Macroeconomic Theory

Hui He
Department of Economics
University of Hawaii at Manoa

October, 2007

Contents

Preface

This is the lecture notes for the ECON607 course that I am currently teaching at University of Hawaii. It is heavily based on Stokey, Lucas and Prescott (1989), Ljungqvist and Sargent (2004), Dirk Krueger's "Macroeconomic Theory" Lecture Notes, and Per Krusell's Lecture Notes for Macroeconomic I. I learned a lot from their books and notes. For this, I have my sincerest thanks.

I learned (and am still learning) Dynamic Macroeconomics from a series of great teachers. Among them, I would like to thank especially Selo Imrohoroglu, Caroline Betts, V.V. Chari, Larry Jones, Ed Prescott, Narayana Kocherlakota, Ellen McGrattan, and Michele Boldrin. This notes is dedicated to them.

I would also like to thank my students at UH for contributing to this notes in their way. Especially I thank John Rush and Sean Doyle for comments on the manuscript.

Part I

Foundation

Chapter 1

An Overview

1.1 How to Describe a Macro Economy?

"Macroeconomics is a language."— Ed. Prescott

Macroeconomics tries to understand the market interactions and the decisions in market settings.

The basic analytical tools for dynamic macroeconomics (or recursive macroeconomics) are:

1. maximization
2. equilibrium

How to describe an economy? An economy consists of

1. A list of objects: e.g., commodity space $L \subseteq R^l$
2. Types of Actors: $i \in I$ for households, $j \in J$ for firms
3. Consumption sets: $X_i \subset L$ and Production sets: $Y_j \subset L$
4. Endowments: $\omega_i \in L$ and profit shares $\theta_{ij} \in R_+$
5. Preference and utility function: $u_i(x_i) : X_i \to R$
6. Information structure: what information do agents possess when they make decisions?

What will happen given all these ingredients?

1. Households will maximize their preferences, subject to a budget constraint set that specifies which combination of commodities they can choose from, given the initial endowment and prices they face.

2. Firms are assumed to maximize profits, subject to their production plans being *technologically feasible*.

3. Markets interact. Supply has to be equal to demand. Therefore, we have to introduce some equilibrium concept.

Definition 1 *We call an allocation* $\{(x_i)_{i=1}^I, (y_j)_{j=1}^J\}$, $x_i \in X_i, y_j \in Y_j$ ***resource feasible*** *if*

$$\sum_{i=1}^I x_i = \sum_{j=1}^J y_j + \sum_{i=1}^I \omega_i.$$

Definition 2 *A* ***competitive equilibrium (CE)*** *is an allocation, and a price system* $p \in R_+^l$*, such that:*
(i). Household i *solves*

$$\begin{aligned} & \max_{x_i} u_i(x_i) \\ & s.t. \\ px_i \leq \; & p\omega_i + \sum_{j=1}^J \theta_{ij} p y_j \end{aligned} \tag{1.1}$$

(ii). Firm j *solves*

$$\begin{aligned} & \max_{y_j} p y_j \\ & s.t. \\ y_j \in \; & Y_j \end{aligned}$$

(iii). Allocation is resource feasible.

Definition 3 *An allocation* $\{(\hat{x}_i)_{i=1}^I, (\hat{y}_j)_{j=1}^J\}$ *Pareto Dominates* $\{(x_i)_{i=1}^I, (y_j)_{j=1}^J\}$ *if* $u_i(\hat{x}_i) \geq u_i(x_i)$*, with inequality for at least one* i*. An allocation is* ***Pareto optimal (PO)*** *when no other feasible allocation Pareto Dominates it.*

In the following lectures, we will demonstrate that a competitive equilibrium is Pareto Optimal. We will also prove that under some conditions, any PO allocation can be supported as a CE. In other words, we will establish the equivalence between CE and PO. In particular, we have following theorems.

Theorem 4 *(First Welfare Theorem): If* u_i *is strictly increasing for all* i*, competitive equilibrium allocations are Pareto Optimal.*

Theorem 5 *(Second Welfare Theorem): Under appropriate convexity assumption, any PO allocation can be supported as a CE.*

1.2 A Simple Dynamic Economy: One-Sector Optimal Growth Model

Let's be more concrete and study a simple dynamic economy. This economy consists of many (more accurately, measure one) identical, infinitely lived households (HH). Time is discrete and indexed by $t = 0, 1, 2, ...$ In each period t a single good y_t (hence the commodity space is R) is produced by many identical firms using two inputs: capital k_t, and labor n_t.

Technology is described by a production function

$$y_t = F(k_t, n_t)$$

Output can be consumed or invested (but allow free disposal)

$$y_t \geq c_t + i_t \tag{1.2}$$

This consumption-savings decision is the only allocation decision the economy must make.

Law of motion for capital stock

$$k_{t+1} = (1-\delta)k_t + i_t \tag{1.3}$$

where $\delta \in (0, 1)$ is the depreciation rate.

Preference

$$u(\{c_t\}_{t=0}^{\infty}) = \sum_{t=0}^{\infty} \beta^t U(c_t) \tag{1.4}$$

Endowments: At period 0 each HH is born with initial capital amount k_0. In addition, in each period every HH is endowed with a unit of time.

Information: No uncertainty. People have perfect foresight.

1.2.1 Optimal Growth: Planner's Problem

The question that a social planner faces is how to choose an allocation $\{(c_t)_{t=0}^{\infty}, (i_t)_{t=0}^{\infty}, (n_t)_{t=0}^{\infty}\}$ to maximize the society's welfare. Here by implicitly assuming the planner assigns equal weight to everyone, maximizing society's welfare is equivalent to maximizing the representative HH's utility, subject to the technological constraint of the economy. Before we proceed to the formal presentation of the planner's problem, let's make some further assumptions about this economy.

Assumption 1: Production function $F : R_+^2 \to R_+$ is continuously differentiable, strictly increasing, homogenous of degree one, and strictly quasi-concave, with

$$\begin{aligned}
F(0,n) &= F(k,0) = 0, F_k(k,n) > 0, F_n(k,n) > 0, \forall k, n > 0 \\
\lambda F(k,n) &= f(\lambda k, \lambda n) \quad \text{(Constant Returns to Scale, CRS)} \\
\lim_{k \to 0} F_k(k,n) &= \infty, \lim_{k \to \infty} F_k(k,n) = 0, \quad \text{(Inada Condition)}
\end{aligned}$$

(Think about what does this assumption mean in economics?)

Assumption 2: Utility function $U : R_+ \to R$ is bounded, continuously differentiable, strictly increasing, and strictly concave, with

$$\lim_{c\to 0} U'(c) = \infty, \lim_{c\to\infty} U'(c) = 0 \quad \text{(Inada Condition)}$$

We have following constraints

$$\begin{aligned} 0 &\leq n_t \leq 1, \forall t \\ c_t + k_{t+1} - (1-\delta)k_t &\leq F(k_t, n_t), \forall t \end{aligned}$$

According to Assumption 1 and 2, two features of any optimum are apparent (Why?).

$$\begin{aligned} n_t &= 1, \forall t \\ c_t + k_{t+1} - (1-\delta)k_t &= F(k_t, n_t), \forall t \end{aligned}$$

Define

$$f(k) = F(k, 1) + (1-\delta)k$$

It is easy to show that

$$f(0) = 0, f'(k) > 0, \lim_{k\to 0} f'(k) = \infty, \lim_{k\to\infty} f'(k) = 1-\delta$$

We have following social planner's problem (SP)

$$\begin{aligned} & \max_{\{c_t, k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(c_t) \\ & s.t. \\ c_t + k_{t+1} &= f(k_t), \forall t \\ c_t &\geq 0, k_{t+1} \geq 0, k_0 > 0 \text{ given} \end{aligned}$$

If we substitute $c_t = f(k_t) - k_{t+1}$, we can further simplify the SP

$$\begin{aligned} & \max_{\{k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(f(k_t) - k_{t+1}) \\ & s.t. \\ 0 &\leq k_{t+1} \leq f(k_t), \forall t \\ k_0 &> 0 \text{ given} \end{aligned} \tag{1.5}$$

Two questions:

1. Why do we want to solve this problem?
2. How do we solve it?

Below I will introduce two methods to solve this dynamic problem.

1.2.2 Solving Optimal Growth Model: The Euler Equation Approach

This is the traditional approach of solving dynamic optimization problems like the one we just proposed. We first look at a finite horizon SP and then at the related infinite-dimensional problem.

Finite Horizon Case

Suppose the horizon in (1.5) is a finite value T instead of infinity.

$$\begin{aligned} & \max_{\{k_{t+1}\}_{t=0}^{T}} \sum_{t=0}^{T} \beta^t U(f(k_t) - k_{t+1}) \\ & s.t. \\ 0 \quad \leq \quad & k_{t+1} \leq f(k_t), \forall t = 0, 1, , ..., T \\ k_0 \quad > \quad & 0 \text{ given} \end{aligned} \tag{1.6}$$

Obviously, we have $k_{T+1} = 0$. (Who cares at the end of the world!) Now we have a well-defined standard concave programing problem. Since the feasible set for sequences $\{k_{t+1}\}_{t=0}^{T}$ is closed and bounded (hence compact), and the objective function is continuous and strictly concave, according to Bolzano-Weierstrass Theorem, there exists a solution. Furthermore, since the constraint set is convex, and the utility function is strictly concave, this solution is actually unique, and it is completely characterized by the Kuhn-Tucker conditions.

Since $f(0) = 0$ and $U'(0) = \infty$, it is clear that the inequality constraints in (1.6) do not bind, and we will have interior solutions. Hence we have first order condition (by directly solving this unconstrainted optimization problem)

$$U'(f(k_t) - k_{t+1}) = \beta U'(f(k_{t+1}) - k_{t+2}) f'(k_{t+1}), t = 0, 1, 2, ..., T \tag{1.7}$$

Question: What is the economic sense of this Euler equation?

How to solve this second-order difference equation? Here is a simple example which allows analytical solution for the SP problem.

Suppose

$$U(c) = \ln c, f(k) = k^{\alpha}, 0 < \alpha < 1$$

We have planner's problem

$$\begin{aligned} & \max_{\{k_{t+1}\}_{t=0}^{T}} \sum_{t=0}^{T} \beta^t U(k_t^{\alpha} - k_{t+1}) \\ & s.t. \\ 0 \quad \leq \quad & k_{t+1} \leq k_t^{\alpha}, t = 0, 1, 2, ..., T \\ k_{T+1} \quad = \quad & 0, k_0 > 0 \text{ given} \end{aligned}$$

Euler equation (or first-order equation FOC) for this problem is

$$\frac{1}{k_t^\alpha - k_{t+1}} = \frac{\beta\alpha k_{t+1}^{\alpha-1}}{k_{t+1}^\alpha - k_{t+2}} \tag{1.8}$$

Now let's use the transformation $z_t = \frac{k_{t+1}}{k_t^\alpha}$ to transform (1.8) to a first-order difference equation in z_t

$$z_{t+1} = 1 + \alpha\beta - \frac{\alpha\beta}{z_t} \tag{1.9}$$

Since we have a boundary condition $z_T = 0$. Start from it, do backward induction (leave as an exercise), we can obtain a general formula

$$z_t = \alpha\beta\frac{1-(\alpha\beta)^{T-t}}{1-(\alpha\beta)^{T-t+1}}$$

and hence

$$\begin{aligned} k_{t+1} &= \alpha\beta\frac{1-(\alpha\beta)^{T-t}}{1-(\alpha\beta)^{T-t+1}}k_t^\alpha \\ c_t &= \frac{1-(\alpha\beta)}{1-(\alpha\beta)^{T-t+1}}k_t^\alpha \end{aligned}$$

Infinite Horizon Case

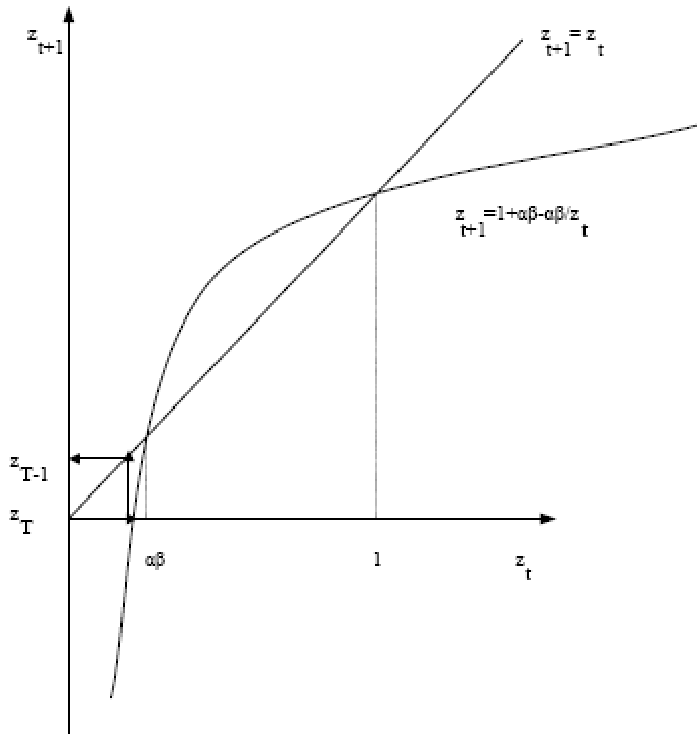
When $T \to \infty$, we have the infinite horizon case. A nature guess is the solution for infinite horizon case is the limit of finite horizon case, i.e., now

$$\begin{aligned} k_{t+1} &= \lim_{T\to\infty} \alpha\beta\frac{1-(\alpha\beta)^{T-t}}{1-(\alpha\beta)^{T-t+1}}k_t^\alpha = \alpha\beta k_t^\alpha \\ c_t &= \lim_{T\to\infty} \frac{1-(\alpha\beta)}{1-(\alpha\beta)^{T-t+1}}k_t^\alpha = (1-\alpha\beta)k_t^\alpha \end{aligned}$$

The conjecture actually is correct.

Notice that in the infinite horizon case, the policy function should be time invariant, hence we should have $z_t = z_{t+1.}$ In steady state, equation (1.9) has two steady states. $z_t = 1$ or $z_t = \alpha\beta$. We can use the following Figure 1 to show intuitively that only $z_t = \alpha\beta$ is a stable steady state.

Starting from $z_{T+1} = 0$ on y-axis, by (1.9), we have $z_T = \frac{\alpha\beta}{1+\alpha\beta} < \alpha\beta$ (it is the intersection of z_{t+1} curve on x-axis). Mirror it against 45^0 line, on y-axis we have the value of z_T. Then Starting again from z_T on y-axis, we have z_{T-1} closer to $\alpha\beta$. Repeating this procedure we can obtain the entire $\{z_t\}_{t=0}^T$ sequence and hence $\{k_{t+1}\}_{t=0}^T$. Note that when t going backwards to zero, z_t approaches $\alpha\beta$. Therefore, when $T \to \infty$, eventually the economy will go to the steady state where $z_t = z = \alpha\beta, \forall t$. By the same argument, it is easy to show that when we do backwards, starting from z_{t+1} around $z = 1$, it will always diverge from it. Thus $z = 1$ is not a stable steady state.



Dynamics for the Simple Economy

For a formal treatment of Euler equation approach on the infinite horizon case, please see Krueger (2002) P. 45-48.

1.2.3 Solving Optimal Growth Model: The Dynamic Programming Approach

Bellman Equation

Define value function

$$v_0(k_0) = \max_{\substack{\{k_{t+1}\}_{t=0}^{\infty} \\ s.t. 0 \le k_{t+1} \le f(k_t), k_0 \text{ given}}} \sum_{t=0}^{\infty} \beta^t U(f(k_t) - k_{t+1})$$

Then we have

$$\begin{aligned} v_0(k_0) &= \max_{\substack{\{k_{t+1}\}_{t=0}^{\infty} \\ s.t. 0 \le k_{t+1} \le f(k_t), k_0 \text{ given}}} \sum_{t=0}^{\infty} \beta^t U(f(k_t) - k_{t+1}) \\ &= \max_{\substack{\{k_{t+1}\}_{t=0}^{\infty} \\ s.t. 0 \le k_{t+1} \le f(k_t), k_0 \text{ given}}} \left\{ U(f(k_0) - k_1) + \beta \sum_{t=1}^{\infty} \beta^{t-1} U(f(k_t) - k_{t+1}) \right\} \\ &= \max_{\substack{k_1 \\ s.t. 0 \le k_1 \le f(k_0), k_0 \text{ given}}} \left\{ U(f(k_0) - k_1) + \beta [\max_{\substack{\{k_{t+1}\}_{t=1}^{\infty} \\ s.t. 0 \le k_{t+1} \le f(k_t), k_1 \text{ given}}} \sum_{t=0}^{\infty} \beta^t U(f(k_t) - k_{t+1})] \right\} \\ &= \max_{0 \le k_1 \le f(k_0), k_0 \text{ given}} \{ U(f(k_0) - k_1) + \beta v_0(k_1) \} \end{aligned}$$

Notice that when the problem is viewed in this recursive way, the time subscripts have become a nuisance. Therefore, we can ignore the time subscript and rewrite the SP problem with current capital stock k as

$$v(k) = \max_{0 \le y \le f(k)} \{ U(f(k) - y) + \beta v(y) \} \tag{1.10}$$

This equation in the unknown function v is called a *functional equation (or Bellman equation).* The study of dynamic optimization problems through the analysis of such functional equations (FE) is called *dynamic programming (DP).*

The current capital stock k is the "state variable" because it completely determines what allocations are feasible from today onwards. The next period capital stock k' is decided (or controlled) by the social planner, it is called the "control variable".

Solving Bellman equation means finding a value function v solving (1.10) and an optimal policy function

$$k' = g(k)$$

But we face several questions associated with Bellman equation (1.10):

1. Under what condition does a solution to (1.10) exist? And if it exists, is it unique?

2. Is there a reliable algorithm that computes the solution?

3. Under what condition can we solve Bellman equation (1.10) and be sure it is also the solution to the planner's problem (1.5)?

4. Can we say something about qualitative properties of v and g?

As a first try, if we know the value function v is differentiable and $y = g(k)$ is interior, first order condition for (1.10) is

$$U'(f(k) - g(k)) = \beta v'[g(k)] \tag{1.11}$$

And the envelope condition is

$$v'(k) = f'(k) U(f(k) - g(k)) \tag{1.12}$$

Substituting (1.12) into (1.11), we have

$$\begin{aligned} U'(f(k)-g(k)) &= \beta f'(g(k))U[f(g(k))-g(g(k))] \\ &\text{i.e.} \\ U'(f(k)-k') &= \beta f'(k')U[f(k')-k''] \end{aligned}$$

This equation should remind you the Euler equation (1.7). Coincidence? It actually hints on the equivalence between SP and FE.

An Example

Consider the previous example $U(c) = \ln c$, $f(k) = k^{\alpha}$. In this economy, the Bellman equation is

$$v(k) = \max_{0\le k'\le k^{\alpha}} \{\ln(k^{\alpha}-k') + \beta v(k')\} \tag{1.13}$$

How can we solve this value function? We are going to do it by using a clever approach called "Guess and Verify".

We will directly guess a particular functional form of the solution to value function v. Let's guess

$$v(k) = A + B\ln(k) \tag{1.14}$$

where A and B are the two coefficients to be determined.

Substituting (1.14) into (1.13), we have

$$v(k) = \max_{0\le k'\le k^{\alpha}} \{\ln(k^{\alpha}-k') + \beta(A + B\ln k')\} \tag{1.15}$$

The FOC for right hand side (RHS) is

$$\frac{1}{k^{\alpha}-k'} = \frac{\beta B}{k'}$$

This leads to

$$k' = \frac{\beta B k^{\alpha}}{1+\beta B} \tag{1.16}$$

Now let's substitute the optimal policy function (1.16) back into Bellman equation (1.15), we have

$$\begin{aligned} v(k) &= \{\ln(k^{\alpha}-k') + \beta(A + B\ln k')\} \\ &= \{\ln(k^{\alpha} - \frac{\beta B k^{\alpha}}{1+\beta B}) + \beta(A + B\ln\frac{\beta B k^{\alpha}}{1+\beta B})\} \\ &= \ln\frac{k^{\alpha}}{1+\beta B} + \beta A + \beta B\ln\frac{\beta B k^{\alpha}}{1+\beta B} \\ &= \beta A - (1+\beta B)\ln(1+\beta B) + \ln\beta B + \alpha(1+\beta B)\ln k \end{aligned}$$

Now substituting our guess (1.14) for $v(k)$, we have

$$A + B\ln(k) = \beta A - (1+\beta B)\ln(1+\beta B) + \ln\beta B + \alpha(1+\beta B)\ln k$$

which means

$$\begin{aligned}(B - \alpha(1 + \beta B)) \ln k &= (\beta - 1)A - (1 + \beta B)\ln(1 + \beta B) + \ln \beta B \\ &= \text{constant}\end{aligned}$$

It implies

$$B - \alpha(1 + \beta B) = 0 \Rightarrow B = \frac{\alpha}{1 - \alpha\beta}$$

Accordingly, we have

$$A = \frac{1}{1-\beta}[\frac{\alpha\beta}{1-\alpha\beta}\ln(\alpha\beta) + \ln(1 - \alpha\beta)]$$

Therefore, the optimal policy function is

$$\begin{aligned}k' &= \frac{\beta B k^{\alpha}}{1 + \beta B} \\ &= \alpha\beta k^{\alpha}\end{aligned}$$

Look familiar? Yes, that is exactly the same as the one solved by Euler equation approach.

