welfare of the representative household as

$$\mathbb{W} = -\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{U_t - U_t^n}{U_c C} \right) = -\frac{1}{2} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\left(\sigma + \frac{\varphi + \alpha}{1 - \alpha} \right) \tilde{y}_t^2 + \frac{\varepsilon}{\lambda} \pi_t^2 \right]$$

so that the expected average welfare loss per period is given by

$$\mathbb{L} = -\frac{1}{2} \left[\left(\sigma + \frac{\varphi + \alpha}{1 - \alpha} \right) \text{Var}(\tilde{y}_t) + \frac{\varepsilon}{\lambda} \text{Var}(\pi_t) \right]$$

We get from this the equivalent rules

$$i_t = \begin{cases} \rho + \phi_\pi \pi_t + \phi_y \hat{y}_t \\ \rho + \phi_\pi \pi_t + \phi_y \tilde{y}_t + v_t \end{cases}$$

where $v_t = \phi_y \hat{y}_t^n$ is a simple rule. Equilibrium dynamics become

$$\begin{bmatrix} \tilde{y}_t \\ \pi_t \end{bmatrix} = \Omega \begin{bmatrix} \sigma & 1 - \beta \phi_\pi \\ \sigma \kappa & \kappa + \beta(\sigma + \phi_y) \end{bmatrix} \begin{bmatrix} \mathbb{E}_t\{\tilde{y}_{t+1}\} \\ \mathbb{E}_t\{\pi_{t+1}\} \end{bmatrix} + \Omega \begin{bmatrix} 1 \\ \kappa \end{bmatrix} (\hat{r}_t^n - v_t)$$

and $\Omega = \frac{1}{\sigma + \phi_y + \kappa \phi_\pi}$ as before. Note that

$$\hat{r}_t^n - v_t = -\psi_{ya} (\sigma(1 - \rho_a) + \phi_y) a_t + (1 - \rho_z) z_t$$

If we take this to the data, we find that it is not optimal to stabilize output if fluctuations are mainly due to productivity shocks, but that strongly stabilizing inflation shocks leads to good outcomes regardless of whether fluctuations are due to inflation or demand shocks.

2.7 Policy Tradeoffs

Remark. In the simple New Keynesian model with only demand and productivity shocks there are no policy trade-offs. Strict inflation targeting is optimal even if we don't care about inflation *per se*. Implicitly, we here assume that the efficient level of output and the natural level of output coincide. What if the two don't coincide?

If output coincides with the natural level of output, there will be no inflation. If it coincides with the efficient level of output, there may be inflation but the condition

$$-\frac{U_{n,t}}{U_{c,t}} = MPN_t$$

will hold. If $y_t^e \neq y_t^n$, we need to modify the Phillips curve to

$$\pi_t = \beta \mathbb{E}_t\{\pi_{t+1}\} + \underbrace{\kappa x_t}_{y_t - y_t^e} + \underbrace{u_t}_{\kappa(y_t^e - y_t^n)}$$

The monetary policy problem is:

$$\min \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t (\pi_t^2 + \vartheta x_t^2)$$

subject to

$$\begin{aligned} x_t &= -\frac{1}{\sigma} (i_t - \mathbb{E}_t\{\pi_{t+1}\} - r_t^e) + \mathbb{E}_t\{x_{t+1}\} \\ \pi_t &= \beta \mathbb{E}_t\{\pi_{t+1}\} + \kappa x_t + u_t \end{aligned}$$

for each t , where $u_t = \rho_u u_{t-1} + \varepsilon_t$ and where $r_t^e = \rho - \sigma \mathbb{E}_t\{\Delta y_{t+1}^e\} + (1 - \rho_z)z_t$. Note that a utility-based criterion requires that $\vartheta = \frac{\kappa}{\varepsilon}$. This simplifies, so that each period the monetary policy authority chooses (x_t, π_t) to minimize

$$\pi_t^2 + \vartheta x_t^2$$

subject to $\pi_t = \kappa x_t + v_t$, with $v_t = \beta \mathbb{E}_t\{\pi_{t+1}\} + u_t$ taken as given. We attain the optimality condition

$$x_t = -\frac{\kappa}{\vartheta} \pi_t$$

Equilibrium requires that

$$\begin{aligned} \pi_t &= \frac{\vartheta}{\kappa^2 + \vartheta(1 - \beta\rho_u)} u_t \\ x_t &= -\frac{\kappa}{\kappa^2 + \vartheta(1 - \beta\rho_u)} u_t \\ i_t &= r_t^e + \frac{\vartheta\rho_u + \sigma\kappa(1 - \rho_u)}{\kappa^2 + \vartheta(1 - \beta\rho_u)} u_t \end{aligned}$$

Implementation is:

$$\begin{aligned} i_t &= r_t^e + \frac{\vartheta\rho_u + \sigma\kappa(1 - \rho_u)}{\kappa^2 + \vartheta(1 - \beta\rho_u)} u_t + \phi_\pi \left(\pi_t - \frac{\vartheta}{\kappa^2 + \vartheta(1 - \beta\rho_u)} u_t \right) \\ &= r_t^e + \Theta_i u_t + \phi_\pi \pi_t \quad \text{where } \Theta_i = \frac{\sigma\kappa(1 - \rho_u) - \vartheta(\phi_\pi - \rho_u)}{\kappa^2 + \vartheta(1 - \beta\rho_u)} \text{ and } \phi_\pi > 1 \end{aligned}$$

Solving the Phillips curve forward gives us

$$\pi_t = \kappa x_t + \beta \sum_{k=1}^{\infty} \beta^k \mathbb{E}_t \{x_{t+k}\} + \frac{1}{1 - \beta \rho_u} u_t$$

By committing to future negative output gaps, the policymaker can reduce the response to inflation today. Given the convex loss function, smoothing over time is optimal.

Under commitment, the policymaker solves the problem of finding a state-contingent policy $\{x_t, \pi_t\}_{t=0}^{\infty}$ that minimizes

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t (\pi_t^2 + \vartheta x_t^2)$$

subject to the sequence of constraints

$$\pi_t = \beta \mathbb{E}_t \{\pi_{t+1}\} + \kappa x_t + u_t$$

which admits the Lagrangian

$$\mathcal{L} = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{2} (\pi_t^2 + \vartheta x_t^2) + \xi_t (\pi_t - \kappa x_t - \beta \pi_{t+1}) \right] + \underbrace{\text{t.i.p.}}_{\text{terms independent of policy}}$$

The optimality conditions for $t = 0, 1, 2, \dots$ are

$$\vartheta x_t - \kappa \xi_t = 0 \quad \text{and} \quad \pi_t + \xi_t + \xi_{t-1} = 0$$

where we set $\xi_{-1} = 0$. Combining, these become for $t = 1, 2, 3, \dots$

$$x_0 = -\frac{\kappa}{\vartheta} \pi_0 \quad \text{and} \quad x_t = x_{t-1} - \frac{\kappa}{\vartheta} \pi_t$$

Alternatively, $x_t = -\frac{\kappa}{\vartheta} \hat{p}_t$ for $t = 0, 1, 2, \dots$, where $\hat{p}_t = p_t - p_{-1}$. In equilibrium, we have

$$\hat{p}_t = \gamma \hat{p}_{t-1} + \gamma \beta \mathbb{E}_t \{\hat{p}_{t+1}\} + \gamma u_t$$

for $t = 0, 1, 2, \dots$, where $\gamma = \frac{\vartheta}{\vartheta(1+\beta) + \kappa^2}$. The stationary solution is

$$\hat{p}_t = \delta \hat{p}_{t-1} + \frac{\delta}{1 - \delta \beta \rho_u} u_t$$

where $\delta = \frac{1 - \sqrt{1 - 4\beta\gamma^2}}{2\gamma\beta} \in (0, 1)$. This implies price-level targeting!

Remark. What we need to know from this:

1. It is possible to derive optimal policy criteria from utility function of representative household
2. With CES utility and decreasing marginal productivity of labor production functions it is optimal to produce the same amount of each good
3. In the presence of only productivity and demand shocks, optimal policy implies complete price stability
4. In the presence of shocks that imply a trade-off between stabilizing output and inflation, the possibility of committing to future policy actions can lead to better outcomes

2.8 Sticky Wages

Remark. Wages are empirically quite sticky. The probability of a wage change is between 5 and 18% per quarter, is higher for hourly workers than salaried workers, there is little heterogeneity, the probability of a wage change increases with unemployment, and wages are less likely to change downwards than upwards, even accounting for inflation. See Barattieri, Basu, & Gottschalk (2010) for more.

Model. *New Keynesian Model with Sticky Wages* We have a production function

$$Y_t(i) = A_t N_t(i)^{1-\alpha}$$

where $N_t(i) = \left(\int_0^1 N_t(i, j)^{1-\frac{1}{\varepsilon_W}} \partial j \right)^{\frac{\varepsilon_W}{\varepsilon_W-1}}$ and $a_t = \log A_t \sim AR(1)$. The cost minimizing labor demand is

$$N_t(i, j) = \left(\frac{W_t(j)}{W_t} \right)^{\varepsilon_W} N_t(i)$$

where $W_t = \left(\int_0^1 W_t(j)^{1-\varepsilon_W} \partial j \right)^{\frac{1}{1-\varepsilon_W}}$ so that W_t is the ideal wage index, implying that

$$\int_0^1 W_t(j) N_t(i, j) \partial j = W_t N_t(i)$$

The firms problem is

$$\max_{P_t^*} \sum_{k=0}^{\infty} \theta_p^k \mathbb{E}_t \{ \Lambda_{t,t+k} (1/P_{t+k}) (P_t^* Y_{t+k|t} - C_{t+k}(Y_{t+k|t})) \}$$

subject to

$$\begin{aligned} C_{t+k}(Y_{t+k|t}) &= W_{t+k}(Y_{t+k|t}/A_{t+k})^{\frac{1}{1-\alpha}} \\ Y_{t+k|t} &= \left(\frac{P_t^*}{P_{t+k}} \right)^{-\varepsilon_p} C_{t+k} \end{aligned}$$

The implied log-linearized price setting rule is

$$p_t^* = \mu^p + (1 - \beta\theta_p) \sum_{k=0}^{\infty} (\beta\theta_p)^k \mathbb{E}_t \{ \psi_{t+k|t} \}$$

where $\psi_{t+k|t} = \log \Psi_{t+k|t}$ and $\mu^p = \log \frac{\varepsilon_p}{\varepsilon_p-1}$, so nothing has changed except for notation.

We write price inflation in terms of the markup gap:

$$\pi_t^p = \beta \mathbb{E}_t \{ \pi_{t+1}^p \} - \lambda_t (\mu_t^p - \mu^p)$$

where $\lambda_p = \frac{(1-\theta_p)(1-\beta\theta_p)}{\theta_p} \frac{1-\alpha}{1-\alpha+\alpha\varepsilon_p}$, and $\mu_t^p - \mu^p = -(\sigma + \varphi) \tilde{y}_t - (\mu_t^w - \mu^w)$.

Households maximize

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t U(C_t, \{\mathcal{N}_t(j)\}; Z_t)$$

subject to

$$\int_0^1 P_t(i) C_t(i) \partial i + Q_t B_t \leq B_{t-1} + \int_0^1 W_t(j) \mathcal{N}_t(j) \partial j + D_t$$

where

$$U(C_t, \{\mathcal{N}_t(j)\}; Z_t) = \begin{cases} \left(\frac{C_t^{1-\sigma}-1}{1-\sigma} - \int_0^1 \frac{\mathcal{N}_t(j)^{1+\varphi}}{1+\varphi} \partial j \right) Z_t & \text{for } \sigma \neq 1 \\ \left(\log C_t - \int_0^1 \frac{\mathcal{N}_t(j)^{1+\varphi}}{1+\varphi} \partial j \right) Z_t & \text{for } \sigma = 1 \end{cases}$$

and $C_t = \left(\int_0^1 C_t(i)^{1-\frac{1}{\varepsilon_P}} \partial i \right)^{\frac{\varepsilon_P}{\varepsilon_P-1}}$ and $z_t = \log Z_t \sim AR(1)$. The optimality conditions require that

$$C_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\varepsilon_P} C_t$$

$$Q_t = \beta \mathbb{E}_t \left\{ \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} \left(\frac{Z_{t+1}}{Z_t} \right) \left(\frac{P_t}{P_{t+1}} \right) \right\}$$

which, in log-linearized form, is

$$c_t = \mathbb{E}_t \{c_{t+1}\} - \frac{1}{\sigma} (i_t + \mathbb{E}_t \{\pi_{t+1}^p\} - \rho) + \frac{1}{\sigma} (1 - \rho_z) z_t$$

Note however that we have no optimal labor supply condition.

Optimal wage setting solves the problem

$$\max_{W_t^*} \mathbb{E}_t \sum_{k=0}^{\infty} (\beta \theta_w)^k \left(C_{t+k}^{-\sigma} \frac{W_t^*}{P_{t+k}} N_{t+k|t} - \frac{N_{t+k|t}^{1+\varphi}}{1+\varphi} \right) Z_{t+k}$$

subject to

$$N_{t+k|t} = \left(\frac{W_t^*}{W_{t+k}} \right)^{-\varepsilon_w} \left(\int_0^1 N_{t+k}(i) \partial i \right)$$

Which admits first order condition

$$\sum_{k=0}^{\infty} (\beta \theta_w)^k \mathbb{E}_t \left\{ N_{t+k|t} Z_{t+k} C_{t+k}^{-\sigma} \left(\frac{W_t^*}{P_{t+k}} - \mathcal{M}_w MRS_{t+k|t} \right) \right\} = 0$$

where $MRS_{t+k|t} = C_{t+k}^{\sigma} N_{t+k|t}^{\varphi}$ and $\mathcal{M}_w = \frac{\varepsilon_w}{\varepsilon_w - 1}$. The log-linearized version is

$$w_t^* = \mu^w + (1 - \beta \theta_w) \sum_{k=0}^{\infty} (\beta \theta_w)^k \mathbb{E}_t \{ mrs_{t+k|t} + p_{t+k} \}$$

or equivalently,

$$w_t^* = \frac{1 - \beta \theta_w}{1 + \varepsilon_w \varphi} \sum_{k=0}^{\infty} (\beta \theta_w)^k \mathbb{E}_t \{ (1 - \varepsilon_w \varphi) w_{t+k} - \hat{\mu}_{t+k}^w \}$$

where $\mu_t^w = w_t - p_t - mrs_t$ and $mrs_t = \sigma c_t + \varphi n_t$. The wage inflation equation is

$$\pi_t^w = \beta \mathbb{E}_t \{ \pi_{t+1}^w \} - \lambda_w (\mu_t^w - \mu^w)$$

where $\lambda_w = \frac{(1-\beta\theta_w)(1-\theta_w)}{\theta_w(1+\varphi\varepsilon_w)}$, so wage stickiness is introduced in a way that is completely analogous to price stickiness.

Equilibrium requires market clearing in goods, meaning that

$$Y_t(i) = C_t(i) \forall i \in [0, 1] \implies Y_t = C_t, \text{ where } Y_t = \left(\int_0^1 Y_t(i)^{1-\frac{1}{\varepsilon_p}} \partial i \right)^{\frac{\varepsilon_p}{\varepsilon_p-1}}$$

and aggregate employment, so we have that

$$\begin{aligned} N_t &= \int_0^1 \int_0^1 N_t(i, j) \partial j \partial i = \int_0^1 N_t(i) \int_0^1 \frac{N_t(i, j)}{N_t(i)} \partial j \partial i = \Delta_{w,t} \int_0^1 N_t(i) \partial i \\ &= \Delta_{w,t} \left(\frac{Y_t}{A_t} \right)^{\frac{1}{1-\alpha}} \int_0^1 \left(\frac{Y_t(i)}{Y_t} \right)^{\frac{1}{1-\alpha}} \partial i = \Delta_{w,t} \Delta_{p,t} \left(\frac{Y_t}{A_t} \right)^{\frac{1}{1-\alpha}} \end{aligned}$$

where $\Delta_{w,t} = \int_0^1 \left(\frac{W_t(j)}{W_t} \right)^{-\varepsilon_w} \partial j$ and $\Delta_{p,t} = \int_0^1 \left(\frac{P_t(i)}{P_t} \right)^{-\frac{\varepsilon_p}{1-\alpha}} \partial i$. Up to a first order approximation, we have that $(1-\alpha)n_t = y_t - a_t$.