1.3 A Stochastic Optimal Growth Model

Now let's relax our assumption about information structure. In this section, we will discuss a optimal growth model with uncertainty.

1.3.1 Markov Process

Markov shock is a stochastic process with the following properties:

1. There are finite number of possible states for each time. $s_t \in S = \{e_1, e_2, ..., e_N\}$. More importantly, no matter what happened before, tomorrow will be represented by one element of a finite set.

2. What only matters for the realization tomorrow is today's state. More intuitively, no matter what kind of history we have, the only thing we need to predict realization of shock tomorrow is today's realization.

$$\Pr(s_{t+1} \mid s_t, s_{t-1}, s_{t-2}, ..., s_{t-k}) = \Pr(s_{t+1} \mid s_t)$$

Define transition probability from state e_i at time t to state e_j at time $t + 1$ as

$$\pi_{ij} = \Pr(s_{t+1} = e_j \mid s_t = e_i).$$

π_{ij} is one generic element of a $N \times N$ matrix called *transition matrix*. We have $\pi_{ij} \geq 0$ and $\sum_{j=1}^{N} \pi_{ij} = 1, \forall i$.

Denote the *history* of events up to time t by $s^t = \{s_1, s_2, ..., s_t\}$. It is easy to show that

$$\pi(s^{t+1}) = \pi(s_{t+1} \mid s_t) \times \pi(s_t \mid s_{t-1}) \times \pi(s_{t-1} \mid s_{t-2}) \times \cdots \times \pi(s_1 \mid s_0) \times \pi(s_0)$$

1.3.2 A Stochastic Model of Optimal Growth

Consider an optimal growth model in which the uncertainty affects the technology only. Assume the production function

$$y_t = z_t f(k_t)$$

where $\{z_t\}$ is a sequence of independently and identically distributed (i.i.d.) random variables. Therefore, the SP becomes

$$\begin{aligned} & \max E[\sum_{t=0}^{\infty} \beta^t U(c_t)] \\ & s.t. \\ c_t + k_{t+1} &\leq z_t f(k_t), \\ c_t, k_{t+1} &\geq 0, k_0 > 0 \text{ given} \end{aligned}$$

Timing of the model: in each period

1. At the beginning of the period t, shock z_t is revealed
2. Output is known, HHs make consumption-savings decision

Thus, the state variable is the pair (k_t, z_t). The social planner does not choose a sequence of numbers anymore rather a sequence of *contingency plans* $\{(c_t(z^t), (k_{t+1}(z^t)\}_{t=0}^{\infty}$.

We can also write down the Bellman equation for this economy.

$$v(k, z) = \max_{0 \leq y \leq f(k)} \{U(zf(k) - y) + \beta E[v(y, z')]\}$$

The optimal capital path is given by

$$\begin{aligned} y &= g(k, z) \\ & \quad or \\ k_{t+1} &= g(k_t, z_t) \end{aligned} \tag{1.17}$$

Equation (1.17) is called *first-order stochastic difference equation.* Obviously the $\{k_{t+1}\}_{t=0}^{\infty}$ generated by this equation is a (first-order) Markov process.

In Chapter 4, we will learn how to solve this type of stochastic dynamic programming problem by using numerical methods.

1.4 Competitive Equilibrium Growth

In this section, I will show that the solutions to social planner's problem can, under appropriate conditions, be interpreted as predictions about the behaviors of market economy. After all, what we are really interested in are allocations and prices arise when firms and consumers interact in markets.

Suppose that we have solved the infinite-horizon optimal growth problem (1.5) and obtained solution $\{(c_t, k_{t+1})\}_{t=0}^{\infty}$. Our goal is to find prices to support these quantities as a competitive equilibrium. However, in order to do this, we have to first specify the ownership rights of households and firms, as well as the structure of markets.

Assume that HHs own all factors of production and all shares in firms. Firms own nothing; they simply hire capital and labor on a rental basis to produce output every period, sell the output produced back to HHs, and return any profits to the shareholders. Finally, assume that all transactions take place in a single once-for-all market that meets in period 0. All trading takes place at that time, so all prices and quantities are determined simultaneously. No further trades are negotiated later. After this market has closed, in the subsequent periods, agents simply deliver the contracts at period 0. We assume that all contracts are perfectly enforceable. This market is often called *Arrow-Debreu* market structure. And the corresponding competitive equilibrium is called *Arrow-Debreu Equilibrium* (ADE).

1.4.1 Definition of Competitive Equilibrium

Given prices $\{p_t, r_t, w_t\}_{t=0}^{\infty}$, the firm's problem is to choose $\{(k_t^d, n_t^d, y_t)\}_{t=0}^{\infty}$ to

$$\begin{aligned} \max \pi &= \sum_{t=0}^{\infty} p_t[y_t - r_t k_t^d - w_t n_t^d] \qquad (1.18) \\ & s.t. \\ y_t &\leq F(k_t^d, n_t^d), t = 0, 1, 2, ... \\ y_t, k_t^d, n_t^d &\geq 0 \end{aligned}$$

Given prices $\{p_t, r_t, w_t\}_{t=0}^{\infty}$, the HH's problem is to choose $\{(c_t, i_t, x_{t+1}, k_t^s, n_t^s)\}_{t=0}^{\infty}$ to

$$\max \sum_{t=0}^{\infty} \beta^t U(c_t) \qquad (1.19)$$

$$\begin{aligned} & s.t. \\ \sum_{t=0}^{\infty} p_t[c_t + i_t] &\leq \sum_{t=0}^{\infty} p_t[r_t k_t^s + w_t n_t^s] + \pi \\ x_{t+1} &= (1-\delta)x_t + i_t, \ \forall t = 0, 1, ... \\ 0 &\leq n_t^s \leq 1, 0 \leq k_t^s \leq x_t, \ \forall t = 0, 1, ... \\ c_t, x_{t+1} &\geq 0, \ \forall t = 0, 1, ... \\ & x_0 \text{ given} \end{aligned}$$

Remark 6 *Note that p_t is the price of a unit of time-t good in terms of time-0 good (Arrow-Debreu price). While w_t and r_t are the prices of a unit of labor and rental capital at time t in terms of time-t good.*

Remark 7 *Note the difference between capital owned x_t and capital supplied to firms k_t^s.*

Definition 8 *A competitive equilibrium is a set of prices $\{(p_t, r_t, w_t)\}_{t=0}^{\infty}$, an allocation $\{(k_t^d, n_t^d, y_t)\}$ for the representative firm, and an allocation $\{(c_t, i_t, x_{t+1}, k_t^s, n_t^s)\}_{t=0}^{\infty}$ for the representative household, such that:*

a. Given prices $\{p_t, r_t, w_t\}_{t=0}^{\infty}$, $\{(k_t^d, n_t^d, y_t)\}$ solves the firm's problem (1.18).

b. Given prices $\{p_t, r_t, w_t\}_{t=0}^{\infty}$, $\{(c_t, i_t, x_{t+1}, k_t^s, n_t^s)\}_{t=0}^{\infty}$ solves the representative HH's problem (1.19).

c. All markets clear: $k_t^d = k_t^s$, $n_t^d = n_t^s$, $c_t + i_t = y_t$, for all t.

Question: How to find a CE?

1.4.2 Characterization of the Competitive Equilibrium

Let's start from a couple of conjectures. First let's simply denote

$$\begin{aligned} k_t &= k_t^d = k_t^s \\ n_t &= n_t^d = n_t^s \end{aligned}$$

Since $U'(c) > 0$, our first guess is $p_t > 0$. (Think about why?) Similarly, since $F_k(k, n) > 0, F_n(k, n) > 0$, we should also have $r_t, w_t > 0$.

Now consider the representation firm. Since prices are strictly positive, we have

$$y_t = F(k_t, n_t), \forall t$$

And the firm does not face a dynamic decision problem (there is no dynamic state variable in firm's problem as k_t in HH's problem), its problem is equivalent to a series of one-period maximization problems.

$$\max_{k_t, n_t \geq 0} \pi_t = p_t[F(k_t, n_t) - r_t k_t - w_t n_t], t = 0, 1, 2, ... \tag{1.20}$$

FOCs are

$$\begin{aligned} r_t &= F_k(k_t, n_t), \forall t \\ n_t &= F_n(k_t, n_t), \forall t \end{aligned}$$

Since the production function $F(\cdot, \cdot)$ is homogeneous degree of one (h.o.d.1), according to Euler's Theorem, we have

$$F(k_t, n_t) = r_t k_t + w_t n_t$$

which implies

$$\pi_t = 0, \forall t$$

Remark 9 *With h.o.d.1, the total profits of the representative firm are equal to zero in equilibrium. This shows that the number of firms in indeterminate in equilibrium. We can view the representative firm is the aggregation of a* *whole bunch (maybe millions) of firms. Therefore, this really justifies that the assumption of a single representative firm is without any loss of generality (as long as we assume that this firm is a price-taker).*[1]

Next we consider the representative HH. Given all the prices are strictly positive, we should have

$$n_t = 1, k_t = x_t, \forall t$$

Also notice that the utility function is strictly increasing, therefore, the budget constraint (BC) will hold with equality. We can rewrite the HH's problem as following

$$\begin{aligned} & \max \sum_{t=0}^{\infty} \beta^t U(c_t) \\ & s.t. \\ \sum_{t=0}^{\infty} p_t[c_t + k_{t+1} - (1-\delta)k_t] &= \sum_{t=0}^{\infty} p_t(r_t k_t + w_t) \quad (1.21) \\ c_t, k_{t+1} &\geq 0, \forall t \\ & k_0 \text{ given} \end{aligned}$$

FOC (or Euler Equation) is

$$\frac{\beta U'(c_{t+1})}{U'(c_t)} = \frac{p_{t+1}}{p_t} = \frac{1}{r_{t+1} + 1 - \delta} \quad (1.22)$$

Or

$$U'(c_t) = \beta U'(c_{t+1})(r_{t+1} + 1 - \delta) \quad (1.23)$$

As we define earlier

$$f(k_t) = F(k_t, 1) + (1-\delta)k_t$$

Therefore

$$\begin{aligned} f'(k_t) &= F_k(k_t, 1) + 1 - \delta \\ &= r_t + 1 - \delta \end{aligned}$$

[1] Note that

$$\begin{aligned} \sum_{j=1}^{J} F(k_{jt}, n_{jt}) &= \sum_{j=1}^{J} (r_t k_{jt} + w_t n_{jt}) \\ &= r_t \sum_{j=1}^{J} k_{jt} + w_t \sum_{j=1}^{J} n_{jt} \\ &= F(\sum_{j=1}^{J} k_{jt}, \sum_{j=1}^{J} n_{jt}). \end{aligned}$$

Substituting into Euler equation (1.23)

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

Looks familiar? This is quite close to the one in the social planner's problem. Next, let me show that it is actually same.

1.4.3 More on Two Welfare Theorems

Let's be more specific about the connections between equilibrium and optimal allocations. First Welfare Theorem says that any CE allocation is PO. In order to prove it, if $\{(c_t^e, k_{t+1}^e, p_t^e, r_t^e, w_t^e)\}_{t=0}^{\infty}$ is a CE, we have to show that $\{(c_t^e, k_{t+1}^e)\}_{t=0}^{\infty}$ is PO.

We use contradiction. Suppose that $\{(c_t^e, k_{t+1}^e)\}_{t=0}^{\infty}$ is not PO. Suppose that there is another feasible allocation $\{(c_t', k_{t+1}')\}_{t=0}^{\infty}$ that $\{c_t'\}$ yields higher total utility in the objective function. Then this allocation must violate BC of (1.21), because otherwise the household would have chosen it. (Recall the Weak Axiom of Revealed Preference, see MWG P. 29) But if BC is violated, we have

$$\begin{aligned} \pi' &= \sum_{t=0}^{\infty} p_t [c_t' + i_t' - r_t k_t' - w_t] \\ &= \sum_{t=0}^{\infty} p_t [c_t' + k_{t+1}' - (1-\delta) k_t' - r_t k_t' - w_t] \\ &> 0 \end{aligned}$$

contracting the hypothesis that $\{(k_t^e, n_t^e = 1)\}$ was a profit-maximizing choice of inputs.

How about the Second Welfare Theorem? We want to show that under some appropriate conditions, any PO allocation can be supported as a CE. Suppose $\{(c_t^*, k_{t+1}^*)\}_{t=0}^{\infty}$ is a solution to the planner's problem, hence it is PO. Then we know that $\{(k_{t+1}^*)\}_{t=0}^{\infty}$ is the unique sequence satisfying the following FOC

$$U'(f(k_{t-1}^*) - k_t^*) = \beta f'(k_t^*) U'[f(k_t^*) - k_{t+1}^*], \forall t$$

and $\{c_t^*\}$ is given by

$$c_t^* = f(k_t^*) - k_{t+1}^*, \forall t$$

To construct a CE, we have to find supporting prices $\{(p_t^*, r_t^*, w_t^*)\}_{t=0}^{\infty}$. From the pricing kernel (1.22) we already know that the good prices must satisfy

$$p_t^* = \frac{p_{t-1}^*}{f'(k_t^*)}$$

Therefore, starting from an arbitrary initial price p_0 (w.l.o.g, assume $p_0 = 1$), once we know $\{(k_{t+1}^*)\}_{t=0}^{\infty}$, we can construct a sequence of good prices $\{p_t^*\}_{t=0}^{\infty}$. Next, we construct factor equilibrium prices as follows

$$\begin{aligned} r_t^* &= f'(k_t^*) - (1-\delta) \\ w_t^* &= f(k_t^*) - k_t^* f'(k_t^*) \end{aligned}$$

It is easy to show that with the allocation $\{(c_t^*, k_{t+1}^*)\}_{t=0}^{\infty}$ and the constructed prices $\{(p_t^e, r_t^e, w_t^e)\}_{t=0}^{\infty}$, the Euler equation for Planner's problem becomes

$$U'(c_t^*) = \beta U'(c_{t+1}^*)(r_{t+1}^* + 1 - \delta)$$

Furthermore, we can show that BC of HH's problem is as same as resource constraint of Planner's problem. Thus these two problems share same constraint set, same objective function and same FOCs. Since the solution to CE and planner problem is unique, therefore, it has to be same.

1.4.4 Sequential Markets Equilibrium (SME)

We all know that ADE has a very weird market arrangement. Agents can make a choice ONCE in Arrow-Debreu world. Now let's relax this unrealistic assumption to make it a little bit closer to the real world. Suppose that the agents meet in a market at the beginning of every period. In the market held in period t, agents trade current-period labor, rental services of existing capital, and final output. In addition, one security (called *Arrow security*) is traded: a claim to one unit of final output in the subsequent period. In each period, factor and bond prices are expressed in terms of current-period goods.

We will assume that people have perfect foresight. Given this expectations hypothesis, the firm's problem is as same as in ADE. But the HH's problem changes

$$\begin{aligned} & \max \sum_{t=0}^{\infty} \beta^t U(c_t) \\ & s.t. \\ c_t + q_t s_{t+1} + k_{t+1} - (1-\delta)k_t \ &\leq \ r_t k_t + w_t n_t + s_t + \pi_t, \forall t \qquad (1.24) \\ 0 \ &\leq \ n_t \leq 1, \forall t \\ c_t, k_{t+1} \ &\geq \ 0 \end{aligned}$$

where s_t is the quantity held of a one period bond. (Think about why I do not put nonnegativity constraint on s_t)

Definition 10 *A sequential market competitive equilibrium (SME) is a set of prices $\{(q_t, r_t, w_t)\}_{t=0}^{\infty}$, an allocation $\{(k_t^d, n_t^d, y_t)\}$ for the representative firm, and an allocation $\{(c_t, i_t, s_{t+1}, k_t^s, n_t^s)\}_{t=0}^{\infty}$ for the representative household, such that:*

a. Given prices $\{q_t, r_t, w_t\}_{t=0}^{\infty}$, $\{(k_t^d, n_t^d, y_t)\}$ solves the firm's problem (1.20).

b. Given prices $\{q_t, r_t, w_t\}_{t=0}^{\infty}$, $\{(c_t, i_t, s_{t+1}, k_t^s, n_t^s)\}_{t=0}^{\infty}$ solves the representative HH's problem (1.24).

c. All markets clear: $k_t^d = k_t^s$, $n_t^d = n_t^s$, $c_t + i_t = y_t$, $s_{t+1} = 0$ for all t.

FOCs for HH's problem:

$$\begin{aligned} c_t \ &: \ \beta^t U'(c_t) = \lambda_t \\ k_{t+1} \ &: \ \lambda_{t+1}(r_{t+1} + 1 - \delta) = \lambda_t \\ s_{t+1} \ &: \ \lambda_t q_t = \lambda_{t+1} \end{aligned}$$

Combining FOCs, we have

$$\begin{aligned} q_t &= \frac{\beta U'(c_{t+1})}{U'(c_t)} \\ U'(c_t) &= \beta U'(c_{t+1})(r_{t+1} + 1 - \delta) \end{aligned}$$

This is exactly as same as ones in ADE. Therefore, we set up the equivalence between SME and ADE.

The intuition for the equivalence is as following. since in the equilibrium, the net supply of bonds is zero, the representative household has a net demand of zero for each of the securities. Hence, if bonds markets are simply shut down, the remaining prices and the real allocation are unaltered. Therefore, we can price Arrow security even though there is not a trade in equilibrium. Extending this idea, we can determine prices of all market instruments even though they are redundant in equilibrium. This is the virtue of Lucas Tree model and it is the fundamental for all finance literature.

It is not hard to show that if we substitute out s_t from sequential BC, we end up with the same BC as in ADE. We leave this as an exercise.

Remark 11 *If markets are complete, then normally ADE is easier to be solved. If markets are not complete, then SME is the tool that you should use.*

1.4.5 Recursive Competitive Equilibrium (RCE)

There is yet a third way in which the solution to the optimal growth model can be interpreted as a competitive equilibrium, one that is closely related to the dynamic programming approach and to the sequence of markets interpretation of equilibrium. The general idea is to characterize equilibrium prices as functions of the economy-wide capital stock k, and to view individual HHs as dynamic programmers whose state variables include both individual level and the economy-wide ones. Notice that in this subsection, as I showed equivalence between sequential problem (SP) and functional equation (FE) of social planner's problemin the previous sections, I am going to show the equivalence among three equilibrium concepts: ADE, SME and RCE.

First, the factor prices can be expressed as the function of state variable k.

$$R(k) = F_k(k, 1) \tag{1.25}$$

$$\varpi(k) = F_n(k, 1) \tag{1.26}$$

Next, I have to develop a DP representing the decision problem faced by a typical HH. To do so, we need to distinguish between the economy-wide capital stock k, and its own capital stock K. Therefore, for individual HH, the state variable is the pair (K, k).

Suppose the economy-wide capital stock accumulates according to law of motion

$$k' = h(k)$$

The HH's problem is

$$V(K,k) = \max_{C,Y}\{U(C) + \beta V[Y, h(k)]\} \tag{1.27}$$

subject to

$$C + [Y - (1-\delta)K] \leq KR(k) + \varpi(k)$$

Let

$$Y = H(K,k)$$

be the optimal policy function for the HH.

Definition 12 *A Recursive Competitive Equilibrium (RCE) is a value function $V : R_+^2 \to R$, a policy function $H : R_+^2 \to R$ for the representative household, an economy-wide law of motion $h : R_+ \to R_t$ for capital, and factor price functions $R : R_+ \to R_+$ and $\varpi : R_+ \to R_+$ such that*

i. Given h, R, ϖ, V solves HH's problem and H is the optimal policy function of HH's problem.

ii. R and ϖ satisfy (1.25) and (1.26).

iii. Individual choice is consistent with the aggreagte law of motion (consistency condition)

$$H(k,k) = h(k), \forall k$$

FWT of RCE: If (V, H, h, R, ϖ) is a RCE, then $v(k) = V(k,k)$ is the value function for social planner's problem, and $g = h$ is the planner's optimal policy function.

SWT of RCE: If v is the value function for the planner's problem and g is the optimal policy function, then the value and policy function for the individual household (V, H) satisfies (1.27) and have properties that $V(k,k) = v(k)$ and $H(k,k) = h(k)$.