We will define a new and useful variable the *wage gap* $\tilde{\omega}_t$, where

$$\tilde{\omega}_t = \omega_t - \omega_t^n$$

where $\omega = w_t - p_t$ and where ω_t^n is the *natural real wage*:

$$\omega_t^n = \log(1-\alpha) + (a_t - \alpha n_t^n) - \mu^p = \log(1-\alpha) + \psi_{wa} a_t - \mu^p$$

where $\psi_{wa} = \frac{1-\alpha\psi_{ya}}{1-\alpha} > 0$ and $\psi_{ya} = \frac{1+\varphi}{\sigma(1-\alpha)+\varphi+\alpha}$.

We define the *price mark-up gap* $\hat{\mu}_t^p$ as

$$\hat{\mu}_t^p = (mpn_t - \omega_t) - \mu^p = -\frac{\alpha}{1-\alpha} \tilde{y}_t - \tilde{\omega}_t$$

Hence,

$$\pi_t^p = \beta \mathbb{E}_t \{ \pi_{t+1}^p \} + \varkappa_p \tilde{y}_t + \lambda_p \tilde{\omega}_t$$

where $\varkappa_p = \frac{\alpha\lambda_p}{1-\alpha}$.

We define the *wage mark-up gap* $\hat{\mu}_t^w$ as

$$\hat{\mu}_t^w = \omega_t - mrs_t - \mu^w = \tilde{\omega}_t - (\sigma \tilde{y}_t + \varphi \tilde{n}_t) = \tilde{\omega}_t - \left(\sigma + \frac{\varphi}{1-\alpha} \right) \tilde{y}_t$$

so that

$$\pi_t^w = \beta \mathbb{E}_t \{ \pi_{t+1}^w \} + \varkappa_w \tilde{y}_t - \lambda_w \tilde{\omega}_t$$

where $\varkappa_w = \lambda_w \left(\sigma + \frac{\varphi}{1-\alpha} \right)$. In addition, we also have that

$$\tilde{\omega}_t = \tilde{\omega}_{t-1} + \pi_t^w - \pi_t^p - \Delta \omega_t^n$$

Finally, we have the IS equation

$$\tilde{y}_t = -\frac{1}{\sigma} (i_t - \mathbb{E}_t \{ \pi_{t+1}^p \} - r_t^n) + \mathbb{E}_t \{ \tilde{y}_{t+1} \}$$

where $r_t^n = \rho - \sigma(1 - \rho_a)\psi_{ya}a_t + (1 - \rho_z)z_t$, and an interest rate rule

$$i_t = \rho + \phi_p \pi_t^p + \phi_w \pi_t^w + \phi_y \hat{y}_t + v_t$$

We can represent this as a dynamical system

$$\mathbf{A}_0^w \mathbf{x}_t = \mathbf{A}_1^w \mathbb{E}_t\{\mathbf{x}_{t+1}\} + \mathbf{B}_0^w \mathbf{u}_t$$

where

$$\mathbf{A}_0^w = \begin{bmatrix} \sigma + \phi_y & \phi_p & \phi_w & 0 \\ -\kappa_p & 1 & 0 & 0 \\ -\kappa_w & 0 & 1 & 0 \\ 0 & -1 & 1 & 1 \end{bmatrix} \quad \mathbf{x}_t = \begin{bmatrix} \tilde{y}_t \\ \pi_t^p \\ \pi_t^w \\ \tilde{\omega}_{t-1} \end{bmatrix} \quad \mathbf{A}_1^w = \begin{bmatrix} \sigma & 1 & 0 & 0 \\ 0 & \beta & 0 & \lambda_p \\ 0 & 0 & \beta & -\lambda_w \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{B}_0^w = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \quad \mathbf{u}_t = \begin{bmatrix} \hat{r}_t^n - v_t - \phi_y \hat{y}_t^n \\ \Delta \omega_t^n \end{bmatrix}$$

For the equilibrium to be unique, it is necessary and sufficient that

$$\phi_p + \phi_w + \phi_y \left(\frac{1 - \beta}{\sigma + \frac{\alpha + \varphi}{1 - \alpha}} \right) \left(\frac{1}{\lambda_p} + \frac{1}{\lambda_w} \right) > 1$$

Example. *Dynamic Responses to a Monetary Policy Shock* We can see in simulation that while it may seem like price and wage stickiness work in opposite directions, they actually don't cancel out the other's inefficiencies. They generally both slow the response of the economy to a shock.

The social planner's problem is now

$$\max U(C_y\{\mathcal{N}_t(j)\}; Z_t)$$

subject to

$$C_t(i) = A_t N_t(i)^{1-\alpha}$$

$$\mathcal{N}_t(j) = \int_0^1 N_t(i, j) \partial i$$

where C_t and $N_t(i)$ are defined as above. Optimality conditions are that

$$C_t(i) = C_t \quad \forall i \in [0, 1]$$

$$N_t(i, j) = \mathcal{N}_t(j) = N_t(i) = N_t \quad \forall i, j \in [0, 1]$$

$$-\frac{U_{n,t}}{U_{c,t}} = MPN_t = (1 - \alpha)A_t N_t^{-\alpha}$$

In the decentralized economy with flexible prices and wages,

$$P_t = \mathcal{M}_p \frac{(1 - \tau)W_t}{MPN_t}$$

$$\frac{W_t}{P_t} = -\frac{U_{n,t}}{U_{c,t}} \mathcal{M}_w$$

for all goods and occupations, with $\mathcal{M}_p = \frac{\varepsilon_p}{\varepsilon_p - 1}$ and $\mathcal{M}_w = \frac{\varepsilon_w}{\varepsilon_w - 1}$. Thus,

$$-\frac{U_{n,t}}{U_{c,t}} = \frac{1}{\mathcal{M}(1-\tau)} MPN_t$$

where $\mathcal{M} = \mathcal{M}_p \mathcal{M}_w$. The condition for efficiency of the natural equilibrium is that $\mathcal{M}(1-\tau) = 1$. Note that this is not generically available with sticky prices and / or wages.

The optimal monetary policy problem is

$$\min \frac{1}{2} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\left(\sigma + \frac{\varphi + \alpha}{1 - \alpha} \right) \tilde{y}_t^2 + \frac{\varepsilon_p}{\lambda_p} \pi_t^p)^2 + \frac{\varepsilon_w(1 - \alpha)}{\lambda_w} (\pi_t^w)^2 \right)$$

subject to

$$\begin{aligned} \pi_t^p &= \beta \mathbb{E}_t \{ \pi_{t+1}^p \} + \kappa_p \tilde{y}_t + \lambda_p \tilde{\omega}_t \\ \pi_t^w &= \beta \mathbb{E}_t \{ \pi_{t+1}^w \} + \kappa_w \tilde{y}_t + \lambda_w \tilde{\omega}_t \\ \tilde{\omega}_{t-1} &= \tilde{\omega}_t - \pi_t^w + \pi_t^p + \Delta \omega_t^n \end{aligned}$$

We consider two types of rules: strict targeting rules, under which $\pi_t^i = 0$, and flexible targeting rules under which $i_t = 0.01 + 1.5\pi_t^i$. With a flexible rule, interest rates can respond to either price or wage inflation (or a composite measure). Looking to the data, we can see that composite inflation targeting can reattain something fairly close to the first-best level of output even with price and wage stickiness. It's doing a fairly good job.

2.9 Unemployment

Remark. One criticism of the baseline New Keynesian model is that it has no role for unemployment. Today, we will reinterpret the NK model to analyze unemployment and introduce an alternative to the search friction-based framework we studied under Ryan.