This equivalence between the planner's problem and RCE can be shown through a comparison of the first-order and envelope conditions for the two dynamic programs. Recall for the planner's problem, they are

$$\begin{aligned} y &: U'(f(k) - g(k)) = \beta v'[g(k)] \\ k &: v'(k) = f'(k)U'(f(k) - g(k)) \end{aligned}$$

For the HH's problem, the analogous conditions are

$$\begin{aligned} Y &: U'[f(k) + f'(k)(K-k) - H(K,k)] = \beta V_1[H(K,k), h(k)] \\ K &: V_1(K,k) = U'[f(k) + f'(k)(K-k) - H(K,k)]f'(k) \end{aligned}$$

In equilibrium, we have $K = k$. Since $H(k,k) = h(k)$, these conditions can be rewritten as

$$\begin{aligned} U'[f(k) - h(k)] &= \beta v_1[h(k), h(k)] \\ v_1(k,k) &= f'(k)U'(f(k) - h(k)) \end{aligned}$$

Thus, if $v_1(k,k) = v'(k)$, the equilibrium conditions of RCE match the conditions for planner's problem.

1.5 Conclusions

Where do we stand now?

We show that for social planner's problem, SP=FE. A sequential representation of planner's problem can be solved using Dynamic Programming approach.

Then we show that solutions to planner's problem, which is PO, is equivalent to allocations in ADE by using two welfare theorems.

We also show the equivalence among ADE, SME, and RCE.

In the following two chapters, we will have more rigorous treatment to establish the equivalence between SP and FE and introduce the DP technique.

Chapter 2

Mathematical Preliminaries

The purpose of this chapter and the next is to show precisely the relationship between the sequential problem (SP) and the functional equation representation of the planner's problem.

Recall the Functional equation

$$\begin{aligned} v(k) &= \max_{c,y}[U(c) + \beta v(y)] \\ & \quad s.t. \\ c + y &\leq f(k) \\ c, y &\geq 0 \end{aligned}$$

In order to study FE, define the following operator T

$$(Tv)(k) = \max_{0 \leq y \leq f(k)} [U(f(k) - y) + \beta v(y)]$$

This operator takes the function v as input and returns a new function Tv. A solution to FE is thus a fixed point of this operator, i.e., a function v^* such that

$$v^* = Tv^*$$

We want to study under what conditions the operator T has a fixed point (existence), under what conditions it is unique (uniqueness), and how can we obtain this fixed point (classic technique is called *method of successive approximation*, start from an arbitrary function v, apply the operator T repeatedly to get convergence of v.)

In Section 1, we will review the basic facts about metric spaces and normed spaces. In Section 2, we will state and prove the Contraction Mapping Theorem (CMT), a very useful fixed-point theorem for FE. In Section 3, we will prove Theorem of Maximum, which is very useful for the next chapter.

2.1 Metric Spaces and Normed Vector Spaces

To be skipped. Professor Nori Tarui will teach you these stuff in ECON627.

2.2 The Contraction Mapping Theorem

In this section we will prove two main results. The first one is Contraction Mapping Theorem (CMT), which gives us the existence and uniqueness of a fixed point of operator T. The second is a set of sufficient conditions, due to Blackwell, for establishing that certain operators are contraction mapping.

Definition 13 *Let (S, ρ) be a metric space and $T : S \to S$ be a function mapping S into itself. The function T is a **contraction mapping** if there exists a number $\beta \in (0, 1)$ satisfying $\rho(Tx, Ty) \leq \beta\rho(x, y)$ for all $x, y \in S$. The number β is called the modulus of the contraction mapping.*

Example 14 *$S = [a, b]$, with $\rho(x, y) = |x - y|$. Then $T : S \to S$ is a contraction mapping if for some $\beta \in (0, 1)$*

$$\frac{|Tx - Ty|}{|x - y|} \leq \beta < 1, \text{ all } x, y \in S \text{ with } x \neq y$$

That is, T is a contraction mapping if it is a function with slope uniformly less than one in absolute value.

It is easy to show that if T is a contraction mapping, then T is continuous. We leave this as an exercise.

We now state and prove CMT.

Theorem 15 *(Contraction Mapping Theorem) Let (S, ρ) be a complete metric space and suppose that $T : S \to S$ is a contraction mapping with modulus β. Then*

a). the operator T has exactly one fixed point $v \in S$ and

b). for any $v_0 \in S$, and any $n \in N$ we have

$$\rho(T^n v_0, v) \leq \beta^n \rho(v_0, v)$$

Part (a) of the Theorem establishs the existence and uniqueness of the fixed point for the contraction mapping. Part (b) assets that no matter what initial guess v_0 is, the sequence $\{v_n\}_{n=0}^{\infty}$ will converge to the fixed point v at a geometric rate of β. This part is extremely important for computation. It is the theoretical foundation of value function iteration method for DP.

Proof. To prove (a), we must find a candidate for v, show that it satisfies $Tv = v$, and show that no other element $\hat{v} \in S$ does.

Define the iterates of T, the mappings $\{T^n\}$, by $T^0 x = x$, and $T^n x = T(T^{n-1}x), n = 1, 2, ...$ Let's choose an arbitrary $v_0 \in S$, and define $\{v_n\}_{n=0}^{\infty}$ by $v_{n+1} = Tv_n$, so that $v_n = T^n v_0$. By the contraction property of operator T,

$$\rho(v_2, v_1) = \rho(Tv_1, Tv_0) \leq \beta\rho(v_1, v_0).$$

By induction

$$\begin{aligned}
\rho(v_{n+1}, v_n) &= \rho(Tv_n, Tv_{n-1}) \leq \beta\rho(v_n, v_{n-1}) \\
&= \beta\rho(Tv_{n-1}, Tv_{n-2}) \leq \beta^2\rho(v_{n-1}, v_{n-2}) \\
&= ...... = \beta^n \rho(v_1, v_0)
\end{aligned}$$

Therefore, for any $m > n$,

$$\begin{aligned}
\rho(v_m, v_n) &\leq \rho(v_m, v_{m-1}) + \rho(v_{m-1}, v_{m-2}) + ... + \rho(v_{n+2}, v_{n+1}) + \rho(v_{n+1}, v_n) \\
&\leq [\beta^{m-1} + \beta^{m-2} + ... + \beta^{n+1} + \beta^n]\rho(v_1, v_0) \\
&= \beta^n[\beta^{m-n-1} + ... + \beta + 1]\rho(v_1, v_0) \\
&\leq \frac{\beta^n}{1-\beta}\rho(v_1, v_0)
\end{aligned}$$

First inequality comes from triangle inequality of metric space. Second one comes from the contraction property we just derived. When $n \to \infty$, we can make $\frac{\beta^n}{1-\beta}\rho(v_1, v_0) < \varepsilon$, where ε is an arbitrary small number. Therefore, by definition (see SLP P. 46), $\{v_n\}_{n=0}^{\infty} \in S$ is a Cauchy sequence. Since S is a complete metric space, every Cauchy sequence in S converges to an element in S (See SLP P. 47), we have $v = \lim_{n\to\infty} v_n \in S$.

To show $Tv = v$, note that for all n and all $v_0 \in S$, we have

$$\begin{aligned}
\rho(Tv, v) &\leq \rho(Tv, v_n) + \rho(v_n, v) \\
&= \rho(Tv, T^n v_0) + \rho(T^n v_0, v) \\
&\leq \beta\rho(v, T^{n-1}v_0) + \rho(T^n v_0, v) \\
&= \beta\rho(v, v_{n-1}) + \rho(v_n, v)
\end{aligned}$$

When $n \to \infty$, since we already show that $\{v_n\}_{n=0}^{\infty}$ is a Cauchy sequence, $\lim_{n\to\infty} v_n = \lim_{n\to\infty} v_{n-1} = v$. Thus we have

$$\rho(Tv, v) = 0, \text{ i.e., } Tv = v$$

Next we need to show v is unique. We prove by contradiction. Suppose $\hat{v} \neq v$ is another solution. Then we have

$$0 < a = \rho(\hat{v}, v) = \rho(T\hat{v}, Tv) \leq \beta\rho(\hat{v}, v) = \beta a$$

Since $\beta \in (0, 1)$, the inequality cannot hold. Contradiction. This concludes part (a).

To prove part (b), observe that for any $n \geq 1$, we have

$$\begin{aligned}
\rho(T^n v_0, v) &= \rho(T(T^{n-1}v_0), Tv) \leq \beta\rho(T^{n-1}v_0, v) \\
&= \beta\rho(T(T^{n-2}v_0), Tv) \leq \beta^2\rho(T^{n-2}v_0, v) \\
&= .... \leq \beta^n\rho(T^0 v_0, v) = \beta^n\rho(v_0, v)
\end{aligned}$$

■

Sometimes, we will find the following theorem is helpful.

Theorem 16 *Let (S, ρ) be a complete metric space, and let $T : S \to S$ be a contraction mapping with fixed point $v \in S$. If S' is a closed subset of S and $T(S') \subseteq S$, then $v \in S'$. If in addition $T(S') \subseteq S'' \subseteq S'$, then $v \in S''$.*

We know CMT is very useful for our analysis of FE. But how can we know a mapping is a contraction mapping? In 1965 Blackwell provided sufficient conditions for an operator to be a contraction mapping. It turns out that these conditions can be easily checked in a lot of applications. However, since they are only *sufficient*, failure of these conditions does not imply that the operator is not a contraction. In these cases we just have to look somewhere else. Here is Blackwell's theorem.

Theorem 17 *(Blackwell's sufficient conditions for a contraction) Let $X \subseteq R^l$, and let $B(X)$ be a space of bounded functions $f : X \to R$, with the sup norm. Let $T : B(X) \to B(X)$ be an operator satisfying*

a). (monotonicity) $f, g \in B(X)$ and $f(x) \leq g(x)$, for all $x \in X$, implies $(Tf)(x) \leq (Tg)(x)$, for all $x \in X$;

b). (discounting) $\exists$ some $\beta \in (0,1)$ such that

$$[T(f+a)](x) \leq (Tf)(x) + \beta a, \text{ all } f \in B(X), a \geq 0, x \in X.$$

[Here $(f+a)(x)$ is a function defined by $(f+a)(x) = f(x) + a$.] Then T is a contraction with modulus β.

Proof. In notation, if $f(x) \leq g(x), \forall x$, then we write $f \leq g$. Now for any $f, g \in B(X)$, it is easy to show that (since we have sup norm) $f \leq g + \|f - g\|$. Thus we have

$$Tf \leq T(g + \|f-g\|) \leq Tg + \beta \|f - g\| \tag{2.1}$$

Notice the first inequality uses property (a), the second comes from part (b). (Note that norm $\|\cdot\|$ is a real number and $\|\cdot\| \geq 0$)

Reversing the roles of f and g, follow the same argument above, we have

$$Tg \leq T(f + \|f-g\|) \leq Tf + \beta \|f - g\| \tag{2.2}$$

Combining (2.1) and (2.2), we have

$$\begin{aligned} Tf - Tg &\leq \beta \|f-g\| \\ Tg - Tf &\leq \beta \|f-g\| \end{aligned}$$

Therefore

$$\sup |Tf - Tg| = \|Tf - Tg\| \leq \beta \|f - g\|$$

■

Now it is time to go back to our FE to check if it is a contarction mapping by using Blackwell's sufficient conditions. Again we define the following operator T

$$(Tv)(k) = \max_{0 \leq y \leq f(k)} [U(f(k) - y) + \beta v(y)]$$

First let's check T is a operator mapping to itself. Since $v \in B(X)$, v is bounded. Since we also assume U is bounded, it is obviously Tv is also bounded, hence $Tv \in B(X)$.

Second, let's check the monotonicity. Suppose $v, w \in B(X)$ and $v \leq w$. Denote $g_v(k)$ an optimal policy function corresponding to v. We will have $\forall k$

$$\begin{aligned} Tv(k) &= U(f(k) - g_v(k)) + \beta v(g_v(k)) \\ &\leq U(f(k) - g_v(k)) + \beta w(g_v(k)) \\ &\leq U(f(k) - g_w(k)) + \beta w(g_w(k)) \\ &= \max_{0 \leq k' \leq f(k)} [U(f(k) - k') + \beta w(k')] \\ &= Tw(k) \end{aligned}$$

Finally, check discounting.

$$\begin{aligned} T(v+a)(k) &= T(v(k) + a) \\ &= \max_{0 \leq y \leq f(k)} [U(f(k) - y) + \beta(v(y) + a)] \\ &= \max_{0 \leq y \leq f(k)} [U(f(k) - y) + \beta(v(y))] + \beta a \\ &= (Tv)(k) + \beta a \end{aligned}$$

Therefore, the neoclassical growth model with bounded utility satisfies the sufficient conditions for a contraction and there is a unique fixed point to the functional equation that can be computed from any starting guess v_0 by repeated application of the T-operator.

2.3 The Theorem of the Maximum

Notice that the Blackwell Sufficient Conditions apply to a space of bounded functions. Since a lot of economic applications focus on continuous function, we would like to study the operator on a space of continuous functions f. And we would like to know the properties of resulting function Tf and the corresponding optimal policy functions. The Thorem of the Maximum (TOM) will help us to establish that.

The operator T is defined by

$$\begin{aligned} (Tv)(x) &= \sup[F(x,y) + \beta v(y)] \\ & \quad s.t.\ y \text{ is feasible, given } x \end{aligned}$$

Obviously, T is an operator to map the space of bounded continuous functions $C(X)$ to itself. We use *correspondence* Γ to describle the feasible set. $\Gamma : X \to Y$ maps each element $x \in X$ into a subset $\Gamma(x)$ of Y. Hence the image of the point x under Γ may consist of more than one point (in contrast to a function, in which the image of x always consists of a singleton).

Define $f(x,y) = F(x,y) + \beta v(y)$, the FE above can be rewritten as a problem with the form $\sup_{y \in \Gamma(x)} f(x,y)$. If for each x, $f(x, \cdot)$ is continuous in y and the set $\Gamma(x)$ is nonempty and compact, according to Weierstrass Theorem (or

Maximum Value Theorem), then for each x the maximum is attained. In this case, the function

$$h(x) = \max_{y \in \Gamma(x)} f(x, y) \tag{2.3}$$

is well-defined and gives the value of the maximization problem. We also define

$$G(x) = \{y \in \Gamma(x) : f(x, y) = h(x)\} \tag{2.4}$$

as the set of y values that attain the maximum.

Since we will talk about continuity, let's make some definitions about continuity first.

Definition 18 *A compact-valued correspondence $\Gamma : X \to Y$ is* ***upper-hemicontinuous*** *(u.h.c.) at a point x if $\Gamma(x) \neq \emptyset$ and if for every sequence $\{x_n\}$ in X converging to some $x \in X$ and every sequence $\{y_n\}$ in Y such that $y_n \in \Gamma(x_n)$ for all n, there exists a convergent subsequence of $\{y_n\}$ that converges to some $y \in \Gamma(x)$. A correspondence is upper-hemicontinuous if it is upper-hemicontinuous at all $x \in X$.*

Definition 19 *A correspondence $\Gamma : X \to Y$ is* ***lower-hemicontinuous*** *(l.h.c.) at a point x if $\Gamma(x) \neq \emptyset$ and if for every $y \in \Gamma(x)$ and every sequence $\{x_n\} \to x$, there exists $N \geq 1$ and a sequence $\{y_n\}_{n=N}^{\infty}$ such that $y_n \to y$ and $y_n \in \Gamma(x_n)$, all $n \geq N$. A correspondence is lower-hemicontinuous if it is lower-hemicontinuous at all $x \in X$.*

Definition 20 *A correspondence $\Gamma : X \to Y$ is continuous if it is both upper-hemicontinuous and lower-hemicontinuous.*

[Insert Figure 3.2 in SLP here]

When trying to establish properties of a correspondence $\Gamma : X \to Y$, it is sometimes useful to deal with its graph

$$A = \{(x, y) \in X \times Y : y \in \Gamma(x)\}.$$

We are now ready to answer the questions: Under what conditions do the function $h(x)$ defined by the maximization problem in (2.3) and the associated set of maximizing y values $G(x)$ defined in (2.4) vary continuously with x? We have the following theorem.

Theorem 21 *(Theorem of Maximum) Let $X \subseteq R^l$ and $Y \subseteq R^m$, let $f : X \times Y \to R$ be a continuous function, and let $\Gamma : X \to Y$ be a compact-valued and continuous correspondence. Then the function $h : X \to R$ defined in (2.3) is continuous, and the correspondence $G : X \to Y$ defined in (2.4) is nonempty, compact-valued, and u.h.c.*

Proof. First we show G is nonempty and compact. Fix $x \in X$. Since by assumption $\Gamma(x)$ is nonempty and compact, and $f(x,\cdot)$ is continuous, hence according to Maximum Value theorem, the maximum in (2.3) is attained, and the set $G(x)$ of maximizers is nonempty. Moreover, since $G(x) \subseteq \Gamma(x)$ and $\Gamma(x)$ is a compact (closed and bounded), obviously $G(x)$ is bounded. Suppose $y_n \to y$, and $y_n \in G(x)$, all n. Since $G(x) \subseteq \Gamma(x)$, hence $y_n \in \Gamma(x)$. and $\Gamma(x)$ is closed, we have $y \in \Gamma(x)$. (A set is closed iff it contains all of the limiting points of convergent sequences in this set. See Rosenlicht P. 47). Also, since $y_n \in G(x)$, we have $h(x) = f(x, y_n)$, all n, and since f is continuous, by definition of continuity, $h(x) = f(x,y)$. Therefore, again by the definition of G, $y \in G(x)$. So $G(x)$ is closed. Since x is arbitrary, thus $G(x)$ is nonempty and compact, for each x.

Next, we will show that G is u.h.c.. Fix x, and let $\{x_n\}$ be any sequence converging to x. Choose $y_n \in G(x_n)$, all n. Since Γ is continuous, hence it is also u.h.c. Therefore, $\exists$ a subsequence $\{y_{n_k}\}$ converging to $y \in \Gamma(x)$. Let $z \in \Gamma(x)$. Since Γ is also l.h.c., $\exists$ a sequence $z_{n_k} \to z$, with $z_{n_k} \in \Gamma(x_{n_k})$, all k. Since $y_{n_k} \in G(x_{n_k})$, by the definition of G, we have $f(x_{n_k}, y_{n_k}) \geq f(x_{n_k}, z_{n_k})$, all k. Again, the continuity of f then guarantee that $f(x,y) \geq f(x,z)$. Since this holds for any $z \in \Gamma(x)$, by the definition of G, it follows that $y \in G(x)$. Hence G is u.h.c.

Finally, we will show that h is continuous. Fix x, and let $\{x_n\}$ be any sequence converging to x. Choose $y_n \in G(x_n)$, all n. Let $\bar{h} = \limsup h(x_n)$ and $\underline{h}= \liminf h(x_n)$. Then there exists a subsequence $\{x_{n_k}\}$ such that $\bar{h} = \lim f(x_{n_k}, y_{n_k})$. But since G is u.h.c., there exists a subsequence of $\{y_{n_k}\}$, call it $\{y'_j\}$, converging to $y \in G(x)$. Hence $\bar{h} = \lim f(x_j, y'_j) = f(x,y) = h(x)$ (the second equality comes from the fact that f is continuous. The third equality comes from the fact that $y \in G(x)$). An analogous argument establishes that $h(x) =\underline{h}$. *Hence* $\{h(x_n)\} \to h(x)$. i.e., $h(x)$ is continuous at x. Since x is arbitrary, h is continuous. ■

Proposition 22 *Suppose that in addition to the hypothesis of the Theorem of the Maximum the correspondence* Γ *is convex-valued and the function* f *is strictly concave in* y. *Then* G *is a single-valued continuous function.*

Chapter 3

Dynamic Programming Under Certainty

From the previous two chapters, we know that we are interested in the sequential problem (SP) with the form

$$\begin{aligned} & \sup_{\{x_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1}) \\ & s.t. \\ x_{t+1} & \in \Gamma(x_t), t = 0, 1, 2, ... \\ x_0 & \in X \text{ given.} \end{aligned}$$

Corresponding to any such problem, we have a functional equation (FE) of the form

$$v(x) = \sup_{y \in \Gamma(x)} [F(x, y) + \beta v(y)],\ \forall x \in X$$

In this chapter we will establish the relationship between solutions to these two problems and develop methods for analyzing the latter.

3.1 The Principle of Optimality

In this section, we study the realtionship between solutions to SP and FE. The general idea is that the solution v to FE, evaluated at x_0, $v(x_0)$, gives the value of the supremum in SP when the initial state is x_0 and that a sequence $\{x_{t+1}\}_{t=0}^{\infty}$ attains the supremum in SP iff it satisfies

$$v(x_t) = F(x_t, x_{t+1}) + \beta v(x_{t+1}), t = 0, 1, 2, ... \tag{3.1}$$

This idea is called the Principle of Optimality by Richard Bellman.

Let's establish some notation first. Let X be the set of possible values for the state variable x. Let $\Gamma : X \to X$ be the correspondence describing the feasibility constraints of next period state variable y, given that today's state is x. Denote A be the graph of Γ:

$$A = \{(x, y) \in X \times X : y \in \Gamma(x)\}$$

$F : A \to R$ is the one-period return function. $\beta \geq 0$ is the discount factor. (X, F, Γ, β) is our "givens" in this problem.