Model. *Unemployment in the New Keynesian Model* We have a representative household with a continuum of members, indexed by $(j, s) \in [0, 1] \times [0, 1]$, where we have a continuum of differentiated occupations indexed by $j \in [0, 1]$ and households get disutility from (indivisible) labor χs^φ for $s \in [0, 1]$, where $\varphi \geq 0$. We assume full consumption risk sharing within the household. The household utility function becomes:

$$U(C_t, \{\mathcal{N}_t(j)\}; Z_t) = \left(\frac{C_t^{1-\sigma} - 1}{1-\sigma} - \chi \int_0^1 \int_0^{\mathcal{N}_t(j)} s^\varphi \partial s \partial j \right) Z_t = \left(\frac{C_t^{1-\sigma} - 1}{1-\sigma} - \chi \int_0^1 \frac{\mathcal{N}_t(j)^{1+\varphi}}{1+\varphi} \partial j \right) Z_t$$

where $X_t = \left(\int_0^1 C_t(i)^{1-\frac{1}{\varepsilon_p}} \partial i \right)^{\frac{\varepsilon_p}{\varepsilon_p-1}}$ with budget constraint

$$\int_0^1 P_t(i) C_t(i) \partial i + Q_t B_t \leq B_{t-1} + \int_0^1 W_t(j) \mathcal{N}_t(j) \partial j + D_t$$

and two optimality conditions:

$$C_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\varepsilon_p} C_t$$

where $P_t = \left(\int_0^1 P_t(i)^{1-\varepsilon_p} \partial i \right)^{\frac{1}{1-\varepsilon_p}}$ which implies that $\int_0^1 P_t(i) C_t(i) \partial i = P_t C_t$; and

$$Q_t = \beta \mathbb{E}_t \left\{ \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} \left(\frac{Z_{t+1}}{Z_t} \right) \left(\frac{P_t}{P_{t+1}} \right) \right\}$$

The average wage dynamics are

$$w_t = \theta_w w_{t-1} + (1 - \theta_w) w_t^*$$

and the optimal wage setting rule is

$$w_t^* = \mu^w + (1 - \beta \theta_w) \sum_{k=0}^{\infty} (\beta \theta_w)^k \mathbb{E}_t \{ mrs_{t+k|t} + p_{t+k} \}$$

where $\mu^w = \log \frac{\varepsilon_w}{\varepsilon_w - 1}$ and $mrs_{t+k|t} = \sigma c_{t+k} + \varphi n_{t+k|t} + \xi$, with wage inflation equation

$$\pi_t^w = \beta \mathbb{E}_t \{ \pi_{t+1}^w \} - \lambda_w (\mu_t^w - \mu^w)$$

where $\pi_t^w = w_t - w_{t-1}$, $\mu_t^w = w_t - p_t - mrs_t$, and $\lambda_w = \frac{(1-\theta_w)(1-\beta\theta_w)}{\theta_w(1+e_w\varphi)}$.

The participation condition for an individual (j, s) is

$$\frac{W_t(j)}{P_t} \geq \chi C_t^\sigma s^\varphi$$

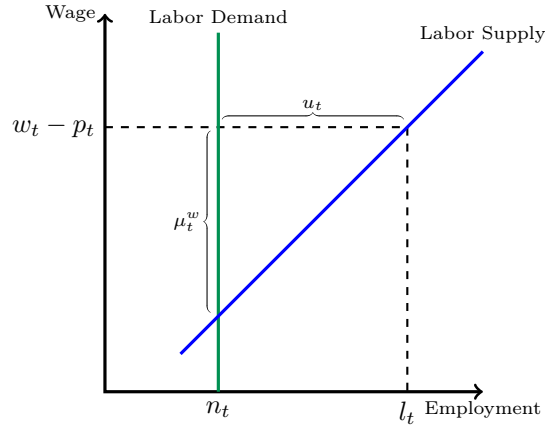
The marginal participant, $L_t(j)$, is given by

$$\frac{W_t(j)}{P_t} \geq \chi C_t^\sigma L_t(j)^\varphi$$

The aggregate labor force is therefore

$$w_t - p_t = \sigma c_t + \varphi l_t + \xi$$

where $w_t \cong \int_0^1 w_t(j) \partial j$ and $l_t = \int_0^1 l_t(j) \partial j$. We can represent this visually as:



The unemployment rate is now $u_t = l_t - n_t$, so we have average wage markup is

$$\mu_t^w = (w_t - p_t) - (\sigma c_t + \varphi n_t + \xi) = \varphi u_t$$

and under flexible wages, $\mu^w = \varphi u^n$, the natural rate of unemployment. We thus have a *New Keynesian Wage Phillips Curve*

$$\pi_t^w = \beta \mathbb{E}_t \{\pi_{t+1}^w\} - \lambda_w \varphi (u_t - u^n)$$

Our production technology is

$$Y_t(i) = A_t N_t(i)^{1-\alpha}$$

where $N_t(i) = \left(\int_0^1 N_t(i, j)^{1-\frac{1}{\varepsilon_w}} \partial j \right)^{\frac{\varepsilon_w}{\varepsilon_w-1}}$, which gives us price dynamics

$$p_t = \theta_p p_{t-1} + (1 - \theta_p) p_t^*$$

and optimal price setting rule

$$p_t^* = \mu^p + (1 - \beta \theta_p) \sum_{k=0}^{\infty} (\beta \theta_p)^k \mathbb{E}_t \{\psi_{t+k|t}\}$$

The implied price inflation equation is

$$\pi_t^p = \beta \mathbb{E}_t \{\pi_{t+1}^p\} - \lambda_p (\mu_t^p - \mu^p)$$

where $\mu_t^p = p_t - \psi_t$ and $\psi_t = w - (a_t - \alpha n_t + \log(1 - \alpha))$, and $\lambda_p = \frac{(1-\theta_p)(1-\beta\theta_p)}{\theta_p} \frac{1-\alpha}{1-\alpha+\alpha\varepsilon_p}$.

The equilibrium non-policy block is:

$$\begin{aligned} \tilde{y}_t &= -\frac{1}{\sigma} (i_t - \mathbb{E}_t \{\pi_{t+1}^p\} - r_t^n) + \mathbb{E}_t \{\tilde{y}_{t+1}\} \\ \pi_t^p &= \beta \mathbb{E}_t \{\pi_{t+1}^p\} + \varkappa_p \tilde{y}_t + \lambda_p \tilde{\omega}_t \\ \pi_t^w &= \beta \mathbb{E}_t \{\pi_{t+1}^w\} - \lambda_w \varphi \hat{u}_t \\ \tilde{\omega}_t &= \tilde{\omega}_{t-1} + \pi_t^w - \pi_t^p - \Delta \omega_t^n \\ \varphi \hat{u}_t &= \hat{u}_t^w = \tilde{\omega}_t - (\sigma \tilde{c}_t + \varphi \tilde{n}_t) = \tilde{\omega} - \left(\sigma + \frac{\varphi}{1-\alpha} \right) \tilde{y}_t \end{aligned}$$

An example of the equilibrium policy block with a Taylor-type rule $i_t = \rho + \phi_p \pi_t^p + \phi_y \hat{y}_t + v_t$ is the natural equilibrium

$$\begin{aligned} \hat{y}_t^n &= \psi_{ya} a_t \\ r_t^n &= \rho - \sigma(1 - \rho_a) \psi_{ya} a_t + (1 - \rho_z) z_t \\ \hat{\omega}_t^n &= \psi_{wa} a_t \end{aligned}$$

with $\psi_{ya} = \frac{1+\varphi}{\sigma(1-\alpha)+\varphi+\alpha}$ and $\psi_{wa} = \frac{1-\alpha\psi_{ya}}{1-\alpha} > 0$, and exogenous $AR(1)$ processes for $\{a_t\}$, $\{z_t\}$ and $\{v_t\}$.

We can compare this Taylor-type rule to the solution to the optimal monetary policy problem

$$\min \frac{1}{2} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\left(\sigma + \frac{\varphi + \alpha}{1 - \alpha} \right) \tilde{y}_t^2 + \frac{\varepsilon_p}{\lambda_p} (\pi_t^p)^2 + \frac{\varepsilon_w(1 - \alpha)}{\lambda_w} (\pi_t^w)^2 \right)$$

subject to

$$\begin{aligned} \pi_t^p &= \beta \mathbb{E}_t \{\pi_{t+1}^p\} + \varkappa_p \tilde{y}_t + \lambda_p \tilde{\omega}_t \\ \pi_t^w &= \beta \mathbb{E}_t \{\pi_{t+1}^w\} + \varkappa_w \tilde{y}_t + \lambda_w \tilde{\omega}_t \\ \tilde{\omega}_t &= \tilde{\omega}_{t-1} + \pi_t^w - \pi_t^p - \Delta \omega_t^n \end{aligned}$$

Optimality conditions are:

$$\begin{aligned}
0 &= \left(\sigma + \frac{\varphi + \alpha}{1 - \alpha} \right) \tilde{y}_t + \varkappa_p \zeta_{1,t} + \varkappa_w \zeta_{2,t} \\
0 &= \frac{\varepsilon_p}{\lambda_p} \pi_t^p - \Delta \zeta_{1,t} + \zeta_{3,t} \\
0 &= \frac{\varepsilon_w(1 - \alpha)}{\lambda_w} \pi_t^2 - \Delta \zeta_{2,t} - \zeta_{3,t} \\
0 &= \lambda_p \zeta_{1,t} - \lambda_w \zeta_{2,t} + \zeta_{3,t} - \beta \mathbb{E}\{\zeta_{3,t+1}\}
\end{aligned}$$

Compare this also to a simple rule with unemployment:

$$i_t = 0.01 + 1.5\pi_t^p - 0.5\hat{u}_t$$

Bringing this to the data, we can see that the simple rule almost perfectly recreates the optimal monetary policy rule, while the Taylor-type rule has a lot of distortions from the optimal.