We call any sequence $\{x_t\}_{t=0}^{\infty}$ in X a *plan*. Given $x_0 \in X$, denote the set of feasible plan starting from x_0

$$\Pi(x_0) = \{\{x_t\}_{t=0}^{\infty} : x_{t+1} \in \Gamma(x_t), t = 0, 1, 2, ...\}$$

That is, $\Pi(x_0)$ is the set of all sequences $\{x_t\}$ satisfying the constraints in SP. Let $\breve{x} = (x_0, x_1, x_2, ...)$ denote a typical element of $\Pi(x_0)$.

We have following assumptions.

Assumption 1: $\Gamma(x)$ is nonempty for all $x \in X$.

Assumption 2: For all $x_0 \in X$ and $\breve{x} \in \Pi(x_0)$,

$$\lim_{n\to\infty} \sum_{t=0}^{n} \beta^t F(x_t, x_{t+1})$$

exists (although it might be $+\infty$ or $-\infty$).

Obviously, if function F is bounded and $0 < \beta < 1$, Assumption 2 is satisfied. (If $0 \leq |F(x, y)| < B < \infty$ and $0 < \beta < 1$, then $\lim_{n\to\infty} \sum_{t=0}^{n} \beta^t F(x_t, x_{t+1}) \leq \lim_{n\to\infty} \sum_{t=0}^{n} \beta^t B = \frac{B}{1-\beta}$.) Thus a sufficient condition for Assumptions 1-2 is that F is bounded above or below and $0 < \beta < 1$. Another sufficient condition is that $\forall x_0 \in X$ and $\breve{x} \in \Pi(x_0)$, $\exists\ \theta \in (0, \beta^{-1})$ and $0 < c < \infty$ such that

$$F(x_t, x_{t+1}) \leq c\theta^t, \text{ all } t$$

Hence we do not need restriction on F. As long as the returns from the sequence do not grow too fast (at a rate higher than $\frac{1}{\beta}$), we are still fine.

Define $u_n : \Pi(x_0) \to R$ by

$$u_n(\breve{x}) = \sum_{t=0}^{n} \beta^t F(x_t, x_{t+1}).$$

Then $u_n(\breve{x})$ is the partial sum of the discounted returns to period 0 through n from the feasible plan $\breve{x}$. Under Assumption 2, we can also define $u : \Pi(x_0) \to \bar{R}$ ($\bar{R} = R \cup \{+\infty, -\infty\}$ is the set of extended real numbers)

$$u(\breve{x}) = \lim_{n\to\infty} u_n(\breve{x}).$$

If Assumptions 1 and 2 both hold, then the set of feasible plans $\Pi(x_0)$ is nonempty and the objective funtion in SP is well defined for every plan $\breve{x} \in \Pi(x_0)$. We can then define the supremum function $v^* : X \to \bar{R}$ by

$$v^*(x_0) = \sup_{\breve{x}\in\Pi(x_0)} u(\breve{x}).$$

Therefore, $v^*(x_0)$ is the supremum in SP given th initial state is x_0. Note that by definition v^* is the unique function satisfying the following three conditions:

a. if $|v^*(x_0)| < \infty$, then

$$v^*(x_0) \geq u(\breve{x}), \forall \breve{x} \in \Pi(x_0); \tag{3.2}$$

This shows v^* is a upper bound of u. And for any $\varepsilon > 0$,

$$v^*(x_0) \leq u(\breve{x}) + \varepsilon, \text{ for } \textit{some } \breve{x} \in \Pi(x_0); \tag{3.3}$$

This shows that v^* is the least upper bound.

b. if $v^*(x_0) = +\infty$, then there exists a sequence $\{\breve{x}^k\}$ in $\Pi(x_0)$ such that $\lim_{k\to\infty} u(\breve{x}^k) = +\infty$; and

c. if $v^*(x_0) = -\infty$, then $u(\breve{x}) = -\infty$, $\forall \breve{x} \in \Pi(x_0)$.

Note that by construction, whenever v^* exists, it is unique (since the supremum of a set of real numbers is always unique).

Accordingly, we say that v^* satisfies FE if the following three conditions hold:

a. if $|v^*(x_0)| < \infty$, then

$$v^*(x_0) \geq F(x_0, y) + \beta v^*(y), \text{ for all } y \in \Gamma(x_0); \tag{3.4}$$

and for any $\varepsilon > 0$,

$$v^*(x_0) \leq F(x_0, y) + \beta v^*(y) + \varepsilon, \text{ for } \textit{some } y \in \Gamma(x_0); \tag{3.5}$$

b. if $v^*(x_0) = +\infty$, then there exists a sequence $\{y^k\}$ in $\Gamma(x_0)$ such that

$$\lim_{k\to\infty} [F(x_0, y^k) + \beta v^*(y^k)] = +\infty; \tag{3.6}$$

c. if $v^*(x_0) = -\infty$, then

$$F(x_0, y) + \beta v^*(y) = -\infty, \text{ for all } y \in \Gamma(x_0). \tag{3.7}$$

Note that FE may not have a unique solution.

We will establish a preliminary result by proving following lemma.

Lemma 23 *Let X, F, Γ, and β satisfy Assumption 2. Then for any $x_0 \in X$ and any $(x_0, x_1, x_2, ...) = \breve{x} \in \Pi(x_0)$, we have*

$$u(\breve{x}) = F(x_0, x_1) + \beta u(\breve{x}'),$$

where $\breve{x}' = (x_1, x_2, ...)$.

Proof. Under Assumption 2, by definition, for any $x_0 \in X$ and any $\breve{x} \in \Pi(x_0)$,

$$\begin{aligned} u(\breve{x}) &= \lim_{n\to\infty} \sum_{t=0}^{n} \beta^t F(x_t, x_{t+1}) \\ &= F(x_0, x_1) + \beta \lim_{n\to\infty} \sum_{t=0}^{n} \beta^t F(x_{t+1}, x_{t+2}) \\ &= F(x_0, x_1) + \beta u(\breve{x}'). \end{aligned}$$

■

Now it is the time to show that the supremum function v^* for the SP problem satisfies FE.

Theorem 24 *Suppose X, F, Γ, and β satisfy Assumptions 1-2. Then the function v^* satisfies FE.*

Proof. If $\beta = 0$, $v^* = F(x_0, x_1)$ in both SP and FE. The result is trival. Hence let's suppose $\beta > 0$, and choose $x_0 \in X$.

Suppose $v^*(x_0)$ is finite. Then (3.2) and (3.3) hold. We want to show that (3.4) and (3.5) also hold for $v^*(x_0)$. To establish (3.4), let $x_1 \in \Gamma(x_0)$ and $\varepsilon > 0$ is given. Then by (3.3), $\exists\ \breve{x}' = (x_1, x_2, ...) \in \Pi(x_1)$ such that $u(\breve{x}') \geq v^*(x_1) - \varepsilon$. Note that $\breve{x} = (x_0, x_1, x_2, ...) \in \Pi(x_0)$. Hence it follows from (3.2) and the lemma above that

$$v^*(x_0) \geq u(\breve{x}) = F(x_0, x_1) + \beta u(\breve{x}') \geq F(x_0, x_1) + \beta v^*(x_1) - \varepsilon.$$

Since $\varepsilon > 0$ can be arbitrarily small, (3.4) follows.

To establish (3.5), choose $x_0 \in X$ and $\varepsilon > 0$. From (3.3) and the lemma above, it follows that we can choose one plan $\breve{x} = (x_0, x_1, x_2, ...) \in \Pi(x_0)$, so that

$$v^*(x_0) \leq u(\breve{x}) + \varepsilon = F(x_0, x_1) + \beta u(\breve{x}') + \varepsilon,$$

where $\breve{x}' = (x_1, x_2, ...)$. It then follows (3.2) that

$$v^*(x_0) \leq F(x_0, x_1) + \beta v^*(x_1) + \varepsilon.$$

Since $x_1 \in \Pi(x_0)$, this establishs (3.5).

If $v^*(x_0) = +\infty$, then there exists a sequence $\{\breve{x}^k\}$ in $\Pi(x_0)$ such that $\lim_{k\to\infty} u(\breve{x}^k) = +\infty$. Since $x_1^k \in \Gamma(x_0)$, all k, and by (3.2) and the lemma above (notice that here $v^*(x_1^k)$ is still finite)

$$u(\breve{x}^k) = F(x_0, x_1^k) + \beta u(\breve{x}'^k) \leq F(x_0, x_1^k) + \beta v^*(x_1^k), \text{ all } k$$

Take a limit on both side of above inequality

$$\lim_{k\to\infty} u(\breve{x}^k) = \lim_{k\to\infty} [F(x_0, x_1^k) + \beta u(\breve{x}'^k)] \leq \lim_{k\to\infty} [F(x_0, x_1^k) + \beta v^*(x_1^k)]$$

Since we know

$$\lim_{k\to\infty} u(\breve{x}^k) = +\infty$$

Therefore, $\lim_{k\to\infty}[F(x_0, x_1^k) + \beta v^*(x_1^k)] = +\infty$ with the sequence $\{y^k = x_1^k\}$.

If $v^*(x_0) = -\infty$, then by lemma above

$$u(\breve{x}) = F(x_0, x_1) + \beta u(\breve{x}') = -\infty, \forall \breve{x} \in \Pi(x_0)$$

Since F is a real-valued function (it does not take on the value $+\infty$ or $-\infty$), it follows that

$$u(\breve{x}') = -\infty, \text{ all } x_1 \in \Gamma(x_0), \text{ all } \breve{x}' \in \Gamma(x_1).$$

Then by (3.3)

$$v^*(x_1) \leq u(\breve{x}') + \varepsilon$$

for any $\varepsilon > 0$, which implies

$$v^*(x_1) = -\infty, \text{ all } x_1 \in \Gamma(x_0).$$

Let $y = x_1$, since F is a real-valued function and $\beta > 0$, (3.7) follows immediately. ■

The next theorem provides a partial converse to the theorem above. It shows that v^* is the only solution to the functional equation that satisfies a certain boundedness condition.

Theorem 25 *Let X, F, Γ, and β satisfy Assumption 1-2. If the function v is a solution to FE and satisfies*

$$\lim_{n\to\infty} \beta^n v(x_n) = 0, \text{ all } (x_0, x_1, ...) \in \Pi(x_0), \text{ all } x_0 \in X, \tag{3.8}$$

then $v = v^*$.

Proof. If $v(x_0)$ is finite, then (3.4) and (3.5) hold. We want to show given this, (3.2) and (3.3) also hold.

Notice that (3.4) implies that for all $\breve{x} \in \Pi(x_0)$,

$$\begin{aligned} v(x_0) &\geq F(x_0, x_1) + \beta v(x_1) \\ &\geq F(x_0, x_1) + \beta F(x_1, x_2) + \beta^2 v(x_2) \\ &\quad .... \\ &\geq u_n(\breve{x}) + \beta^{n+1} v(x_{n+1}), \; n = 0, 1, 2, ... \end{aligned}$$

Take the limit on both side as $n \to \infty$ and using the definition of $u(\breve{x})$ and boundary condition (3.8), we find (3.2) holds.

Next, we show (3.3) also holds. Fix $\varepsilon > 0$ and choose $\{\delta_t\}_{t=1}^{\infty}$ in R_+ such that $\sum_{t=1}^{\infty} \beta^{t-1}\delta_t \leq \varepsilon$. Since (3.5) holds, we can choose $x_1 \in \Gamma(x_0)$, $x_2 \in \Gamma(x_1)$, ... so that

$$v(x_t) \leq F(x_t, x_{t+1}) + \beta v(x_{t+1}) + \delta_{t+1}, \; t = 0, 1, ...$$

Then $\breve{x} = (x_0, x_1, x_2, ...) \in \Pi(x_0)$, and by repeatedly applying the inequality above, we have

$$\begin{aligned} v(x_0) &\leq F(x_0, x_1) + \beta v(x_1) + \delta_1 \\ &\leq F(x_0, x_1) + \beta(F(x_1, x_2) + \beta v(x_2) + \delta_2) + \delta_1 \\ &\quad .... \\ &\leq \sum_{t=0}^{n} \beta^t F(x_t, x_{t+1}) + \beta^{n+1} v(x_{n+1}) + \sum_{t=0}^{n} \beta^t \delta_{t+1} \\ &\leq u_n(\breve{x}) + \beta^{n+1} v(x_{n+1}) + \varepsilon \end{aligned}$$

Now thake limit when $n \to \infty$, (3.8) implies $v^*(x_0) \leq u(\breve{x}) + \varepsilon$. (3.3) holds.

If $v(x_0) = +\infty$, choose $n \geq 0$ and $(x_0, x_1, x_2, ..., x_n)$ such that $x_t \in \Gamma(x_{t-1})$ and $v(x_t) = +\infty$ for $t = 0, 1, ..., n$ and $v(x_{n+1}) < +\infty$ for all $x_n \in \Gamma(x_{n+1})$. Clearly (3.8) implies that n is finite. Fix any $A > 0$. Since $v(x_n) = +\infty$, (3.6) implies that we can choose $x_{n+1}^A \in \Gamma(x_n)$ such that

$$F(x_n, x_{n+1}^A) + \beta v(x_{n+1}^A) \geq \beta^{-n}[A + 1 - \sum_{t=0}^{n-1} \beta^t F(x_t, x_{t+1})].$$

Since $v(x_{n+1}) < +\infty$, we can choose $\breve{x}_{n+1}^A \in \Pi(x_{n+1}^A)$ such that by (3.3) (remember we already proved that it holds)

$$u(\breve{x}_{n+1}^A) \geq v(x_{n+1}^A) - \beta^{-(n+1)}$$

Then $\breve{x}^A = (x_0, x_1, ..., x_n, \breve{x}_{n+1}^A) \in \Pi(x_0)$, and

$$\begin{aligned} u(\breve{x}^A) &= \sum_{t=0}^{n-1} \beta^t F(x_t, x_{t+1}) + \beta^n F(x_n, x_{n+1}^A) + \beta^{n+1} u(\breve{x}_{n+1}^A) \text{ (using the Lemma)} \\ &\geq \sum_{t=0}^{n-1} \beta^t F(x_t, x_{t+1}) + \beta^n F(x_n, x_{n+1}^A) + \beta^{n+1}(v(x_{n+1}^A) - \beta^{-(n+1)}) \\ &= \sum_{t=0}^{n-1} \beta^t F(x_t, x_{t+1}) + \beta^n F(x_n, x_{n+1}^A) + \beta^{n+1} v(x_{n+1}^A) - 1 \\ &= \sum_{t=0}^{n-1} \beta^t F(x_t, x_{t+1}) + \beta^n (F(x_n, x_{n+1}^A) + \beta v(x_{n+1}^A)) - 1 \\ &\geq \sum_{t=0}^{n-1} \beta^t F(x_t, x_{t+1}) + A + 1 - \sum_{t=0}^{n-1} \beta^t F(x_t, x_{t+1}) - 1 \\ &= A. \end{aligned}$$

Since $A > 0$ is arbitrary, we find a sequence $\breve{x}^A \in \Pi(x_0)$ such that $\lim_{A \to \infty} u(\breve{x}^A) = +\infty$.

If (3.8) holds, then (3.7) implies that v has to be bounded. Therefore, it cannot take the value $-\infty$. ■

This theorem immediately implies that FE can only have at most one solution satisfying (3.8). The following example is a case where FE has an extraneous solution in addition to v^*.

Consider the following consumption problem of an infinitely lived HH. The household has initial wealth $x_0 \in X = R$. He can borrow or lend at a gross interest rate $1 + r = \frac{1}{\beta}$. So the price of a bond that pays off one unit of consumption next period is $q = \frac{1}{1+r} = \beta$. There are no borrowing constraints, so the sequential budget constraint is

$$c_t + \beta x_{t+1} \leq x_t$$

Now the social planner's problem is

$$\begin{aligned} & \max \sum_{t=0}^{\infty} \beta^t c_t \\ & s.t. \\ 0 \leq \; & c_t \leq x_t - \beta x_{t+1} \\ & x_0 \text{ given.} \end{aligned}$$

Since there are no borrowing constraints, the consumer can assure herself infinite utility by just borrowing an infinite amount in period 0 and then rolling over the debt by even borrowing more in the future. i.e., $x_{t+1} = -\infty \Rightarrow c_t = +\infty$. Such a strategy is called a *Ponzi scheme*. Consumption is unbounded, the sup function is obviously $v^*(x) = +\infty$, for all x. Now consider the recursive formulation of this problem. The return function is $F(x, y) = x - \beta y$, and the correspondence describing the feasible set is $\Gamma(x) = (-\infty, \beta^{-1}x]$. The FE is

$$v(x) = \sup_{y \leq \beta^{-1}x} [x - \beta y + \beta v(y)].$$

Obviously the function $v(x) = +\infty$ satisfies this functional equation, SP⇒FE. But $v(x) = x$, $\forall x$ is feasible, also satisfies FE (Check!). However, since the sequence $\{x_t = \beta^{-t}x_0\}_{t=0}^{\infty}$ is in $\Gamma(x_0)$, while

$$\lim_{n\to\infty} \beta^n v(x_n) = \lim_{n\to\infty} \beta^n x_n = x_0 > 0,$$

(3.8) does not hold here. Hence $v(x) = x$, although is a solution to FE, not a solution to SP. FE⇏SP.

Remark 26 *It is actually easy to understand that value function $v(x) = x$ with feasible allocation $\{x_t = \beta^{-t}x_0\}_{t=0}^{\infty}$ cannot be the solution to SP in terms of economic sense. Allocation $\{x_t = \beta^{-t}x_0\}_{t=0}^{\infty}$ implies $\{c_t = 0\}_{t=0}^{\infty}$, which in turn will yield zero life-time utility.*

Now we want to establish a similar equivalence between SP and FE with respect to the optimal policies/plans. We call a feasible plan $\breve{x} \in \Pi(x_0)$ an *optimal plan from* x_0 if it attains the supremum in SP, that is, if $u(\breve{x}) = v^*(x_0)$. The next two theorems deal with the relationship between optimal plans and those that satisfy the policy equation (3.1) for $v = v^*$. The following theorem shows that optimal plans satisfy (3.1).

Theorem 27 *Let X, F, Γ, and β satisfy Assumption 1-2. Let $\breve{x}^* \in \Pi(x_0)$ be a feasible plan that attains the supremum in SP for initial state x_0. Then*

$$v^*(x_t^*) = F(x_t^*, x_{t+1}^*) + \beta v(x_{t+1}^*), t = 0, 1, 2, ... \tag{3.9}$$

Proof. Since $\breve{x}^*$ attains the supremum, we have

$$\begin{aligned} v^*(x_0) &= u(\breve{x}^*) \\ &= F(x_0, x_1^*) + \beta u(\breve{x}^{*\prime}) \\ &\geq u(\breve{x}) \\ &= F(x_0, x_1) + \beta u(\breve{x}'), \text{ all } \breve{x} \in \Pi(x_0). \end{aligned} \tag{3.10}$$

The first equality comes from the definition of v^*. The second equality comes from Lemma 23. The third inequality is from equation (3.2). The fourth equality again is using Lemma 23. In particular, the inequality holds for all plans with $x_1 = x_1^*$. Since $(x_1^*, x_2, x_3, ...) \in \Pi(x_1^*)$ implies that $(x_0, x_1^*, x_2, x_3, ...) \in \Pi(x_0)$, it follows that

$$u(\breve{x}^{*\prime}) \geq u(\breve{x}'), \text{ all } \breve{x} \in \Pi(x_1^*).$$

Therefore $u(\breve{x}^{*\prime})$ is a supremum. By definition, $u(\breve{x}^{*\prime}) = v(x_1^*)$. Substituting this into (3.10) gives (3.9) for $t = 0$. Continuing by induction establishes (3.9) for all t. ■

The next theorem provides a partial converse to Theorem 27. It says that any sequence satisfying (3.9) and a boundedness condition is an optimal plan.

Theorem 28 *Let X, F, Γ, and β satisfy Assumption 1-2. Let $\breve{x}^* \in \Pi(x_0)$ be a feasible plan from x_0 satisfying (3.9), and with*

$$\lim_{t\to\infty} \sup \beta^t v^*(x_t^*) \leq 0. \tag{3.11}$$

Then $\breve{x}^*$ attains the supremum in SP for initial state x_0.