Remark. Wage stickiness and differentiated labor types allow for reinterpreting the New Keynesian model to include unemployment; and the labor force includes all agents who would work for the aggregate real wage.

Criticisms of the basic New Keynesian Model.

1. We can introduce involuntary unemployment into the NK model, but not search unemployment or any adverse consequences of unemployment. [Trigari \(2009\)](#) introduced search unemployment into the NK model.
2. The basic model uses a representative household, but in reality not all households earn the same wage and there is large heterogeneity in wealth holdings. The costs are that the model makes inaccurate predictions if well-paid and wealthy households behave differently from low-paid and poor households, and that the model gives inaccurate normative recommendations if we care about inequality. [McKay & Reis \(2016\)](#) and [Kaplan, Moll, & Violante \(2016\)](#) explore some *Heterogenous Agent New Keynesian (HANK) models*.
3. Consumption (and output) decision is completely determined by the expected real interest rate. There is little empirical evidence that individual spending and investment decisions respond to expected real interest rates. The cost is that the main model mechanism is not a good description of what drives consumption and investment in reality. Heterogeneous agent models allow for wealth effects and precautionary motives to affect the spending / saving decision.

2.10 State Space Models (and the Kalman Filter)

The most general way to write linear models is as state space systems:

$$\begin{aligned}
X_t &= A_t X_{t-1} + C_t \mathbf{u}_t : \mathbf{u}_t \sim \mathcal{N}(0, I) && \text{(State Equation)} \\
Z_t &= D_t X_t + \mathbf{v}_t : \mathbf{v}_t \sim \mathcal{N}(0, \Sigma_v) && \text{(Measurement Equation)}
\end{aligned}$$

This nests “observable” VAR(p), MA(p), and VARMA(p, q) processes as well as systems with latent variables. As an example, the VAR(p) model

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + u_t$$

can be written as

$$\begin{aligned} X_t &= A_t X_{t-1} + C_t \mathbf{u}_t \\ Z_t &= D_t X_t + \mathbf{v}_t \end{aligned}$$

where

$$A = \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_p \\ I & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & I & 0 \end{bmatrix} ; \quad C = \begin{bmatrix} I \\ 0 \\ 0 \\ 0 \end{bmatrix} \mathbf{u}_t ; \quad D = \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix}^T ; \quad \Sigma_{vv} = 0$$

The MA(1) process $y_t = \varepsilon_t + \theta \varepsilon_{t-1}$ can be written as

$$\begin{aligned} \begin{bmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{bmatrix} &= \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_{t-1} \\ \varepsilon_{t-2} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \varepsilon_t \\ y_t &= \begin{bmatrix} 1 & \theta \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{bmatrix} \end{aligned}$$

Definition. The *Kalman Filter* is used for mainly two purposes: (i) to estimate the unobservable state X_t , and (ii) to evaluate the likelihood function associated with a state space model.

For state spaces of the form

$$\begin{aligned} X_t &= A_t X_{t-1} + C_t \mathbf{u}_t \\ Z_t &= D_t X_t + \mathbf{v}_t \end{aligned}$$

the Kalman filter recursively computes estimates of X_t conditional on the history of observations Z_t, Z_{t-1}, \dots, Z_0 and an initial estimate (or prior) $X_{0|0}$ with variance $P_{0|0}$. The form of the filter is

$$X_{t|t} = A_t X_{t-1|t-1} + K_t (Z_t - D_t X_{t|t-1})$$

and the task is thus to find the *Kalman gain* K_t so that the estimates $X_{t|t}$ are in some sense ‘optimal’. We define

$$X_{t|t-s} = \mathbb{E} [X_t | Z^{t-s}] \quad \text{and} \quad P_{t|t-s} = \mathbb{E} [(X_t - X_{t|t-s})(X_t - X_{t|t-s})']$$

Example. A Simple Example Let’s say that we have noisy measures z_1 of an unobservable process x such that $z_1 = x + v_1$, where $v_1 \sim \mathcal{N}(0, \sigma_1^2)$. Since the signal is unbiased, the minimum variance estimate $\mathbb{E}[x | z_1] = \hat{x}$ of x is given by $\hat{x} = z_1$, and its variance is equal to the variance of the noise $\mathbb{E}[\hat{x} - x]^2 = \sigma_1^2$.

Now say we have a second measure z_2 of x so that $z_2 = x + v_2$, where $v_2 \sim \mathcal{N}(0, \sigma_2^2)$. How can we combine the information in the two signals to find the minimum variance estimate of x ? Clearly, we will restrict to linear estimators of the form

$$\hat{x} = (1 - a)z_1 + az_2$$

and we will simply minimize $\mathbb{E} [(1 - a)z_1 + az_2 - x]^2$ with respect to a . We rewrite the variance as

$$\mathbb{E} [(1 - a)(x + v_1) + a(x + v_2) - x]^2 = \sigma_1^2 - 2a\sigma_1^2 + a^2\sigma_1^2 + a^2\sigma_2^2$$

Our first order condition attains $a = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$. The minimum variance estimate of x is thus

$$\hat{x} = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} z_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} z_2$$

which has conditional variance

$$\mathbb{E}[\hat{x} - x]^2 = \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)^{-1}$$

As long as $\sigma_2^2 < \infty$, this strictly improves on the variance of the estimate with only one signal.

Example. *Scalar Kalman Filter* Consider the process

$$\begin{aligned} x_t &= \rho x_{t-1} + u_t \\ z_t &= x_t + v_t \\ \begin{bmatrix} u_t \\ v_t \end{bmatrix} &\sim \mathcal{N} \left(0, \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix} \right) \end{aligned}$$

We want to form an estimate of x_t conditional on $z^t = \{z_t, z_{t-1}, \dots, z_1\}$. In addition to the system above, we have a prior on the initial value of the state x_0 such that $x_{0|0} = \bar{x}_0$ and $\mathbb{E}(\bar{x}_0 - x_0)^2 = p_0$. With this information we can form a prior on x_1 . Using the state transition equation we get that

$$x_{1|0} = \mathbb{E}[x_1 | x_{0|0}] = \rho x_{0|0} \quad \text{with variance} \quad \mathbb{E}(x_{1|0} - x_1)^2 = \rho^2 p_0 + \sigma_u^2$$

Note that $\rho^2 p_0$ is the uncertainty from period 0 carried to period 1, and σ_u^2 is the uncertainty in period 0 about the period 1 innovation to x_t . We denote the *prior variance* as $p_{1|0} = \rho^2 p_0 + \sigma_u^2$. We can combine the information in the signal z_1 with the information in the prior in exactly the same way as we combined the two signals in the simple example above. The optimal weight k_1 in the equation $x_{1|1} = (1 - k_1)x_{1|0} + k_1 z_1$ is thus given by

$$k_1 = \frac{p_{1|0}}{p_{1|0} + \sigma_v^2}$$

and the period 1 posterior error covariance $p_{1|1}$ is

$$p_{1|1} = \left(\frac{1}{p_{1|0} + \frac{1}{\sigma_v^2}} \right)^{-1}$$

or equivalently, $p_{1|1} = p_{1|0} - p_{1|0}^2 (p_{1|0} + \sigma_v^2)^{-1}$. We can propagate this forward to get

$$p_{2|1} = \rho^2 \left(p_{1|0} - p_{1|0}^2 (p_{1|0} + \sigma_v^2)^{-1} \right) + \sigma_u^2$$

and in general we have the difference equation

$$p_{t|t-1} = \rho^2 \left(p_{t-1|t-2} - p_{t-1|t-2}^2 (p_{t-1|t-2} + \sigma_v^2)^{-1} \right) + \sigma_u^2$$