Proof. Suppose that $\breve{x}^* \in \Pi(x_0)$ satisfies (3.9) and (3.11). Then it follows by an induction on (3.9) that

$$\begin{aligned} v^*(x_0) &= F(x_0, x_1^*) + \beta v^*(x_1^*) \\ &= F(x_0, x_1^*) + \beta[F(x_1^*, x_2^*) + \beta v^*(x_2^*)] \\ &= F(x_0, x_1^*) + \beta F(x_1^*, x_2^*) + \beta^2[F(x_2^*, x_3^*) + \beta v^*(x_3^*)] \\ &\quad ...... \\ &= \sum_{t=0}^{n} \beta^t F(x_t^*, x_{t+1}^*) + \beta^{n+1} v^*(x_{n+1}^*), \ n = 1, 2, ... \end{aligned}$$

Then take limit on both sides and using (3.11), we find that $v^*(x_0) \leq u(\breve{x}^*)$. Since $\breve{x}^* \in \Pi(x_0)$, by definition of $v^*(x_0) = \sup_{\breve{x}\in\Pi(x_0)} u(\breve{x})$, we also have $v^*(x_0) \geq u(\breve{x}^*)$. This establishes the result. ■

The consumption example (or "Cake-eating problem") used after Theorem 25 can be modified to illustrate why the boundary condition (3.11) is needed. Now we impose a borrowing constraint of zero (you can only save, not borrow).

The sequential problem is

$$\begin{aligned} & \max_{\{x_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t (x_t - \beta x_{t+1}) \\ & s.t. \\ 0 \ \leq \ & x_{t+1} \leq \beta^{-1} x_t, \ t = 0, 1, 2, ... \\ & x_0 \text{ given} \end{aligned}$$

Obviously the supremum function is $v^*(x_0) = x_0$, all $x_0 \geq 0$. It is also clear that this v^* satisfies FE

$$v^*(x) = \max_{0 \leq y \leq \beta^{-1} x} [x - \beta y + \beta v^*(y)], \text{ all } x$$

as Theorem 24 implies.

Now consider plans that attain the optimum. Obviously, $(x_0, 0, 0, 0, ...)$ (only consume in period 0) is a feasible plan. It is clear that this plan satisfies (3.9). Since it is also clear that $\lim_{t \to \infty} \sup \beta^t v^*(x_t^*) = 0$ for this plan, as Theorem 28 states, this plan attains the supremum in SP. Similarly, plans like $(x_0, \beta^{-1} x_0, 0, 0, ...)$ (only consume in period 1), $(x_0, \beta^{-1} x_0, \beta^{-2} x_0, 0, 0, ...)$ (only consume in period 2) and all convex combinations of them attain the supremum in SP. But feasible plan $\{x_t = \beta^{-t} x_0, t = 0, 1, 2, ...\}$ (never consume, save forever) does not attain the supremum because

$$\lim_{t \to \infty} \sup \beta^t v^*(x_t^*) = \lim_{t \to \infty} \sup \beta^t (\beta^{-t} x_0) = x_0 > 0.$$

(3.11) is violated hence Theorem 28 does not apply.

3.2 Dynamic Programming with Bounded Returns

In this section we study functional equations of the form

$$v(x) = \max_{y \in \Gamma(x)} [F(x, y) + \beta v(y)] \tag{3.12}$$

under the assumption that function F is bounded and the discount factor β is strictly less than one.

Let's make some further assumptions.

Assumption 3: X is a convex subset of R^l, and the correspondence $\Gamma : X \to X$ is noempty, compact-valued, and continuous.

Assumption 4: The function $F : A \to R$ is bounded and continuous, and $0 < \beta < 1$.

Recall that A is the graph of Γ

$$A = \{(x, y) \in X \times X : y \in \Gamma(x)\}.$$

Clearly, under Assumption 3-4, Assumption 1-2 hold. Therefore, the theorems in the previous section apply. We define the policy correspondence $G : X \to X$ by

$$G(x) = \{y \in \Gamma(x) : v(x) = F(x, y) + \beta v(y)\}. \tag{3.13}$$

If a sequence $\breve{x} = (x_0, x_1, x_2, ...)$ satisfies $x_{t+1} \in \Gamma(x_t)$, $t = 0, 1, 2, ...$, we will say that $\breve{x}$ is *generated from* x_0 *by* G. So Theorem 27 and 28 imply that for any $x_0 \in X$, a sequence $\{x_t^*\}$ attains the supremum in SP iff it is generated by G. (Thorem 25 establish the "onlt if" part. Theorem 26 together with the facts that v^* is bounded and $0 < \beta < 1$ guarantee that $\lim_{t\to\infty} \sup \beta^t v^*(x_t^*) = 0$, hence proves the "if" part.)

Define the operator T on the space of bounded continuous function $C(X)$ by

$$(Tf)(x) = \max_{y\in\Gamma(x)} [F(x, y) + \beta f(y)] \tag{3.14}$$

Therefore, functional equation (3.12) becomes $v = Tv$, i.e., v is a fixed point of the functional equation.

The following theorem establishes that $T : C(X) \to C(X)$, that T has a unique fixed point in $C(X)$, and that the policy correspondence defined in (3.13) is nonempty and u.h.c.

Theorem 29 *Let X, F, Γ, and β satisfy Assumption 3-4, and let $C(X)$ be the space of bounded continuous functions $f : X \to R$, with the sup norm. Then the operator T maps $C(X)$ into itself, $T : C(X) \to C(X)$; T has a unique fixed point $v \in C(X)$; and for all $v_0 \in C(X)$,*

$$\|T^n v_0 - v\| \leq \beta^n \|v_n - v\|, n = 0, 1, 2, ...$$

Moreover, given v, the optimal policy correspondence $G : X \to X$ defined by (3.13) is compact-valued and u.h.c.

Proof. Under Assumptions 3-4, for each $f \in C(X)$, and $x \in X$, the problem in (3.14) is to maximize the continuous function $[F(x, \cdot) + \beta f(\cdot)]$ over the compact set $\Gamma(x)$. Hence according to maximum value theorem, the maximum is attained. Since both F and f are bounded, clearly Tf is bounded. And since F and f are continuous, and Γ is compacted-valued and continuous, it follows from the Theorem of Maximum (Theorem 21) that Tf is continuous. Therefore, T is a mapping from $C(X)$ to itself.

Then it is easy to show that T satisfies Blackwell's sufficient conditions for a contraction mapping (Theorem 15). According to Theorem 3.1 in SLP, $C(X)$ with a sup norm is a complete normed vector space. Therefore, Contraction Mapping Theorem (Theorem 15) applies, that T has a unique fixed point $v \in C(X)$, ans we have $\|T^n v_0 - v\| \leq \beta^n \|v_n - v\|$. The stated properties of G then follow from the Theorem of Maximum. ■

It follows immediately from Theorem 25 that under the hypothesis of Theorem 29, the unique bounded continuous function v satisfying FE is also the supremum function for SP. Therefore, supremum function in SP is also bounded and continuous. Moreover, it then follows from Theorem 28 and 29 that there exists at least one optimal plan: any plan generated by the correspondence G is optimal.

In order to characterize v and G more sharply, we need further information about F and Γ.

Assumption 5: For each y, $F(\cdot, y)$ is strictly increasing in each of its first l arguments.

Assumption 6: Γ is monotone in the sense that $x \leq x'$ implies $\Gamma(x) \subseteq \Gamma(x')$.

Under these assumptions, we can show that the value function v is strictly increasing.

Theorem 30 *Let X, F, Γ, and β satisfy Assumption 3-6, and let v be the unique solution to FE (3.12). Then v is strictly increasing.*

Proof. Obviously. See SLP P. 80. ■

We have a similar result in spirit if we make assumptions about the curvature of the return function and the convexity of the constraint set.

Assumption 7: F is strictly concave, i.e.,

$$\begin{aligned} F[\theta(x,y) + (1-\theta)(x',y')] &\geq \theta F(x,y) + (1-\theta)F(x',y'), \\ \text{all } (x,y),(x',y') &\in A, \text{ and all } \theta \in (0,1) \end{aligned}$$

and the inequality is strict if $x \neq x'$.

What is the economic meaning of this assumption? (strictly convex preference, love of variety)

Assumption 8: Γ is convex in the sense that for any $\theta \in [0,1]$, and $x, x' \in X$, $y \in \Gamma(x)$ and $y' \in \Gamma(x')$ implies

$$\theta y + (1-\theta)y' \in \Gamma[\theta x + (1-\theta)x']$$

What is the economic meaning of this assumption? (no increasing returns)

Theorem 31 *Let X, F, Γ, and β satisfy Assumption 3-4 and 7-8; let v be the unique solution to FE (3.12); and let policy correspondence G satisfy (3.13). Then v is strictly concave and G is a continuous, single-valued function.*

Proof. Let $C'(X) \subset C(X)$ be the set of bounded, continuous, *weakly concave functions on* X, and let $C''(X) \subset C'(X)$ be the set of *strictly concave* function. Obviously, $C'(X)$ is a closed subset of the complete metric space $C(X)$. According to Theorem 29 and Theorem 16, in order to show $v \in C''(X)$, it is sufficient to show that $T[C'(X)] \subseteq C''(X)$.

To verify this, let $f \in C'(X)$ and let

$$x_0 \neq x_1,\ \theta \in (0,1),\text{ and } x_\theta = \theta x_0 + (1-\theta)x_1.$$

Let $y_i \in \Gamma(x_i)$ attain $(Tf)(x_i)$, for $i = 0, 1$. Then by Assumption 8, we know $y_\theta = \theta y_0 + (1-\theta)y_1 \in \Gamma(x_\theta)$. It follows that

$$\begin{aligned}
(Tf)(x_\theta) &\geq F(x_\theta, y_\theta) + \beta f(y_\theta) \\
&> [\theta F(x_0, y_0) + (1-\theta)F(x_1, y_1)] + \beta[\theta f(y_0) + (1-\theta)f(y_1)] \\
&= \theta[F(x_0, y_0) + \beta f(y_0)] + (1-\theta)[F(x_1, y_1) + \beta f(y_1)] \\
&= \theta(Tf)(x_0) + (1-\theta)(Tf)(x_1).
\end{aligned}$$

Where the first weak inequality comes from the definition of (Tf) (see equation (3.14)) and the fact $y_\theta \in \Gamma(x_\theta)$; the second strict inequality uses the hypothesis that $f \in C'(X)$, i.e., f is weakly concave and Assumption 7, F is strictly concave. and the last equality comes from the way y_0 and y_1 were selected. Since x_0 and x_1 are arbitrary, it follows that Tf is strictly concave, and since f is arbitrary, we have $T[C'(X)] \subseteq C''(X)$.

Now since $T[C'(X)] \subseteq C''(X) \subset C'(X)$, according to Theorem 16, we have $v \in C''(X)$. v is strictly concave. Therefore, objective function $F(x, y) + \beta f(y)$ is strictly concave. By Assumption 8, the constraint set $\Gamma(x)$ is strictly convex, it follows that the maximum in mapping T problem (3.14) is attained at a unique y value. Hence G is a single-valued function. The continuity of G then follows from the fact that it is u.h.c. (See Theorem 29). ∎

Since now, we call this single-valued policy correspondence G *policy function,* we denote it as g. From the Contraction Mapping Theorem, we know that we can obtain the value function v by repeatly iterating FE, eventually $v_n \to v$. Can we also have a similar result for policy function?

Theorem 32 *(Convergence of the policy functions) Let $X, F, \Gamma,$ and β satisfy Assumption 3-4 and 7-8; let v be the unique solution to FE (3.12); and let policy function g satisfies (3.13). Let $C'(X)$ be the set of bounded, continuous, concave functions $f : X \to R$ and let $v_0 \in C'(X)$. Let $\{(v_n, g_n)\}$ be defined by*

$$v_{n+1} = Tv_n, n = 0, 1, 2, ...$$

and

$$g_n(x) = \arg\max_{y \in \Gamma(x)} [F(x, y) + \beta f_n(y)], n = 0, 1, 2, ...$$

Then $g_n \to g$ pointwise. If X is compact, then the convergence is uniform.

This theorem is the theoretical foundation of the "policy function iteration" method to solving dynamic programming.

Proof. Skip. See SLP P. 82. ∎

Finally we state a result about the differentiability of the value function v, the famous envelope theorem (some people call it the Benveniste-Scheinkman theorem).

We begin with the theorem originally proved by Benveniste and Scheinkman.

Theorem 33 *(Benveniste-Scheinkman) Let $X \subseteq R^l$ be a convex set, let $V : X \to R$ be concave, let $x_0 \in intX$, and let D be a neighborhood of x_0. If there* *is a concave, differentiable function $W : D \to R$, with $W(x_0) = V(x_0)$ and with $W(x) \leq V(x)$ for all $x \in D$, then V is differentiable at x_0, and*

$$V_i(x_0) = W_i(x_0), i = 1, 2, ..., l,$$

where V_i is the first-order derivative of V w.r.t. the ith argument.

Proof. Any subgradient p of V at x_0 ($p = V_i(x_0)$) must satisfy

$$p \cdot (x - x_0) \geq V(x) - V(x_0) \geq W(x) - W(x_0), \text{ all } x \in D$$

where the first inequality uses the definition of a subgradient and the fact that V is concave. The second uses the fact that $W(x) \leq V(x)$, with equality at x_0. Since W is differentiable at x_0, p is unique, and any unique function with a unique subgradient at an interior point x_0 is differentiable at x_0. ■

Remark 34 *We call V is the envelope for W in the case above.*

[Insert Figure 4.1 in SLP here.]

Apply this theorem to dynamic programs is straightforward, but we need the following additional restriction.

Assumption 9: F is continuously differentiable on the interior of A.

Theorem 35 *(Differentiability of the value function) Let X, F, Γ, and β satisfy Assumption 3-4 and 7-9; let v be the unique solution to FE (3.12); and let policy function g satisfies (3.13). If $x_0 \in intX$ and $g(x_0) \in int\Gamma(x_0)$, then v is continuously differentiable at x_0, with derivatives given by*

$$v_i(x_0) = F_i[x_0, g(x_0)], i = 1, 2, ..., l.$$

Proof. Since $g(x_0) \in int\Gamma(x_0)$ and Γ is continuous, it follows that $g(x_0) \in int\Gamma(x)$, for all x in some neighborhood D of x_0. Define W on D by

$$W(x) = F[x, g(x_0)] + \beta v[g(x_0)].$$

Since F is concave and differentiable in its first l arguments (Assumption 7), and differentiable (Assumption 9), it follows that W is concave and differentiable w.r.t. x. Moreover, since $g(x_0) \in int\Gamma(x)$, for all $x \in D$, it follows that

$$W(x) \leq \max_{y \in \Gamma(x)} [F(x, y) + \beta v(y)] = v(x), \ \forall x \in D,$$

with equality at x_0. Hence v and W satisfy the hypothesis of Benveniste-Scheinkman Theorem, it thus establishes the result. ■

This theorem gives us an easy way to derive Euler equations from the recursive formulation of the neoclassical growth model. Recall FE

$$v(k) = \max_{0 \leq k' \leq f(k)} U(f(k) - k') + \beta v(k')$$

Take the FOC w.r.t. k', we get

$$U'(f(k) - k') = \beta v'(k')$$

Denote by $k' = g(k)$ the optimal policy function. The problem is that we do not know v'. That is the reason why we need Benveniste-Scheinkman theorem. According to Theorem 35, we have

$$v'(k) = U'(f(k) - k')f'(k)$$

Substituting it into FOC

$$\begin{aligned} U'(f(k) - k') &= \beta v'(k') \\ &= \beta U'(f(k') - k'')f'(k') \end{aligned}$$

That is the same Euler equation (second-order difference equation) as we obtain by using Euler equation approach to solve SP.

Chapter 4

Computational Dynamic Programming

The previous chapter lays out the theoretical foundation of dynamic programming. Under ordinary assumptions of objective function and constraint set, we show the existence and uniqueness of the value function. We also show that this value function is strictly increasing, strictly concave, and continuously differentiable. And there is a unique and time-invariant optimal policy function g for the functional equation.

The goal of this chapter is to show how we can solve FE

$$v(x) = \max_{y \in \Gamma(x)} [F(x, y) + \beta v(y)] \tag{4.1}$$

practically. We first introduce three main types of computational methods for solving DP. Then we will talk about two popular methods for obtaining numerical approximations of value and policy functions.

4.1 Three Computational Methods

There are three widely used computational methods for solving dynamic programming problems like the one in (4.1).

4.1.1 Value Function Iteration

The first method proceeds by constructing a sequence of value functions and the associated policy functions. The sequence is created by iterating on the following equation, starting from $v_0 = 0$, and continuing until v_n has converged[1]:

$$v_{n+1}(x) = \max_{y \in \Gamma(x)} [F(x, y) + \beta v_n(y)].$$

[1] In practice, we will use a tolerance say $tol = 10^{-8}$ for the stopping criteria of the converegnce. i.e., we stop when $\|v_{n+1} - v_n\| < tol$.

As we said before, the Contraction Mapping Theorem is the theoretical foundation for this method. It assets that no matters what initial guess v_0 is, the sequence $\{v_n\}_{n=0}^{\infty}$ will converge to fixed point v at a geometric rate of β.

4.1.2 Guess and Verify

As the name of this method, "Guess and Verify" method involves guessing a functional form of the value function v and then substituting the guess back into FE to verify that it is indeed a correct guess. This method relies on the uniqueness of the solution to FE. But because it relies on luck in making a good guess, it is not generally available.

The optimal growth model with utility function $U(c) = \ln c$ and Cobb-Douglas production function $f(k) = k^{\alpha}$ is a example that we can apply "Guess and Verify" method. Please see Section 1.2.3 for a detailed explanation.

Here is another example that we can use "Guess and Verify".

Example 36 *Consider optimal growth model with $U(c) = \ln c$ and production function $f(k) = k^2$. This problem is called "Cake-Eating". FE is*

$$v(k) = \max_{0 \le k' \le k} [\ln(k - k') + \beta v(k')] \tag{4.2}$$

Let's guess $v(k) = B \ln k$, B is the coefficient to be determined. Substituting our guess into FE above, we obtain

$$B \ln k = v(k) = \max_{0 \le k' \le k} [\ln(k - k') + \beta B \ln k']$$

FOC w.r.t. k' is

$$\frac{1}{k - k'} = \frac{\beta B}{k'}$$

$\Rightarrow$

$$k' = \frac{\beta B}{1 + \beta B} k$$

Substituting this optimal policy function back into FE

$$\begin{aligned} B \ln k &= v(k) \\ &= \ln(\frac{1}{1+\beta B} k) + \beta B \ln(\frac{\beta B}{1 + \beta B} k) \\ &= (1 + \beta B) \ln k + \beta B \ln \frac{\beta B}{1 + \beta B} + \ln(\frac{1}{1 + \beta B}) \end{aligned}$$

which implies

$$\begin{aligned} B &= 1 + \beta B \\ &\Rightarrow \\ B &= \frac{1}{1 - \beta} \end{aligned}$$

[2] Obviously, the "cake-eating" problem is a special case for the general case with log utility function and C-D production technology.

Hence

$$\begin{aligned} v(k) &= \frac{1}{1-\beta}\ln k \\ k' &= g(k) = \beta k \end{aligned}$$

4.1.3 Policy Function Iteration

A third method, known as *policy function iteration* or *Howard's improvement algorithm*, consists of the following steps:

1. Pick a feasible policy, $y = g_0(x)$, and compute the value function associated with operating forever with this policy

$$v_{h_n}(x) = \sum_{t=0}^{\infty} \beta^t F[x_t, g_n(x_t)]$$

where $x_{t+1} = g_n(x_t)$ with $n = 0$.

2. Generate a new policy $y = g_{n+1}(x)$ that solves the two-period problem for each x

$$\max_y [F(x, y) + \beta v_{h_n}(y)].$$

3. Iterate over n to convergence on steps 1 and 2.

The idea about the policy function iteration method is following: Suppose we start from an initial guess of policy function $y = g_0(x)$. Plug this policy function into SP, we can calculate the associated value function. Now given this value function, we go back to FE, to solve a dynamic problem that economy will follow the suggested policy function since tomorrow to forever. We solve this FE to obtain a new policy function. If this new policy function is as same as the one guessed before, we achieve the time-invariant optimal policy function, hence we stop. Otherwise, we use the newly solved one to replace the initial guess to start the process again until it converges.

The theoretical foundation for policy function iteration is Theorem 32. Since this algorithm does not require maximization (maximization is usually the expensive step in these computations), this method ofen converges faster than does value function iteration.

So far, only two special types of dynamic problems can be analytically solved. One is with logarithmic preference and Cobb-Douglas constraints, which is studied in this chapter. Another is an economy with linear constraints and quadratic preference which will be studied in the next chapter. Therefore we have to rely on computer to do the iterations for us.

4.2 Practical Dynamic Programming

What can we do if we cannot obtain closed form solutions for iterating on the Bellman equation (or functional equation)? We have to adopt some numerical approximations. This section will describe two popular methods for obtaining numerical approximations for value function: discrete-state dynamic programming and polynomial approximations.

4.2.1 Discrete-state Dynamic Programming

We introduce the method of discretization of the state space in the context of a particular discrete-state version of a stochastic optimal savings problem. An infinitely lived household likes to consume one good, which it can acquire by using labor income or accumulated savings. The household has an endowment of labor at time t , s_t, that evolves according to an m-state Markov chain with transition matrix P (See Section 1.3.1 for details). If the realization of the process at t is $\bar{s}_i$, then at time t the household receives labor income of amount $w\bar{s}_i$. The wage w is fixed over time. For simplicity, we sometimes assume that m is 2, and that s_t takes on value 0 in an unemployed state and 1 in an employed state. In this case, w has the interpretation of being the wage of employed workers.

The household can choose to hold a single asset in discrete amount $a_t \in \mathcal{A}$, where $\mathcal{A}$ is a grid $[a_1 < a_2 < ... < a_n]$. The interest rate for asset holding r is constant over time.

For given values of $\{w, r\}$ and given initial values of $\{a_0, s_0\}$, the HH's optimization problem is

$$\begin{aligned} & \max_{\{c_t, a_{t+1}\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t U(c_t) \\ & s.t. \\ c_t + a_{t+1} \leq & \ (1+r)a_t + ws_t \\ c_t \geq & \ 0, a_{t+1} \in \mathcal{A} \end{aligned}$$

The associated Bellman equation (FE) is

$$v(a, s) = \max_{a' \in \mathcal{A}} \{U[(1+r)a + ws - a'] + \beta E v(a', s') \mid s\}.$$

Or more specifically, in terms of discrete-state space, for each $i \in [1, 2, ..., m]$ and $h \in [1, 2, ..., n]$,

$$v(a_h, \bar{s}_i) = \max_{a' \in \mathcal{A}} \{U[(1+r)a_h + w\bar{s}_i - a'] + \beta \sum_{j=1}^{m} \pi_{ij} v(a', \bar{s}_j)\} \tag{4.3}$$

A solution of this problem is a value function $v(a, s)$ that satisfies equation (4.3) and an associated policy function $a' = g(a, s)$ mapping this period's state variables (a, s) pair into an optimal choice of assets to carry into next period.

Value Function Iteration

We pick value function iteration method to slove this Bellman equation numerically. But how to write a computer program (for example Matlab code) to do iteration on the Bellman equation (4.3)?

Let $\mathcal{A} = [a_1, a_2, ..., a_n]$ be the asset space and $\mathcal{S} = [s_1, s_2]$ be the discrete space for employment status. Define two $n \times 1$ vectors v_j, $j = 1, 2$, whose ith rows are determined by $v_j(i) = v(a_i, s_j)$, $i = 1, 2, ..., n$. Define $\mathbf{1}$ be the $n \times 1$ vector consisting entirely of ones. i.e., $\mathbf{1} = [1, 1, ..., 1]'$. Define two $n \times n$ matrices R_j, $j = 1, 2$, whose (i, h) element is

$$R_j(i, h) = U[(1 + r)a_i + ws_j - a_h], \ i = 1, ..., n, h = 1, ..., n.$$

In words, $R_j(i, h)$ is the utility when the employment status at time t is s_j, the asset holding at the beginning of period t is a_i, and the agent will choose to hold asset amount a_h through this period.

Define an operator $T([v_1, v_2])$ that maps a pair of vectors $[v_1, v_2]$ into a pair of vector $[tv_1, tv_2]$:

$$\begin{aligned} \underbrace{tv_1}_{n\times 1} &= \max\{\underbrace{R_1}_{n\times n} + \beta\pi_{11}\underbrace{\mathbf{1}}_{n\times 1}\underbrace{v_1'}_{1\times n} + \beta\pi_{12}\underbrace{\mathbf{1}}_{n\times 1}\underbrace{v_2'}_{1\times n}\} \\ tv_2 &= \max\{R_2 + \beta\pi_{21}\mathbf{1}v_1' + \beta\pi_{22}\mathbf{1}v_2'\}. \end{aligned} \tag{4.4}$$

Here the "max" operator applied to a $n \times n$ matrix M returns an $(n \times 1)$ vector whose ith element is the maximum of the ith row of the matrix M. These two equations can be written compactly as

$$\begin{bmatrix} tv_1 \\ tv_2 \end{bmatrix} = \max\{\begin{bmatrix} R_1 \\ R_2 \end{bmatrix} + \beta(P \otimes \mathbf{1})\begin{bmatrix} v_1' \\ v_2' \end{bmatrix}\},$$

where $\otimes$ is the Kronecker product.

The Bellman equation then can be represented as a fixed point

$$(v_1, v_2) = T([v_1, v_2]),$$

and can be solved by iterating to converegnce on

$$[v_1, v_2]_{m+1} = T([v_1, v_2]_m).$$

Let's briefly summarize the procedure of a standard value function iteration method:

1. Choosing a functional form for the utility function and production technology.

2. Discretizing the state and control variables.

3. Building a computer code to perform value function iteration.

4. Evaluating the value and the policy function.

Policy Function Iteration

Ofen time computation speed is very important. As we mentioned before, policy improvement algorithm can be much faster than directly iterating on the Bellman equation. It is also easy to implement the Howard improvement algorithm in the present setting.

Consider a given feasible policy function $a' = g(a, s)$. For each h, define the $n \times n$ matrices J_h by

$$J_h(a, a') = \begin{cases} 1 & \text{if } g(a, s_h) = a' \\ 0 & \text{otherwise} \end{cases}$$

Here $h = 1, 2, ..., m$ where m is the number of possible values for state s_t, and $J_h(a, a')$ is the element of J_h with rows corresponding to initial assets a and columns to terminal assets a'. $J_h(a, a')$ is an index function. When a' satisfies the policy function, it is one; otherwise zero.

For a given policy function $a' = g(a, s)$, define the $n \times 1$ vector r_h with rows corresponding to

$$r_h(a) = U[(1 + r)a + ws_h - g(a, s_h)] \tag{4.5}$$

for $h = 1, 2, ..., m$. Therefore, vector r_h is the utility level when the state is s_h, the asset holding at the beginning of current period is a, and the next period asset holding is determined by policy function $g(a, s_h)$. The ith element of r_h is corresponded to the current asset holding a_i and the next period asset holding $g(a_i, s_h)$.

Suppose the policy function $a' = g(a, s)$ will be used forever. Let the value function associated with using $g(a, s)$ forever be represented by the m vectors $[v_1, v_2, ..., v_m]$. Each vector v_h, $h = 1, 2, ..., m$ is a $n \times 1$ vector where its ith element $v_h(a_i)$ is the value of lifetime utility starting from state (a_i, s_h). Suppose that $m = 2$, the vectors $[v_1, v_2]$ obey

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} + \begin{bmatrix} \beta\pi_{11}J_1 & \beta\pi_{12}J_1 \\ \beta\pi_{21}J_2 & \beta\pi_{22}J_2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}. \tag{4.6}$$

Then

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \left[I - \beta \begin{pmatrix} \pi_{11}J_1 & \pi_{12}J_1 \\ \pi_{21}J_2 & \pi_{22}J_2 \end{pmatrix} \right]^{-1} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \tag{4.7}$$

where I is a $2n \times 2n$ identity matrix.

Here is how to implement the Policy Function Iteration algorithm.

Step 1. For an initial feasible policy function $g_j(a, s)$ for $j = 1$, form the r_h matrices using equation (4.5), then use equation (4.7) to evaluate the vectors of values $[v_1^j, v_2^j]$ implied by using that policy forever.

Step 2. Use $[v_1^j, v_2^j]$ as the terminal value vectors in Bellman equation (4.4), and perform one step on the Bellman equation to find a new policy function $g_{j+1}(a, s)$ for $j + 1 = 2$. Use this policy function, update j, and repeat step 1.

Step 3. Iterate to converegnece on steps 1 and 2. i.e., if $\left\| [v_1^{j+1}, v_2^{j+1}] - [v_1^j, v_2^j] \right\| < \varepsilon$, stop.

Note that having a good initial guess of policy function is crucial. Given a good guess close to the actual policy function, policy function iteration often takes a few iterations. Unfortunately, the inverse operation in step 2 may be expensive, making each iteration costly. Researchers have had an acceleration procedure called *modified policy iteration with k steps*, implements the basic idea of policy iteration without intensively computing the inverse $\left[I - \beta \begin{pmatrix} \pi_{11}J_1 & \pi_{12}J_1 \\ \pi_{21}J_2 & \pi_{22}J_2 \end{pmatrix}\right]^{-1}$. The idea is for $k \geq 2$, iterate k times on Bellman equation (4.4). Take the resulting policy and value function and use (4.7) to produce a new candidate value function. Then starting from this terminal value function, perform another k iterations on the Bellman equation. Continue in this fashion until the decision rule converges. In this way, we avoid intensively repeating inverse operation in step 2 by mixing the advantages of value function iteration and policy function iteration.

4.3 Polynomial Approximations

In this section, we describe a numerical method for iterating on the Bellman equation using a polynomial to approximate the value function and a numerical optimizer to perform the optimization at each iteration.

We describe this method in the context of the Bellman equation for a particular economic problem: Hopenhayn and Nicolini (1997)'s model of optimal umemployment insurance. In their model, a planner wants to provide incentives to an unemployed worker to search for a new job while also partially insuring the worker against bad luck in the search process. The planner seeks to deliver discounted expected utility V to an unemployed worker at minimum cost while providing proper incentives to search for work. Hopenhayn and Nicolini show that the minimum cost $C(V)$ satisfies the Bellman equation as following

$$C(V) = \min_{V^u} \{c + \beta[1 - p(a)]C(V^u)\} \tag{4.8}$$

where c, a are given by

$$c = U^{-1}[\max(0, V + a - \beta\{p(a)V^e + [1 - p(a)]V^u\})], \tag{4.9}$$

and

$$a = \max\{0, \frac{\log[r\beta(V^e - V^u)]}{r}\}. \tag{4.10}$$

Here V is a discounted present value that an insurer (social planner) has promised to an umemployed worker. V^e is the expected sum of discounted utility of an employed worker. V^u is a value for next period that the insurer promises the worker if he remains unemployed. $1 - p(a)$ is the probability of remaining unemployed if the worker exerts search effort a, and c is the worker's consumption level.

$$\begin{aligned} V^e &= \frac{U(w)}{1-\beta} \\ V^u &= \max_{a \geq 0}\{U(0) - a + \beta[p(a)V^e + (1-p(a))V^u]\} \end{aligned}$$

Hopenhayn and Nicolini assume that $p(a) = 1 - \exp(ra), r > 0$.

4.3.1 Computational Strategy

To approximate the solution of the Bellman equation (4.8), we use a polynomial to approximate the ith iteration $C_i(V)$ of $C(V)$. This polynomial is stored on the computer in terms of $n+1$ coefficients. Then at each iteration, the Bellman equation is to be solved at a small number $m \geq n+1$ values of V. This procedure gives values of the ith iterate of the value function $C_i(V)$ at those particular V's. Then we interpolate (or "connect the dots") to fill in the continuous function $C_i(V)$. Substituting this approximation $C_i(V)$ for $C(V)$ in equation (4.8), we pass the minimum problem on the right side of equation (4.8) to a numerical minimizer. Programming languages like Matlab and Gauss have easy-to-use algorithms for minimizing continuous functions of several variables. We solve one such numerical problem minimization for each node value for V. Doing so yields optimized value $C_{i+1}(V)$ at those node points. We then interpolate to build up $C_{i+1}(V)$. We iterate on this scheme to convergence.

4.3.2 Chebyshev Polynomials

According to the Weierstarss Theorem, any continuous real-valued function f defined on a bounded interval $[a, b]$ of the real line can be approximated to any degree of accuracy using a polynomial. But what kind of polynomials are we going to use to approximate $C(V)$? The widely used one is called *Chebyshev Polynomials. (*Thoerem in Numerical Analysis shows that Chebyshev polynomials are optimal.)

for nonnegative integer n and $x \in R$, the nth Chebyshev polynomial is

$$T_n(x) = \cos(\frac{n}{\cos x})$$

Given coefficients c_j, $j = 0, 1, ..., n$, the nth-order Chebyshev polynomial approximator is

$$C_n(x) = c_0 + \sum_{i=1}^{n} c_j T_j(x). \tag{4.11}$$

And it is defined over the interval $[-1, 1]$. For computational purposes, we want to form an approximator to a real-valued function f of the form (4.11). Note that a nice thing about this polynomial is this approximator simply can be stored as the $n+1$ coefficients c_j, $j = 0, 1, ..., n$. To form the approximator, we evaluate $f(x)$ at $n+1$ carefully chosen points (called "*nodes*"), then use a least square formula to form the c_j's in equation (4.11).

In particular, to interpolate a function of a single variable x with domain $x \in [-1, 1]$, Judd (1998) recommends evaluating the function f at the $m \geq n+1$ points x_k, $k = 1, 2, ..., m$, where

$$x_k = \cos(\frac{2k-1}{2m}\pi), k = 1, 2, ..., m. \tag{4.12}$$

Here x_k is the zero of the kth Chebyshev polynomial on $[-1, 1]$. Given the $m \geq n+1$ values of $f(x_k)$ for $k = 1, 2, ..., m$, we choose the "least square" values of coefficient c_j

$$c_j = \frac{\sum_{k=1}^{m} f(x_k)T_j(x_k)}{\sum_{k=1}^{m} T_j(x_k)^2}, j = 0, 1, ..., n \tag{4.13}$$

4.3.3 Numerical Algorithm

In summary, applied to the Hopenhayn-Nicolini model, the numerical procedure consists of the following steps:

1. Choose upper and lower bounds for V^u, so that V and V^u will be in the interval $[\underline{V}^u, \bar{V}^u]$. In particular, set $\bar{V}^u = V^e - \frac{1}{\beta p'(0)}$, the bound required to assure positive search effort $a > 0$.

2. Choose a degree n for the approximator, a Chebyshev polynomial, and a number $m \geq n + 1$ of nodes or grid points.

3. Generate the m zeros of the Chebyshev polynomial on the domain $[-1, 1]$, given by (4.12).

4. By a change of scale, transform the x_k's to the corresponding points V_l^u in $[\underline{V}^u, \bar{V}^u]$.

5. Choose initial values of the $n+1$ coefficients in the Chebyshev polynomial, for example, $c_j = j, \forall j = 0, 1, ..., n$. Use these coefficients to define the function $C_i(V^u)$ for the iteration number $i = 0$.

6. Compute the function $\tilde{C}_i(V) \equiv c + \beta[1 - p(a)]C_i(V^u)$, where c, a are determined as functions of $V, V^e(= \frac{U(w)}{1-\beta}), V^u$ from equation (4.9) and (4.10).

7. For each point V_l^u, use a numerical minimization program to find $C_{i+1}(V_l^u) = \min_{V^u} \tilde{C}_i(V^u)$.

8. Using these m values of $C_{i+1}(V_l^u)$, compute new values of the coefficients in the Chebyshev polynomials by using "least squares" (equation (4.13)). Return to step 5 and iterate to convergence.

Chapter 5

Linear Quadratic Dynamic Programming

This chapter describes the class of dynamic programming problems in which the return function is quadratic and the transition function (law of motion) is linear. This specification leads to the widely used optimal linear regulator problem, for which the Bellman equation can be solved analytically using linear algebra. We consider the special case in which the return function and transition function are both time invariant, although the mathematics is almost identical when they are permitted to be deterministic functions of time.

Linear quadratic dynamic programming has two uses for us. A first is to study optimum and equilibrium problems arising for linear rational expectations models. Here the dynamic decision problems naturally take the form of an optimal linear regulator. A second is to use a linear quadratic dynamic program to approximate one that is not linear quadratic.

5.1 The Optimal Linear Regulator Problem

5.1.1 The Linear Quadratic Problem

The undiscounted optimal linear regulator problem (OLRP) is to maximize over choice of $\{u_t\}_{t=0}^{\infty}$

$$\begin{aligned} & \max_{\{u_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \{x_t' R x_t + u_t' Q u_t\} \\ & s.t. \\ x_{t+1} \quad = \quad & A x_t + B u_t \\ & x_0 \text{ given.} \end{aligned}$$

Here x_t is an $n \times 1$ vector of state variables, u_t is an $k \times 1$ vector of control variables, R is a negative semidefinite (NSD) symmetric matrix, Q is a positive definite (PD) symmetric matrix, A is an $n \times n$ matrix, and B is an $n \times k$ matrix.

We guess (recall the "Guess and Verify" method) that the value function is quadratic,

$$v(x) = x'Px$$

where P is a positive semidefinite (PSD) symmetric matrix. We want to know what does P look like.

Using the transition law $x_{t+1} = Ax_t + Bu_t$ to eliminate next period's state, the Bellman equation becomes

$$\begin{aligned} v(x_t) &= \max\{x_t'Rx_t + u_t'Qu_t + v(x_{t+1})\} \\ &= \max\{x_t'Rx_t + u_t'Qu_t + {x_{t+1}}'Px_{t+1}\} \\ &= \max\{x_t'Rx_t + u_t'Qu_t + (Ax_t + Bu_t)'P(Ax_t + Bu_t)\} \end{aligned}$$

We can ignore the time subscript and rewrite RHS of Bellman equation as following

$$v(x) = \max_u\{x'Rx + u'Qu + x'A'PAx + x'A'PBu + u'B'PAx + u'B'PBu\}$$

FOC w.r.t. u is[1]

$$(Q + Q')u + B'P'Ax + B'PAx + (B'PB + B'P'B)u = 0$$

Since Q and P are symmetric, we have

$$2Qu + 2B'PBu + 2B'PAx = 0$$

$\Rightarrow$

$$(Q + B'PB)u = B'PAx,$$

which implies the feedback rule (policy function) for u:

$$\begin{aligned} u &= -(Q + B'PB)^{-1}B'PAx \\ &\quad \text{or} \\ u &= -Fx, \text{ where } F = (Q + B'PB)^{-1}B'PA \end{aligned} \tag{5.1}$$

Substituting back the policy function (5.1) to Bellman equation, we have[2]

$$\begin{aligned} v(x) &= x'Rx + (-x'F')Q(-Fx) + x'A'PAx + x'A'PB(-Fx) - x'F'B'PAx + (-x'F')B'PB(-Fx) \\ &= x'Rx + x'F'QFx + x'A'PAx - x'A'PBFx - x'F'B'PAx + x'F'B'PBFx \\ &= x'Rx + x'A'PAx - x'A'PBFx - x'F'B'PAx + x'F'(Q + B'PB)Fx \\ &= x'Rx + x'A'PAx - x'A'PBFx - x'F'B'PAx + x'F'(Q + B'PB)(Q + B'PB)^{-1}B'PAx \\ &= x'Rx + x'A'PAx - x'A'PBFx - x'F'B'PAx + xF'B'PAx \\ &= x'Rx + x'A'PAx - x'A'PBFx \\ &= x'(R + A'PA - A'PBF)x \end{aligned}$$

[1] We use the following rules for differentiating quadratic and bilinear matrix forms: $\frac{\partial x'Ax}{\partial x} = (A + A')x$; $\frac{\partial y'Bz}{\partial y} = Bz$; $\frac{\partial y'Bz}{\partial z} = B'y$.

[2] Here we use transpose and inverse properties of matrix. $(A')' = A; (A + B)' = A' + B'; (AB)' = B'A'; AA^{-1} = I; (A^{-1})' = (A')^{-1}$. We also use the fact that P, Q, R are symmetric matrix.

Note that by guess, we have

$$v(x) = x'Px$$

This implies that in order to justify this guess, we should have

$$\begin{aligned} P &= R + A'PA - A'PBF \\ &= R + A'PA - A'PB(Q + B'PB)^{-1}B'PA \end{aligned} \tag{5.2}$$

Equation (5.2) is called the *algebraic matrix Riccati equation.* It expresses the matrix P as an implicit function of the exogenously given matrices R, Q, A, B. Solving this equation for P requires a computer whenever P is larger than a 2×2 matrix.

5.1.2 Value Function Iteration

How to solve Riccati equation practically? Esstentially, Riccati equation is a Bellman equation (think P as a value function v). Therefore, we can also apply the numerical methods for solving Bellman equation here.

Under particular conditions, we can show that equation (5.2) has a unique negative semidefinite solution, which is approached in the limit as $j \to \infty$ by iterations on the matrix diference Riccati equation:

$$P_{j+1} = R + A'P_jA - A'P_jB(Q + B'P_jB)^{-1}B'P_jA,$$

starting from initial guess $P_0 = 0$. The policy function associated with P_j is

$$F_{j+1} = (Q + B'P_jB)^{-1}B'P_jA$$

5.1.3 Discounted Optimal Linear Regulator Problem

The discounted OLRP is

$$\begin{aligned} & \max_{\{u_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \{x_t'Rx_t + u_t'Qu_t\} \\ & s.t. \\ x_{t+1} &= Ax_t + Bu_t \\ & x_0 \text{ given,} \end{aligned}$$

where discount factor $\beta \in (0, 1)$.