We thus have the Kalman gain $k_t = p_{t|t-1} (p_{t|t-1} + \sigma_v^2)^{-1}$ which allows us to compute

$$x_{t|t} = (1 - k_t)x_{t|t-1} + k_t z_t$$

The state and measurement equations:

$$x_t = \rho x_{t-1} + u_t \quad \text{and} \quad z_t = x_t + v_t$$

give us the Kalman update equations:

$$\begin{aligned}x_{t|t} &= \rho x_{t-1|t-1} + k_t(z_t - \rho x_{t-1|t-1}) \\k_t &= p_{t|t-1}(p_{t|t-1} + \sigma_v^2)^{-1} \\p_{t|t-1} &= \rho^2 \underbrace{\left(p_{t-1|t-2} - p_{t-1|t-2}^2(p_{t-1|t-2} + \sigma_v^2)^{-1}\right)}_{p_{t-1|t-1}} + \sigma_u^2\end{aligned}$$

Remark. There are two things worth noting about the difference equation for the prior error variances:

1. The prior error variance is bounded from below and above so that

$$\sigma_u^2 \leq p_{t|t-1} \leq \frac{\sigma_u^2}{1 - \rho^2}$$

2. For $|\rho| \in [0, 1)$, this is a contradiction

The upper bound is given by the optimality of the filter, where we cannot do worse than making the unconditional mean x_t for all t . The lower bound is given by the fact that the future is inherently uncertain, so even with a perfect estimate of x_{t-1} , x_t will still not be known with certainty.

Question. What determines the Kalman gain k_t ?

The Kalman filter will optimally combine information in the prior $\rho x_{t-1|t-1}$ and the signal z_t to form a posterior estimate $x_{t|t}$ with covariance $p_{t|t}$:

$$x_{t|t} = (1 - k_t)\rho x_{t-1|t-1} + k_t z_t$$

We put more weight on the signal, meaning a higher Kalman gain, if prior variance is large or the signal is very precise. The prior variance can be large either because the previous state estimate was imprecise (*i.e.* $p_{t-1|t-1}$ is large) or because the variance of state innovations is large (*i.e.* σ_u^2 is large).

Remark. If $\rho < 1$ and if ρ , σ_u^2 , and σ_v^2 are constant, the prior variance of the state estimate will converge to

$$p = \rho^2 (p - p^2(p + \sigma_v^2)^{-1}) + \sigma_u^2$$

and the Kalman gain will converge to

$$k = p(p + \sigma_v^2)^{-1}$$

Example. *The Multivariate Kalman Filter* For state space systems of the form

$$\begin{aligned}X_t &= A_t X_{t-1} + C_t \mathbf{u}_t \\Z_t &= D_t X_t + \mathbf{v}_t\end{aligned}$$

the Kalman filter recursively computes estimates of X_t conditional on the history of observations Z_t, Z_{t-1}, \dots, Z_0 and an initial estimate $X_{0|0}$ with variance $P_{0|0}$. The form of the filter is

$$X_{t|t} = A_t X_{t-1|t-1} + K_t (Z_t - D_t X_{t|t-1})$$

and the task is again to find the Kalman gain K_t so that the estimates $X_{t|t}$ are optimal. We further assume that $X_{0|0} - X_0$ is uncorrelated with the shock processes $\{\mathbf{u}_t\}$ and $\{\mathbf{v}_t\}$. We can find a brute force linear minimum variance estimator by solving the period t problem

$$\min_{\alpha} \mathbb{E} \left[X_t - \sum_{j=0}^t \alpha_j Z_{t-j} \right] \left[X_t - \sum_{j=0}^t \alpha_j Z_{t-j} \right]'$$

We want to find the linear projection of X_t on the history of observables Z_t, Z_{t-1}, \dots, Z_1 . From the projection theorem, the linear combination $\sum_{j=1}^t \alpha_j Z_{t-j+1}$ should imply errors that are orthogonal to Z_t, Z_{t-1}, \dots, Z_1 so that the following holds:

$$\left(X_t - \sum_{j=0}^t \alpha_j Z_{t-j} \right) \perp \{Z_j\}_{j=1}^t$$

We could compute the α 's directly as

$$\mathbb{P}\{X_t \mid Z_t, Z_{t-1}, \dots, Z_1\} = \mathbb{E} \left(X_t [Z'_t Z'_{t-1} \dots Z'_1]' \right) \times \left(\mathbb{E} [Z'_t Z'_{t-1} \dots Z'_1] [Z'_t Z'_{t-1} \dots Z'_1]' \right)^{-1} \times [Z'_t Z'_{t-1} \dots Z'_1]'$$

However, as $t \rightarrow \infty$ that gets really messy. We will use two tricks to find a nice recursive formulation:

1. Gram-Schmidt Orthogonalization

2. Exploit properties of projections onto mutually orthogonal variables

Definition. *Gram-Schmidt Orthogonalization* is defined in \mathbb{R}^m when we have a matrix $Y \in \mathbb{R}^{m \times n}$ with columns y_1, y_2, \dots, y_n . We will keep the first column, and then subtract the projection of y_2 on y_1 from y_2 and define a new column vector \tilde{y}_2 as:

$$\tilde{y}_2 = y_2 - y_1(y'_1 y_1)^{-1} y'_1 y_2 \iff \tilde{y}_2 = (I - \mathcal{P}_{y_1}) y_2$$

and then subtract the projection of y_3 on $[y_1 \ y_2]$ from y_3 to construct \tilde{y}_3 , and so on.

Lemma 2.1. *Let Z and Y be two uncorrelated mean zero variables where $\mathbb{E}[ZY'] = 0$. Then*

$$\mathbb{E}[X \mid Z, Y] = \mathbb{E}[X \mid Z] + \mathbb{E}[X \mid Y]$$

Proof. Follows directly from the projection formula, if the variables we are projecting onto are orthogonal, we will be taking the inverse of a diagonal matrix. \square

With these, we can find the Kalman gain, which as a reminder is defined by the equation

$$X_{t|t} = A_t X_{t-1|t-1} + K_t (Z_t - D_t X_{t|t-1})$$

We start from the first-period problem of how to optimally combine the information in the prior $X_{0|0}$ and the signal Z_1 . We will use that $Z_1 = D_1 A_1 X_0 + D_1 C \mathbf{u}_1 + \mathbf{v}_1$ and that we know that \mathbf{u}_t and \mathbf{v}_t are orthogonal to $X_{0|0}$ to first find the optimal projection of Z_1 on $X_{0|0}$: $Z_{1|0} = D_1 A_1 X_{0|0}$. We can then define the period 1 innovation \tilde{Z}_1 in Z_1 as $\tilde{Z}_1 = Z_1 - Z_{1|0}$. We know that

$$\mathbb{E}(X_1 \mid \tilde{Z}_1, X_{0|0}) = \mathbb{E}(X_1 \mid \tilde{Z}_1) + \mathbb{E}(X_1 \mid X_{0|0})$$

since $\tilde{Z}_1 \perp X_{0|0}$ and $\mathbb{E}(Z_1 \mid X_{0|0}) = D_1 A_1 X_{0|0}$. From the projection theorem, we will look for a K_1 such that the inner product of the projection error and \tilde{Z}_1 is zero:

$$\langle X_1 - K_1 \tilde{Z}_1, \tilde{Z}_1 \rangle = 0$$

Defining the inner product $\langle X, Y \rangle$ as $\mathbb{E}(XY')$, we get

$$\begin{aligned}\mathbb{E} \left[\left(X_1 - K_t \tilde{Z}_1 \right) \tilde{Z}_1' \right] &= 0 \\ \mathbb{E} \left[Z_1 \tilde{Z}_1' \right] - K_1 \mathbb{E} \left[\tilde{Z}_1 \tilde{Z}_1' \right] &= 0 \\ K_1 &= \mathbb{E} \left[X_1 \tilde{Z}_1' \right] \left(\mathbb{E} \left[\tilde{Z}_1 \tilde{Z}_1' \right] \right)^{-1}\end{aligned}$$

We thus need to evaluate the two expectational expressions. Before doing so it helps to define the state innovation $\tilde{X}_1 = X_1 - X_{1|0}$, so \tilde{X}_1 is the one period error. The first expectation factor of K_1 can now be manipulated as follows:

$$\mathbb{E} \left[X_1 \tilde{Z}_1' \right] = \mathbb{E} \left(\tilde{X}_1 + X_{1|0} \right) \tilde{Z}_1' = \mathbb{E} \tilde{X}_1 \tilde{Z}_1' = \mathbb{E} \tilde{X}_1 \left(\tilde{X}_1 D' + \mathbf{v}_1' \right) = P_{1|0} D'$$