It is easy to show (we leave it as an exercise) that the matrix Riccati difference equation for this problem is

$$P_{j+1} = R + \beta A'P_jA - \beta^2 A'P_jB(Q + \beta B'P_jB)^{-1}B'P_jA.$$

The associated policy function is

$$F_{j+1} = \beta(Q + \beta B'P_jB)^{-1}B'P_jA$$

5.1.4 Policy Function Iteration

The policy improvement algorithm can be applied to solve both the undiscounted and discounted optimal linear regulator problem. Starting from an initial F_0 for which the eigenvalues of $A - BF_0$ are less than $\frac{1}{\sqrt[2]{\beta}}$ in modulus[3], the algorithm iterates on the two equations

$$\begin{aligned} P_j &= R + F_j'QF_j + \beta(A - BF_j)'P_j(A - BF_j) \\ F_{j+1} &= \beta(Q + \beta B'P_jB)^{-1}B'P_jA \end{aligned} \tag{5.3}$$

The first equation is a result for substituting $x_{t+1} = Ax_t + Bu_t = Ax_t + B(-Fx_t) = (A - BF)x_t$ and $u_t = -Fx_t$ into the Bellman equation. It pins down the matrix for the quadratic form in the value function associated with using a fixed rule F_j forever. The solution of this equation can be represented in the form

$$P_j = \sum_{k=0}^{\infty} \beta^k((A - BF_j)')^k(R + F_j'QF_j)(A - BF_j)^k \tag{5.4}$$

The algorithm for policy function iteration is as following:

1. Pick initial guess for policy function F_0. By using (5.4), obtain P_0.
2. Substituting P_0 into policy function equation (5.3) to obtain F_1.
3. If $\|F_1 - F_0\| < \varepsilon$, stop. Otherwise, go back to step 1 and repeat until the converegnce.

5.1.5 Stochastic Optimal Linear Regulator Problem

The stochastic discounted linear optimal regulator problem is to choose a decision rule for u_t to maximize

$$\begin{aligned} & E_0 \max_{\{u_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t\{x_t'Rx_t + u_t'Qu_t\} \\ & s.t. \\ x_{t+1} &= Ax_t + Bu_t + \varepsilon_{t+1} \\ & x_0 \text{ given.} \end{aligned}$$

ε_{t+1} is an $(n \times 1)$ vector random variables that is independently and identically distributed (i.i.d.) through time and obeys the normal distribution with mean vector zero and convariance matrix Σ. i.e., $\varepsilon_{t+1} \sim N(0, \Sigma)$ and we have

$$E\varepsilon_t = 0; E\varepsilon_t\varepsilon_t' = \Sigma.$$

[3]This requirement is for the stability of the dynamic system.

We guess value function takes the form

$$v(x) = x'Px + d \tag{5.5}$$

where d is just a scalar. Now let's use "Guess and Verify" method to verify what is P and d.

Substituting guess (5.5) into the Bellman equation

$$\begin{aligned} v(x) &= \max_u \{x'Rx + u'Qu + \beta E[(Ax + Bu + \varepsilon)'P(Ax + Bu + \varepsilon)] + \beta d\} \\ &= \max_u \{x'Rx + u'Qu + \beta E[x'A'PAx + x'A'PBu \\ &\quad + x'A'P\varepsilon + u'B'PAx + u'B'PBu + u'B'P\varepsilon \\ &\quad + \varepsilon'PAx + \varepsilon'PBu + \varepsilon'P\varepsilon] + \beta d\}. \end{aligned}$$

Note that $E\varepsilon = 0$,

$$\begin{aligned} v(x) &= \max_u \{x'Rx + u'Qu + \beta x'A'PAx + \beta x'A'PBu \\ &\quad + \beta u'B'PAx + \beta u'B'PBu + \beta E\varepsilon'P\varepsilon + \beta d\} \end{aligned}$$

FOC w.r.t. u is

$$(Q + \beta B'PB)u = -\beta B'PAx$$

Hence we have policy function

$$u = -Fx, \; F = \beta(Q + \beta B'PB)^{-1}B'PA \tag{5.6}$$

which is identical to the certainty case.

Since $E\varepsilon_i\varepsilon_j = 0$ for $i \neq j$ (recall ε is i.i.d.), we have $E\varepsilon'P\varepsilon = tr(E\varepsilon'P\varepsilon) = trP(E\varepsilon'\varepsilon) = tr(P\Sigma)$. Substituting policy function (5.6) into the Bellman equation, and using (5.5), we obtain

$$P = R + \beta A'PA - \beta^2 A'PB(Q + \beta B'PB)^{-1}B'PA$$

and

$$d = \beta(1-\beta)^{-1}tr(P\Sigma)$$

Theorem 37 *(Certainty Equivalence Principle): The decision rule that solves the stochastic optimal linear regulator problem is identical with the decision rule for the corresponding nonstochastic linear optimal regulator problem.*

The remarkable thing about this principle is that although through d the value function depends on Σ, the optimal decision rule is not. In other words, the optimal decision rule is independent of the problem's noise statistics. The certainty equivalence principle is a special property of the optimal linear regulator problem and comes from the quadratic objective function, the linear transition equation, and the property of random variable ε. Certainty equivalence does not characterize stochastic control problems generally.

5.2 Linear-Quadratic Approximations

This section describes an important use of the OLRP: to approximate the solution of more complicated dynamic programs. We already know that for most of the dynamic decision problems there are no closed-form solutions. But the beauty of the LQ problem is that we can have solution to Bellman equation by quickly using linear algebra. Therefore, the idea is that if we can approximate a dynamic problem to a corresponding LQ problem, we can overcome this difficulty.

Optimal linear regulator problems are often used to approximate problems of the following form:

$$\begin{aligned} & \max_{\{u_t\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t r(z_t) \\ & s.t. \\ x_{t+1} &= Ax_t + Bu_t + Cw_{t+1} \end{aligned} \tag{5.7}$$

where $\{w_{t+1}\}$ is a vector of i.i.d. random disturbances with mean zero and finite variance. $r(z_t)$ is a concave and twice continuously differentiable function of $z_t \equiv \begin{pmatrix} x_t \\ u_t \end{pmatrix}$. Here x_t is the state variable and u_t is the control variable.

Let's use the sotchastic optimal growth model, which is the workhorse of so-called Real Business Cycle (RBC) theory, to describe how can we transform it to a LQ problem. The social planner's problem is

$$\begin{aligned} & \max_{\{c_t,k_{t+1}\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t \ln c_t \\ & s.t. \\ c_t + i_t &= A\theta_t k_t^{\alpha} \\ k_{t+1} &= (1-\delta)k_t + i_t \\ \ln \theta_{t+1} &= \rho \ln \theta_t + w_{t+1} \end{aligned} \tag{5.8}$$

where $\{w_{t+1}\}$ is an i.i.d. random disturbances with mean zero and finite variance. θ_t is a technology shock. Define $\tilde{\theta}_t \equiv \ln \theta_t$. To transform this problem into the form as in (5.7), we can define

$$\begin{aligned} x_t &= \begin{pmatrix} k_t \\ \tilde{\theta}_t \end{pmatrix}, u_t = i_t \\ r(z_t) &= \ln(Ak_t^{\alpha} \exp \tilde{\theta}_t - i_t). \end{aligned}$$

Therefore, we can rewrite the laws of motion as

$$\begin{pmatrix} 1 \\ k_{t+1} \\ \tilde{\theta}_{t+1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & (1-\delta) & 0 \\ 0 & 0 & \rho \end{pmatrix} \begin{pmatrix} 1 \\ k_t \\ \tilde{\theta}_t \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} i_t + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} w_{t+1}$$

where it is convenient to add the constant 1 as the first component of the state vector.

Now here is the solution algorithm for solving problems as in (5.7).

1. Choose a point about which to expand the return function. In most cases, this point is the steady state of the deterministic version of the model economy, call it $\bar{z}$, that we obtain when we substitute the random variables with their unconditional means.

2. Construct a quadratic approximation of $r(z_t)$ around $\bar{z}$ by using Taylor's expansion.

3. Compute the optimal value function $v(x)$ by value function iteration method.

Now let's explain more detailedly step by step.

How to determine $\bar{z}$? Usually, $\bar{z}$ is chosen as the (optimal) stationary state of the non-stochastic version of the original nonlinear model:

$$\begin{aligned} & \max_{\{u_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t r(z_t) \\ & s.t. \\ x_{t+1} &= Ax_t + Bu_t. \end{aligned}$$

This stationary point is obtained in following steps:

1. Find the Euler equations.

2. Substitute $z_{t+1} = z_t = \bar{z}$ into the Euler equations and transition laws, and solve the resulting system of nonlinear equations for $\bar{z}$.

For example, for the following deterministic neoclassical growth model

$$\begin{aligned} & \max \sum_{t=0}^{\infty} \beta^t \ln c_t \\ & s.t. \\ c_t + i_t &= k_t^{\alpha} \\ k_{t+1} &= (1-\delta)k_t + i_t \end{aligned}$$

We know the Euler equation for this economy is

$$c_{t+1} = \beta c_t[\alpha k_t^{\alpha-1} + 1 - \delta]$$

Imposing the steady state condition $c_{t+1} = c_t = \bar{c}$ and $k_t = \bar{k}$, we have

$$\begin{aligned} \bar{k} &= \left(\frac{\alpha}{\frac{1}{\beta} - 1 + \delta}\right)^{1/(1-\alpha)} \\ \bar{\imath} &= \delta\bar{k} \end{aligned}$$

Therefore, the steady state vector for this economy is

$$\bar{z} = \begin{pmatrix} \bar{k} \\ \bar{\theta} \\ \bar{\imath} \end{pmatrix} = \begin{pmatrix} (\frac{\alpha}{\frac{1}{\beta}-1+\delta})^{1/(1-\alpha)} \\ 0 \\ \delta(\frac{\alpha}{\frac{1}{\beta}-1+\delta})^{1/(1-\alpha)} \end{pmatrix}.$$

Next, we want to show how to replace return function $r(z_t)$ by a quadratic form $z_t' M z_t$. We use the method described in Kydland-Prescott (1982). The basic isea is to choose a point $\bar{z}$ and approximate with the first two terms of a Taylor series:

$$\hat{r}(z) = r(\bar{z}) + (z - \bar{z})' \frac{\partial r(\bar{z})}{\partial z} + \frac{1}{2}(z - \bar{z})' \frac{\partial^2 r(\bar{z})}{\partial z \partial z'}(z - \bar{z}). \tag{5.9}$$

Note that z_t is an $(n + k + 1) \times 1$ vector (recall the first element of z_t is the constant one). Let e be the $(n + k) \times 1$ vector with 0's everywhere except for a 1 in the row corresponding to the location of the constant unity in the state vector, i.e., $e' = (1, 0, 0, ..., 0)$. Therefore, we have $e' z_t \equiv 1$ for all t.

Repeatedly using $e'z = z'e = 1$, we can express equation (5.9) as (leave as an exercise for verifying it)

$$\hat{r}(z) = z' M z,$$

where

$$\begin{aligned} M \quad = \quad & \underbrace{e[r(\bar{z}) - (\frac{\partial r(\bar{z})}{\partial z})' \bar{z} + \frac{1}{2} \bar{z}' \frac{\partial^2 r(\bar{z})}{\partial z \partial z'} \bar{z}] e'}_{M_{11}} \\ & + \frac{1}{2} \underbrace{[\frac{\partial r(\bar{z})}{\partial z} e' - e \bar{z}' \frac{\partial^2 r(\bar{z})}{\partial z \partial z'} - \frac{\partial^2 r(\bar{z})}{\partial z \partial z'} \bar{z} e' + e(\frac{\partial r(\bar{z})}{\partial z})']}_{M_{12} = M_{21}} \\ & + \frac{1}{2} \underbrace{(\frac{\partial^2 r(\bar{z})}{\partial z \partial z'})}_{M_{22}}. \end{aligned}$$

The partial derivatives are evaluated at point $\bar{z}$. Partition M, we have

$$\begin{aligned} z' M z \quad & \equiv \quad \begin{pmatrix} x \\ u \end{pmatrix}' \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} \\ & = \quad \begin{pmatrix} x \\ u \end{pmatrix}' \begin{pmatrix} R & W \\ W' & Q \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} \\ & = \quad x' R x + u' Q u + 2u' W' x \end{aligned}$$

Exercise 38 *Transform the objective function in the stochastic growth model in (5.8) to the LQ form $z'Mz$, then decompose it to the form $x'Rx + u'Qu + 2u'W'x$.*

Step 3: Value function iteration. Now we transform the stochastic growth model into a LQ framework

$$\begin{aligned} & \max E_0 \sum_{t=0}^{\infty} \beta^t \{x_t' R x_t + u_t' Q u_t + 2u_t' W' x_t\} \\ & s.t. \\ x_{t+1} &= A x_t + B u_t + C w_{t+1} \end{aligned}$$

We can easily to show (leave as an exercise) that the optimal policy function takes the form

$$u_t = -(Q + \beta B' P B)^{-1} (\beta B' P A + W') x_t$$

and value function is

$$v(x) = x' P x + d$$

where P solves the algebraic matrix Riccati equation

$$P = R + \beta A' P A - (\beta A' P B + W)(Q + \beta B' P B)^{-1} (\beta B' P A + W'),$$

and

$$d = \beta (1 - \beta)^{-1} tr(PCC').$$

We can use either value function iteration or policy function iteration to solve this LQ problem.

Part II

Applications

Chapter 6

Economic Growth

Growth is a vast literature in macroeconomics, which seeks to explain some facts in the long-term behavior of economies. This chapter is an introduction to basic nonstochastic models of sustained economic growth[1]. It is divided in three sections. In the first section, we introduce the motivation for the theory: the empirical regularity which the theory seeks to explain. The second section is about exogenous growth models, i.e. models in which an exogenous change in the production technology results in income growth as a theoretical result. Finally, the third section introduces technological change as a decision variable, and hence the growth rate becomes endogenously determined, i.e., somehow the growth is chosen by the households in the economy.

6.1 Stylized Facts About Economic Growth

6.1.1 Kaldor's stylized facts

The first "five facts" refer to the long-run behavior of economic variables in an economy, whereas the sixth one involves an inter-country comparison.

1. The growth rate of output per worker ($y = \frac{Y}{L}$) $g_y = \frac{\triangle y}{y}$ is relatively constant over time.

2. The capital-labor ratio $\frac{K}{L}$ grows at a relatively constant rate.

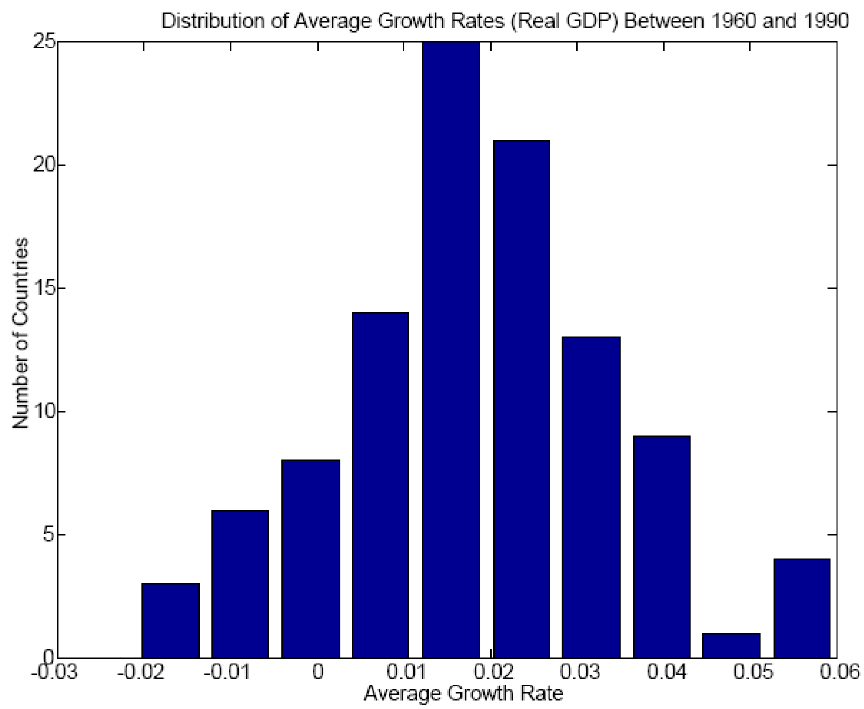
3. As a result of Facts 1 and 2, the capital-income ratio $\frac{K}{Y}$ is constant.

4. Real return to capital r is relatively constant over time.

5. As a result of Facts 3 and 4, capital and labor shares of income ($\alpha = \frac{rK}{Y}, 1 - \alpha = \frac{wL}{Y}$) are roughly constant over time.

[1]This chapter is based on Krusell's Lecture Notes for Macroeconomics I at Princeton University, Chapter 8, Dirk Kreuger's "Macroeconomic Theory" lecture notes, Chapter 9, and LS, Chapter 14.

6. Growth rates of output persistently vary across countries. Figure below shows the distribution of average yearly growth rates from 1960 to 1990. The majority of countries grew at average rates of between 1% and 3% (these are growth rates for real GDP per worker). Note that some countries posted average growth rates in excess of 6% (Singapore, Hong Kong, Japan, Taiwan, South Korea) whereas other countries actually shrunk, i.e., had negative growth rates (Venezuela, Nicaragua, Guyana, Zambia, Benin, Ghana, Mauretania, Madagascar, Mozambique, Malawi, Uganda, Mali). We will sometimes call the first group *growth miracles*, the second group *growth disasters*. Note that not only did the disasters' relative position worsen, but that these countries experienced absolute declines in living standards. The US, in terms of its growth experience in the last 30 years, was in the middle of the pack with a growth rate of real GDP per worker of 1.4% between 1960 and 1990.



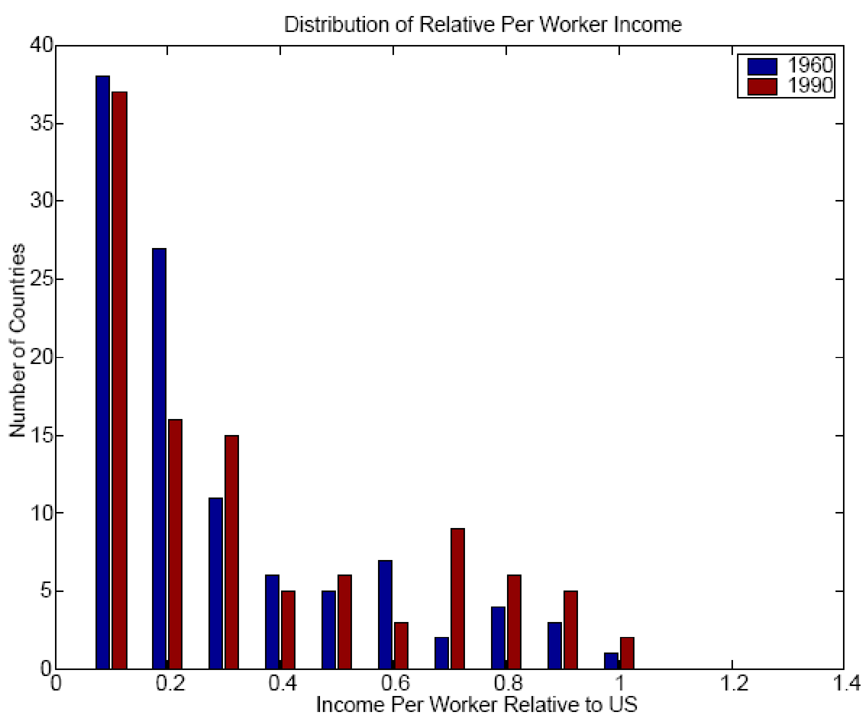
In particular, the first five facts imply that economies exhibit *Balanced Growth Path* (BGP) over time, i.e., the scale of the economy increases over time, but the composition of output does not. These stylized facts motivated the development of the neoclassical growth model, the Solow growth model, to be discussed below. The Solow model has spectacular success in explaining the stylized growth facts raised by Kaldor.

The sixth fact poses questions of what determines growth rate in different countries. Solow model does a poor job to explain the cross-country growth rate differences. This fact motivates so-called "*New Growth Theory*".

6.1.2 Other Facts

Besides these classical facts, there are also other empirical regularities which growth theory must account for. These are:

1. Output per worker $\frac{Y}{L}$ is very dipersed across countries. The poorest countries have about 5% of per capita GDP of US per capita GDP. This fact makes a statement about dispersion in income *levels* instaed of *growth rates*. When we look at Figure below, we see that out of the 104 countries in the data set, 37 in 1990 and 38 in 1960 had per worker incomes of less than 10% of the US level. The richest countries in 1990, in terms of per worker income, are Luxembourg, the US, Canada and Switzerland with over $30,000, the poorest countries, without exceptions, are in Africa. Mali, Uganda, Chad, Central African Republic, Burundi, Burkina Faso all have income per worker of less than $1000. Not only are most countries extremely poor compared to the US, but most of the world's population is poor relative to the US.