Evaluating the second expectation factor as follows

$$\mathbb{E} \left[\tilde{Z}_1 \tilde{Z}_1' \right] = \mathbb{E} \left[\left(D_1 \tilde{X}_1 + \mathbf{v}_1 \right) \left(D_1 \tilde{X}_1 + \mathbf{v}_1 \right)' \right] = D_1 P_{1|0} D_1' + \Sigma_{vv}$$

gives us the last component needed to find K_1 :

$$K_1 = P_{1|0} D_1' \left(D_1 P_{1|0} D_1' + \Sigma_{vv} \right)^{-1}$$

where we know that $P_{1|0} = A_1 P_{0|0} A_1' + C_0 C_0'$. We can add the projections of X_1 on \tilde{Z}_1 and $X_{0|0}$ to recover the linear minimum variance estimate $X_{1|1}$

$$X_{1|1} = \mathbb{E} \left(X_1 \mid X_{0|0} \right) + \mathbb{E} \left(X_1 \mid \tilde{Z}_1 \right) = A_1 X_{0|0} + K_1 \tilde{Z}_1$$

It remains to find an expression for the period t covariance. We can rewrite $X_{t|t} = K_t \tilde{Z}_t + X_{t|t-1}$ as

$$X_1 - X_{t|t} + K_t \tilde{Z}_t = X_t - X_{t|t-1}$$

Since the period t error $X_t - X_{t|t}$ is orthogonal to \tilde{Z}_t , the variance of the right hand side must be equal to the sum of the variances of the terms on the left hand side. We thus have

$$P_{t|t} + K_t \left(D P_{t|t-1} D' + \Sigma_{vv} \right) K_t' = P_{t|t-1}$$

Rearranging,

$$P_{t|t} = P_{t|t-1} - K_t \left(D P_{t|t-1} D' + \Sigma_{vv} \right) K_t' = P_{t|t-1} - P_{t|t-1} D_t' \left(D_t P_{t|t-1} D_t' + \Sigma_{vv} \right)^{-1} D_t P_{t|t-1}$$

It is then straightforward to show that

$$P_{t+1|t} = A_{t+1} P_{t|t} A_{t+1}' + C C' = A_{t+1} \left(P_{t|t-1} - P_{t|t-1} D_t' \left(D_t P_{t|t-1} D_t' + \Sigma_{vv} \right)^{-1} D_t P_{t|t-1} \right) A_{t+1}' + C C'$$

Remark. In summation, for the state space system

$$\begin{aligned}X_t &= A_t X_{t-1} + C_t \mathbf{u}_t \\ Z_t &= D_t X_t + \mathbf{v}_t \\ \begin{bmatrix} \mathbf{u}_t \\ \mathbf{v}_t \end{bmatrix} &\sim \mathcal{N} \left(0, \begin{bmatrix} I_n & \mathbf{0}_{n \times I} \\ \mathbf{0}_{I \times n} & \Sigma_{vv} \end{bmatrix} \right)\end{aligned}$$

we have the state estimate update equations

$$\begin{aligned} X_{t|t} &= A_t X_{t-1|t-1} + K_t (Z_t - D_t X_{t|t-1}) \\ K_t &= P_{t|t-1} D_t' (D_t P_{t|t-1} D_t' + \Sigma_{vv})^{-1} \\ P_{t+1|t} &= A_{t+1} \left(P_{t|t-1} - P_{t|t-1} D_t' (D_t P_{t|t-1} D_t' + \Sigma_{vv})^{-1} D_t P_{t|t-1} \right) A_{t+1}' + CC' \end{aligned}$$

The innovation sequence can be computed recursively from the innovation representation

$$\tilde{Z}_t = Z_t - D_t X_{t|t-1} \quad ; \quad X_{t+1|t} = A_{t+1} X_{t|t-1} + A_{t+1} K_t \tilde{Z}_t$$

The Kalman filter can be used to estimate latent variables in a state space system, and to evaluate the likelihood function for a given parameterized state space system. We will explore the latter later.

2.11 Calibration and Moment Matching

Remark. Structural macroeconomic models can be used to make predictions about the effects of monetary policy changes, changes in taxes, etc. However, model predictions often depend on model parameters. How should we choose parameters for the model? There are three main categories of methods:

1. Calibration
2. Moment matching
3. Likelihood-based methods

Today, we do the first two.

Definition. *Calibration* is using evidence from existing studies to choose parameters, solving the model, and comparing the output to actual macroeconomic data (that was *not* used when choosing the parameters). If the output resembles the actual data, we can claim success.

Kydland & Prescott (1996) defined a *computational experiment* as the following process:

1. Ask a well-posed question
2. Write down a suitable model that can address the question
3. Choose parameters from empirical evidence that is not directly related to the question at hand
4. Solve the model
5. Compute the answer

Example. The original experiment is Kydland & Prescott (1982), where they ask whether an equilibrium model can generate realistic business cycles. Their model is a RBC model where it takes more than one period to install new capital, households choose how much labor to supply and how much to consume, and investment is income minus consumption.

They used data from the average labor share of GNP to calibrate the production function, the investment share of GDP and steady-state capital-to-output ratio to set the depreciation parameter, and the average risk-free real interest rate to set the discount rate. Importantly, parameters were not chosen based on the features of the data that the model is used to address. In addition, some parameters were chosen without reference to any empirical observations.

Remark. We could do similarly in the New Keynesian model.

Remark. There are several valid criticisms of calibration:

- The choice of empirical facts used for calibration is arbitrary

- The choice of which moments to use to judge the performance of the model is arbitrary
- The mapping between microeconomic studies to parameters is not always straightforward or innocuous
- There is no formal way to take into account parameter uncertainty
- What is judged a success is not based on any formal criteria

Definition. *Moment matching* is choosing parameters to minimize the discrepancy between the model and the data. It allows for a more formal statistical treatment of choosing parameters and testing over-identified restrictions. Note the difference from calibration: the data moments are used to choose parameter values rather than evaluate the fit of the model.

Example. Consider a data moment $h(z_t)$ that has a clear correspondence to a model moment $h(y_t, \Theta)$. Given simulated model data $\{y_i\}_{i=1}^N$, we can compute the discrepancy between the data and the model moment as

$$g(Z, \Theta) = \left(\frac{1}{T} \sum_{t=1}^T h(z_t) \right) - \left(\frac{1}{N} \sum_{i=1}^N h(y_i, \Theta) \right)$$

We should then choose parameters Θ to minimize the weighted discrepancy between model and data moments

$$\Gamma(\Theta) = g(Z, \Theta)' \times \left[\Sigma \left(1 + \frac{T}{N} \right) \right]^{-1} \times g(Z, \Theta)$$

where Σ in the weighting matrix is the covariance of numerical standard deviation of $h(z_t) - h(y_t, \Theta)$ under the null hypothesis. Since the weighting matrix Σ depends on the parameters Θ , multiple iterations are necessary in order to find the Θ that minimizes $\Gamma(\Theta)$.

2.12 Likelihood-Based Estimation

Example. *Estimating the parameters in a state space system* For a given state space system

$$\begin{aligned} X_t &= AX_{t-1} + C\mathbf{u}_t & \mathbf{u}_t &\sim \mathcal{N}(0, I) \\ Z_t &= DX_t + \mathbf{v}_t & \mathbf{v}_t &\sim \mathcal{N}(0, \Sigma_{vv}) \end{aligned}$$

How can we find the A , C , D , and Σ_{vv} that best fit the data?

1. **The Kalman Filter:** We saw earlier that the associated Kalman filter for this system is given by

$$\begin{aligned} X_{t|t} &= A_t X_{t-1|t-1} + K_t (Z_t - D_t X_{t-1|t-1}) \\ K_t &= P_{t|t-1} D_t' (D_t P_{t|t-1} D_t' + \Sigma_{vv})^{-1} \\ P_{t+1|t} &= A_{t+1} \left(P_{t|t-1} - P_{t|t-1} D_t' (D_t P_{t|t-1} D_t' + \Sigma_{vv})^{-1} D_t P_{t|t-1} \right) A_{t+1}' + C_{t+1} C_{t+1}' \end{aligned}$$

2. **The Likelihood Function:** We can use the fact that the vector of innovations \tilde{Z}_t are conditionally independent Gaussians to write the log likelihood function as

$$L(Z \mid \Theta) = (-nT/2) \log(2\pi) - \frac{T}{2} \log |\Omega_t| - \frac{1}{2} \sum_{t=1}^T \tilde{Z}_t' \Omega_t^{-1} \tilde{Z}_t$$

where

$$\begin{aligned}\tilde{Z}_t &= Z_t - DAX_{t-1|t-1} \\ X_{t|t} &= AX_{t-1|t-1} + K_t(Z_t - DAX_{t-1|t-1}) \\ \Omega_t &= DP_{t|t-1}D' + \Sigma_{vv}\end{aligned}$$

But how do we find the MLE?

Question. How can we estimate parameters when we cannot maximize likelihood analytically? We need to (i) be able to evaluate the likelihood function for a given set of parameters, and (ii) find a way to evaluate a sequence of likelihoods conditional on different parameter vectors so that we can feel confident that we have found the parameter vector that maximizes the likelihood.

There are four main numerical maximization methods:

1. Grid search
2. Steepest ascent
3. Newton-Raphson algorithms
4. Simulated annealing

We will see examples of the first and fourth.

Example. *Unobserved Component Model* The *Unobserved Component model of inflation*

$$\pi_t = \tau_t + \eta_t \quad \text{where} \quad \tau_t = \tau_{t-1} + \varepsilon_t$$

decomposes inflation into permanent τ and transitory η components. This fits the data well, but we might be concerned about having an actual unit root in inflation on theoretical grounds.

We wait to (i) estimate the parameters of the system, meaning σ_η^2 and σ_ε^2 , and (ii) find an estimate of the permanent component τ_t at different points in time. The above state implies that $A = 1$, $D = 1$, $C = \sqrt{\sigma_\varepsilon^2}$, and $\Sigma_v = \sigma_\eta^2$. The likelihood function is the same as always, but how do we choose initial values for the Kalman recursions? Unconditional variance is infinite because of the unit root in the permanent component of inflation. A good choice is to choose ‘neutral’ values, meaning something akin to uninformative priors. One good choice is $X_{0|0} = \pi_1$ and $P_{0|0}$ very large and constant.

Since the dimension of the parameter vector is low, we can use grid search. Define a grid of variances:

$$\begin{aligned}\sigma_\varepsilon^2 &= \{0, 0.001, 0.002, \dots, \bar{\sigma}_\varepsilon^2\} \\ \sigma_\eta^2 &= \{0, 0.001, 0.002, \dots, \bar{\sigma}_\eta^2\}\end{aligned}$$

and evaluate the likelihood function for all combinations. How do we choose the boundaries of the grid? Variances are necessarily non-negative, and both of the estimates should be smaller than or equal to the sample variance of inflation, so we set $\bar{\sigma}_\eta^2 = \bar{\sigma}_\varepsilon^2 = \frac{1}{T} \sum \pi_t^2$. If we do grid search, we estimate $\hat{\sigma}_\varepsilon^2 = 0.0028$ and $\hat{\sigma}_\eta^2 = 0.0051$. This also allows us to estimate the permanent component of inflation.

Remark. With a fine enough grid, grid search always finds the global maximum as long as the parameter space is bounded. However, it is computationally infeasible for models with a large number of parameters.

Definition. *Steepest ascent methods* are for when grid search is not feasible. We need to be able to evaluate the likelihood function for a given set of parameters, and find a way to evaluate a sequence of likelihoods conditional on the difference parameter vectors so that we can feel confident that we have found the parameter vector that maximizes the likelihood. This is useful when the likelihood service may not be so well-behaved. The procedure is:

1. Make an initial guess of $\Theta = \Theta^{(0)}$

2. Find the direction of *steepest ascent* by computing the gradient

$$\mathbf{g}(\Theta) = \frac{\partial \mathcal{L}(Z | \Theta)}{\partial \Theta}$$

which is a vector which can be approximated element by element as

$$\frac{\partial \mathcal{L}(Z | \Theta^{(0)})}{\partial \theta_i} \approx \frac{\mathcal{L}(Z | \theta_j = \theta_j^{(0)} + \varepsilon : j = i; \theta_j = \theta_j^{(0)} \text{ otherwise}) - \mathcal{L}(Z | \Theta^{(0)})}{\varepsilon}$$

for each θ_j in $\Theta = \{\theta_1, \theta_2, \dots, \theta_J\}$.

3. Take a step proportional to the gradient by setting the new value of the parameter vector as

$$\Theta^{(1)} = \Theta^{(0)} + s \cdot \mathbf{G}(\Theta)$$

4. Repeat steps (2) and (3) until convergence.

Remark. This is feasible for models with a large number of parameters, but can be hard to calibrate even for simple models to achieve the right rate of convergence. Additionally, it can converge to a local maximum.

Definition. The *Newton-Raphson method* is similar to steepest ascent, but also computes the step size. The step size depends on the second derivative, and this may converge faster than steepest ascent. However, Newton-Raphson requires concavity, so it is less robust when the shape of the likelihood function is unknown.

Definition. *Simulated Annealing* comes from Goffe et al. (1994), which uses language from thermodynamics and combines elements of grid search with (strategically chosen) random movements in the parameter space. This has a good record in practice, but cannot be proven to reach a global maximum quicker than grid search.

The main algorithm inputs are $\Theta^{(0)}$, temperature T , the boundaries of Θ , the temperature reduction parameter r_T , and the function to be optimized $f(\Theta)$. The algorithm is:

1. $\theta'_j = \theta_j^{(0)} + r \cdot v_j$, where $r \sim U[-1, 1]$ and v_i is an element of the step size vector V .
2. Evaluate $f(\Theta')$ and compare with $f(\Theta^{(0)})$. If $f(\Theta') > f(\Theta^{(0)})$, set $\Theta^{(1)} = \Theta'$. If $f(\Theta') < f(\Theta^{(0)})$, set $\Theta^{(1)} = \Theta'$ with probability $\exp((f(\Theta') - f(\Theta^{(0)}))/T)$ and $\Theta^{(1)} = \Theta^{(0)}$ with probability $1 - \exp((f(\Theta') - f(\Theta^{(0)}))/T)$.
3. After N_s loops through steps (1) and (2), step length vector V is adjusted in direction so that approximately 50% of all moves are accepted.
4. After N_t loops through (1) and (3), temperature is reduced so that $T' = r_T \cdot T$ so that fewer downhill steps are accepted.

Example. *Estimating DSGE using Simulated Annealing* Recall the benchmark model:

$$\begin{aligned}\pi_t &= \beta \mathbb{E}_t \pi_{t+1} + \kappa \tilde{y}_t \\ \tilde{y}_t &= \mathbb{E}_t \tilde{y}_{t+1} - \frac{1}{\sigma} (i_t - \mathbb{E}_t \pi_{t+1} - r_t^n) \\ i_t &= \rho + \phi_\pi \pi_t + \phi_y \hat{y}_t + v_t \\ r_t^n &= \rho - \sigma(1 - \rho_a) \psi_{ya} a_t\end{aligned}$$

We want to estimate the parameters in:

$$\Theta = \{\sigma, \beta, \varphi, \varepsilon, \phi_\pi, \phi_y, \theta, \alpha, \rho_a, \rho_v, \rho_z, \sigma_a, \sigma_v, \sigma_z\}$$

The solved model can be put in state space form

$$\begin{aligned}X_t &= AX_{t-1} + C\mathbf{u}_t \\Z_t &= DX_t + \mathbf{v}_t\end{aligned}$$

where A , C , D , and Σ_v are functions of the model's 'deep' parameters. The matrix D and the covariance of \mathbf{v}_t depend on what observable variables will be included in Z_t . We can evaluate the log likelihood using the simulated annealing code provided (in `LLDSGE.m`).

Remark. In summation, the DSGE model takes as an input a vector of parameters Θ and outputs a state space system

$$f(\Theta) \rightarrow \{A, C, D, \Sigma_{vv}\}$$

The Kalman Filter estimates latent variables in the state space system, and evaluates the likelihood function for a given parameterized state space system, and an optimization algorithm finds a parameter vector Θ that maximizes the likelihood given the data.