1. The distribution of $\frac{Y}{L}$ does not seem to spread out (although the variance has increased somewhat).

2. Countries change their *relative* position in the (international) income distribution. Growth disasters fall, growth miracles rise in the relative cross-country income distribution. A classical example of a growth disaster is Argentina. At the turn of the century Argentina had a per-worker income that was comparable to that in the US. In 1990 the per-worker income of Argentina was only on a level of one third of the US, due to a healthy growth experience of the US and a disastrous growth performance of Argentina. Countries that dramatically moved up in the relative income distribution include Italy, Spain, Hong Kong, Japan, Taiwan and South Korea, countries that moved down are New Zealand, Venezuela, Iran, Nicaragua, Peru and Trinidad&Tobago.

3. Countries with low incomes in 1960 did not show on average higher subsequent growth (this phenomenon is sometimes referred to as "no absolute (β) convergence").

4. There is "conditional convergence": Within groups classified by 1960 human capital measures (such as schooling), 1960 savings rates, and other indicators, a higher initial income y_0 (in 1960) was positively correlated with a lower growth rate g_y. This is studied by performing the "growth regression":

$$g_{y,i}^{1960-1990} = \alpha + \beta \log y_{0i} + \gamma \log edu_{0i} + \varepsilon_i, \ i = 1, ..., n.$$

   By controlling for the initial level of education, the growth rate was negatively correlated with initial income for the period 1960-1990: $\hat{\beta} < 0$. Whereas if the regression is performed without controlling for the level of education, the result for the period is $\hat{\beta} = 0$, i.e., no absolute convergence as mentioned above.

5. There is a positive correlation between the fraction of GDP spent on government goods and services and income levels.

6. Foreign trade volume seems to correlate positively with growth.

7. Demographic growth (fertility rate) is negatively correlated with income.

8. Growth in factor inputs (capital, labor) does not suffice in explaining output growth. The idea of an "explanation" of growth is due to Solow, who envisaged the method of "growth accounting". Based on a neoclassical production function

$$y = zF(K, L)$$

   the variable z captures the idea of technological change. If goods production is performed using a constant-returns-to-scale technology, operated under perfect competition, then (by an application of the Euler Theorem) it is possible to estimate how much out of total production growth is due to each production factor, and how much to the technological factor z. The empirical studies have shown that the contribution of z (the *Solow* *residual*) to output growth is very significant. In the US, the TFP growth can account for almost half the growth in output.

9. Workers tend to migrate into high-income countries.

6.2 Exogenous Growth Model

In this section we will study the basic framework to model output growth by introducing an exogenous change in the production technology that takes place over time. Mathematically, this is just a simple modification of the standard neoclassical growth model that we have studied before.

Two basic questions arise, one on the technique itself, and one on its reach. First, we may ask how complicated it will be to analyze the model. The answer is quite reassuring: it will be just a relatively easy transformation of material we have seen before. The second question is what is the power of this model: what types of technological change can be studied with these tools?

We will separate the issue of growth into two components. One is a technological component: is growth feasible with the assumed production technology? The second one is the decision making aspect involved: will a central planner choose a growing path? Which types of utility function allow for what we will call a "balanced growth path"?

We will first introduce famous Solow model.

6.2.1 Solow Growth Model

The Neoclassical production function is

$$Y_t = F(K_t, \gamma^t N_t)$$

where N_t is the total working hours. $\gamma > 1$ is the (gross) growth rate of labor-augumenting technological change. Therefore, $\gamma^t N_t$ represents the efficient units at time t. The production function satisfies all the standard assumptions (continuously differentiable, increasing in K and N, concave, CRS, Inada condition).

The condition of CRS implies that output can be written as

$$Y_t = F(K_t, \gamma^t N_t) = \gamma^t N_t F(K_t/\gamma^t N_t, 1)$$

Therefore, output per efficient unit $y_t \equiv Y_t/\gamma^t N_t$ is

$$y_t = F(K_t/\gamma^t N_t, 1)$$

The resource constraint in the economy is

$$C_t + I_t = F(K_t, \gamma^t N_t) \tag{6.1}$$

The capital accumulation law follows

$$K_{t+1} = (1-\delta)K_t + I_t \tag{6.2}$$

And investment is a fraction s (*exogeneously* given) of current output

$$I_t = sY_t, \;\; s \in (0,1) \tag{6.3}$$

6.2.2 Balanced Growth Path

Our question is: given this setting, is the sustained growth possible? Our object of study is to find a so-called *balanced growth path (BGP)*, in which all economic variables grow at constant rates (but they could be different). In this case, this would imply that for all t, the value of each variable in the model is given by

$$\begin{aligned} Y_t &= Y_0 g_Y^t \\ C_t &= C_0 g_C^t \\ K_t &= K_0 g_K^t \\ I_t &= I_0 g_I^t \\ N_t &= N_0 g_N^t \end{aligned}$$

Our task is to find the growth rate for each variable in a balanced growth path, and check whether such a path is consistent. We begin by guessing one of the growth rates as follows. From the capital accumulation law equation (6.2), divide both sides by K_t

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

Obviously $\frac{I_t}{K_t}$ =constant, $\forall t$, which implies both I_t and K_t grow at the same rate.[2] Hence we have

$$g_K = g_I$$

By the same type of reasoning, from the investment function (6.3), divide both sides by Y_t

$$\frac{I_t}{Y_t} = s$$

which implies

$$g_I = g_Y$$

Finally, from the resource constraint (6.1), divide both sides by I_t

$$\frac{C_t}{K_t} + 1 = \frac{Y_t}{I_t} = \frac{1}{s}$$

which implies

$$g_C = g_K$$

If we further assume the production function takes the form of Cobb-Douglas as following

$$Y_t = F(K_t, \gamma^t N_t) = K_t^\alpha (\gamma^t N_t)^{1-\alpha}$$

[2] If $I_t = cK_t$, where c is a constant, then we have $g_I \equiv \frac{I_{t+1}}{I_t} = \frac{cK_{t+1}}{cK_t} = \frac{K_{t+1}}{K_t} \equiv g_K$.

Taking into account the fact that the time endowment is bounded, actual working hours can not grow beyond a certain upper limit (usually normalized to 1), we must have $N_t = N_{t+1} = 1$ (i.e., $g_N = 1$) along BGP. We obtain

$$\begin{aligned}\frac{Y_{t+1}}{Y_t} &= g_Y = \frac{K_{t+1}^{\alpha}(\gamma^{t+1}N_{t+1})^{1-\alpha}}{K_t^{\alpha}(\gamma^t N_t)^{1-\alpha}} \\ &= (\frac{K_{t+1}}{K_t})^{\alpha}\gamma^{1-\alpha} \\ &= g_K^{\alpha}\gamma^{1-\alpha}\end{aligned}$$

Recall $g_Y = g_K = g$, substituting into the equation above, we have

$$g = \gamma$$

Therefore, along BGP, we have

$$g_Y = g_C = g_I = g_K = \gamma$$

And this growth path is technologically feasible.

6.2.3 Emprical Evaluation of Solow Model

Solow model implies along BGP:

1. $g_{(Y/N)} = g_Y = \gamma$, The growth rate of output per worker is constant over time.
2. Capital-efficent labor ratio $K_t/\gamma^t N_t$ is constant over time.
3. $g_Y = g_K = \gamma$ implies the capital-income ratio is constant.
4. Capital and labor shares of income are constant by construction. (C-D production function)
5. $r_t = (\frac{\gamma^t N_t}{K_t})^{1-\alpha}$. Therefore,

$$\begin{aligned}\frac{r_{t+1}}{r_t} &= \frac{(\gamma^{t+1}N_{t+1})^{1-\alpha}}{(\gamma^t N_t)^{1-\alpha}} \cdot (\frac{K_t}{K_{t+1}})^{1-\alpha} \\ &= \gamma^{1-\alpha} \cdot (1/\gamma)^{1-\alpha} = 1.\end{aligned}$$

   Real rates of return are constant. While the real wage $w_t = \frac{\partial F}{\partial N} = \gamma^{(1-\alpha)t}(\frac{K_t}{N_t})^{\alpha}$. The growth rate of wage is

$$\begin{aligned}\frac{w_{t+1}}{w_t} &= \gamma^{1-\alpha}(\frac{K_{t+1}}{K_t})^{\alpha} \\ &= \gamma^{1-\alpha}\gamma^{\alpha} \\ &= \gamma.\end{aligned}$$

   But the real wage for an efficient unit is constant.
6. In order to have cross-country difference in growth rate, we have to allow γ is different across countries.

6.2.4 Choosing Growth (Ramsey-Case-Koopmans Model)

The next issue to address is whether an individual who inhabits an economy in which there is some sort of exogenous technological progress, and in which the production technology is such that sustained growth is feasible, will choose a growing output path or not.

Initially, Solow overlooked this issue by assuming that the saving rate is exogenous and constant

$$I_t = sY_t.$$

From the text above, It is clear that such a rule can be consistent with a balanced growth path. Then the underlying premise is that the consumers' preferences are such that they choose a growing path for output. Clearly, not all types of preferences will work. In this section, We will restrict our attention to the usual time-separable preference relations. Hence the problem faced by a social planner will be of the form:

$$\begin{aligned} & \max \sum_{t=0}^{\infty} \beta^t u(C_t, N_t) \qquad (6.4) \\ & s.t. \\ C_t + I_t &= F(K_t, \gamma^t N_t) \\ K_{t+1} &= (1-\delta)K_t + I_t \\ & K_0 \text{ given} \end{aligned}$$

In order to allow HHs choose optimally the BGP, we do have a restriction on the preferences in the following theorem.

Theorem 39 *Balanced Growth Path is possible as a solution to the social planner's problem (6.4) if and only if*

$$\begin{aligned} u(C,N) &= \frac{C^{1-\sigma}v(1-N)-1}{1-\sigma}, \text{ if } 0<\sigma<1, \text{ and } \sigma>1 \qquad (6.5) \\ u(C,N) &= \log C + v(1-N), \text{ if } \sigma = 1 \end{aligned}$$

where the time endowment is normailized to one and $v(\cdot)$ is a function with leisure as an argument.

The proof can be seen in King, Plosser, and Rebelo (1988). Intuitively, to be consistent with BGP, we have to impose two restrictions on preferences: (1) the intertemporal elasticity of substitution (IES) in consumption must be invariant to the scale of consumption; (2) the income and substitution effects associated with sustained growth in labor productivity must offset each other.

This kind of social planner's problem will have an Euler equation looks like

$$g(\frac{C_{t+1}}{C_t}) = \beta(1+r)$$

Therefore, since along BGP, real interest rate r is constant, so does the right hand side of the equation above. Since consumption is growing at a constant rate, the IES, which is a power parameter in function $g(\cdot)$, must be constant and independent of the level of consumption.

The second condition is required bacause hours worked cannot grow ($g_N = 1$) along BGP. To reconcile this with a growing marginal product of labor - induced by labor-augumenting technological change - income and substitution effects of productivity growth must have exactly offsetting effects on labor supply (N).

Note that $v(1-N) = 1$ (i.e., $u(c) = \frac{c^{1-\sigma}-1}{1-\sigma}$) fits the theorem assumptions. Hence non-valued leisure is also consistent with balanced growth path.

The existence of BGP does also put restriction on the way we model technological change. In order for the model to have a BGP with constant growth rates, the technological progress must take the labor-augumenting form.

Theorem 40 *In neoclassical growth model, only labor-augumenting technological change is consistent with the existence of a BGP.*

Proof. Assume a neoclassical production function that includes both the labor-augumenting and capital-augumenting technological change

$$Y_t = F(q^t K_t, \gamma^t N_t)$$

Notice if $q = \gamma$, then the technological change is Hicks neutral.

Due to CRS of the neoclassical production function, we have

$$\begin{aligned} Y_t &= q^t K_t F(1, \gamma^t N_t / q^t K_t) \\ &= q^t K_t f((\frac{\gamma}{q})^t \frac{N_t}{K_t}) \end{aligned}$$

where $f((\frac{\gamma}{q})^t \frac{N_t}{K_t}) \equiv F(1, \gamma^t N_t / q^t K_t)$. Thus

$$\frac{Y_t}{K_t} = q^t f((\frac{\gamma}{q})^t \frac{N_t}{K_t}) \tag{6.6}$$

We know along BGP, Y and K grow at the same rate, which implies $\frac{Y}{K}$ is constant. Besides, the third of Kaldor's stylized facts says that in the data, $\frac{Y}{K}$ is pretty constant. But to make the RHS of equation (6.6) constant, there are only two ways:

1. $q = 1$ and $g_K = \gamma$, which is the labor-augumenting technological change case we already showed in the previous subsection.

2. $q > 1$ (i.e., we have capital-augumenting technological change) and the term $f((\frac{\gamma}{q})^t \frac{N_t}{K_t})$ exactly offsets the growth term q^t. You can show that only Cobb-Douglas production function can allow this case happen, i.e., the production function now takes the form

$$\begin{aligned} Y_t &= (q^t K_t)^\alpha (\gamma^t N_t)^{1-\alpha} = K_t^\alpha N_t^{1-\alpha} q^{t\alpha} \gamma^{t(1-\alpha)} \\ &= K_t^\alpha [\chi^t N_t]^{1-\alpha} \end{aligned}$$

where $\chi = q^{t\alpha/t(1-\alpha)}\gamma$. In other words, we can always express technological change as purely labor-augumenting as at the rate χ.

■

6.2.5 A Little Digress on Utility Function

The utility function as in (6.5) is widely used in recursive macroeconomics. It is worth spending some time to get familiar with it.

This utility function has several nice properties.

1. Homotheticity. Define the marginal rate of substitution (MRS) between consumption at any two dates t and $t+s$ as

$$MRS_{t,t+s} = \frac{\partial u(c)/\partial c_{t+s}}{\partial u(c)/\partial c_t}.$$

Then a function u is said to be homothetic if $MRS(c_t, c_{t+s}) = MRS(\lambda c_t, \lambda c_{t+s})$ for all $\lambda > 0$ and c. It is easy to verify that for our utility function (6.5) satisfies homotheticity.

$$\begin{aligned} MRS(c_t, c_{t+s}) &= \frac{\beta^{t+s}\sigma c_{t+s}^{-\sigma} v(1-N_{t+s})}{\beta^t \sigma c_t^{-\sigma} v(1-N_t)} \\ &= \frac{\beta^{t+s}\sigma c_{t+s}^{-\sigma} v(1-N)}{\beta^t \sigma c_t^{-\sigma} v(1-N)} \\ &= \beta^s (\frac{c_{t+s}}{c_t})^{-\sigma} \\ &= \beta^s (\frac{\lambda c_{t+s}}{\lambda c_t})^{-\sigma} \\ &= MRS(\lambda c_t, \lambda c_{t+s}) \end{aligned}$$

Homotheticity means that consumption allocations are independent of the units of measurement employed.

Another definition of homotheticity is that if $u(x_1) = u(x_2) \Longleftrightarrow u(\lambda x_1) = u(\lambda x_2)$. It is easy to show that this definition is equivalent to the one using MRS. Basing on this definition, we have following claim.

Claim 41 *If utility function u is homogeneous of degree η, $\forall \eta > 0$, then it is also homothetic. But if u is homothetic, it may not be homogeneous.*

For example, consider a homogeneous of degree η utility function such that $u(\lambda x) = \lambda^\eta u(x)$. Then if $u(x_1) = u(x_2) \Longleftrightarrow u(\lambda x_1) = \lambda^\eta u(x_1) = \lambda^\eta u(x_2) = u(\lambda x_2)$. u is homothetic. It is easy to show our utility function here as in (6.5) is homogeneous of degree $1-\sigma$, therefore it is also homothetic.

Now consider utility function $u(x, y) = a \log x + b \log y$. It is easy to show that it is homothetic, but it is not homogeneous.

2. Define Arrow-Pratt Relative Risk Aversion Coefficient as $-\frac{u''(c)c}{u'(c)}$. Then it is easy to verify that the utility function above has constant relative risk aversion (CRRA) coefficient which is σ. Since now, we will call this family of utility function CRRA utility function.

3. The Intertemporal Elasticity of Substitution (IES) $\varepsilon(c_t, c_{t+1})$ measures by how many percentage the demand for consumption in period $t+1$, relative to demand for consumption in period t, $\frac{c_{t+1}}{c_t}$, declines as the relative price of consumption in $t+1$ to consumption in period t, $q_t = \frac{1}{1+r_t}$ changes by one percent. i.e.,

$$\varepsilon(c_t, c_{t+1}) = -\frac{\frac{d(c_{t+1}/c_t)}{(c_{t+1}/c_t)}}{\frac{d(1/(1+r_t))}{1/(1+r_t)}} = -\frac{d(\ln(c_{t+1}/c_t))}{d(\ln(1/(1+r_t)))}$$

For the optimal growth model with CRRA utility function, we will end up with an Euler equation as following

$$\frac{\beta u'(c_{t+1})}{u'(c_t)} = \frac{p_{t+1}}{p_t} = \frac{1}{1+r_{t+1}}$$

$\Rightarrow$

$$\beta(\frac{c_{t+1}}{c_t})^{-\sigma}(\frac{v(1-N_{t+1})}{v(1-N_t)}) = \frac{1}{1+r_{t+1}}$$

Take log on both sides, we have

$$-\sigma \ln(\frac{c_{t+1}}{c_t}) = \ln(\frac{1}{1+r_{t+1}}) + \text{other terms}$$

Therefore, for CRRA utility function, IES is $1/\sigma$, the inverse of the coefficient of risk aversion. When σ is higher, HHs are more risk averse, therefore, they are more reluctant to shift consumption intertemporally, IES thus is lower.

6.2.6 Solving the Balanced Growth Path

We now describe the steps of solving the Ramsey-Cass-Koopmans model for a balanced growth path.

1. Assume that preferences are represented by the CRRA utility function

$$u(C, N) = \frac{C^{1-\sigma}v(1-N) - 1}{1-\sigma}.$$

2. Transform the current model into a stationary one.

3. Take first order conditions of the tranformed planner's problem.

Now we will show step by step how to solve this model. Recall the original problem is

$$
\begin{aligned}
& \max \sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\sigma} v(1-N_t) - 1}{1-\sigma} \qquad (6.7) \\
& s.t. \\
C_t + I_t &= \gamma^t N_t F(\frac{K_t}{\gamma^t N_t}, 1) \\
K_{t+1} &= (1-\delta)K_t + I_t \\
& K_0 \text{ given}
\end{aligned}
$$

We know that the balanced growth path solution to this growth Model (6.4) has all variables growing at rate γ except for labor. We define transformed variables by dividing each original variable by its growth rate:

$$
c_t \equiv \frac{C_t}{\gamma^t}, i_t \equiv \frac{I_t}{\gamma^t}, k_t \equiv \frac{K_t}{\gamma^t}. \qquad (6.8)
$$

Thus we transform the original problem (6.7) into a transformed model

$$
\begin{aligned}
& \max \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\sigma} \gamma^{t(1-\sigma)} v(1-N_t) - 1}{1-\sigma} \\
& s.t. \\
(c_t + i_t)\gamma^t &= \gamma^t N_t F(\frac{k_t \gamma^t}{\gamma^t N_t}, 1) \\
k_{t+1}\gamma^{t+1} &= [i_t + (1-\delta)k_t]\gamma^t \\
& K_0 \text{ given}
\end{aligned}
$$

Cleaning it up, we obtain

$$
\begin{aligned}
& \max \sum_{t=0}^{\infty t} \hat{\beta}^t \frac{c_t^{1-\sigma} v(1-N_t) - 1}{1-\sigma} \\
& s.t. \\
c_t + i_t &= F(k_t, N_t) \\
k_{t+1}\gamma &= i_t + (1-\delta)k_t \\
& K_0 \text{ given}
\end{aligned}
$$

where $\hat{\beta} = \beta\gamma^{(1-\sigma)}$. In order to guarantee finiteness of life-time utility, we require $\hat{\beta} < 1$. Recall that $\gamma > 1$ and $0 < \beta < 1$, thus we have:

Case 1: If $\sigma > 1$, $\gamma^{(1-\sigma)} < 1$, so $\beta\gamma^{(1-\sigma)} < 1$ holds.

Case 2: If $\sigma = 1$, $\hat{\beta} = \beta < 1$ also holds.

Case 3: If $0 < \sigma < 1$, then for some parameter values of γ and β, we may run into an ill-defined problem.

Therefore, as long as $\sigma \geq 1$, we will be fine.

Now we are back to the standard neoclassical growth model without trend that we have been dealing with before (for example in the "Foundation" part). The only differences are that there is a γ factor in the capital accumulation equation, and the discount factor is modified.

Next, we can set up Lagrangian for this transformed problem

$$\begin{aligned}\mathcal{L} &= \sum_{t=0}^{\infty t} \hat{\beta}^t \frac{c_t^{1-\sigma} v(1-N_t) - 1}{1-\sigma} \\ &+ \sum_{t=0}^{\infty t} \mu_t [F(k_t, N_t) - c_t - \gamma k_{t+1} + (1-\delta) k_t].\end{aligned}$$

FOCs are

$$c_t \quad : \quad \hat{\beta}^t c_t^{-\sigma} v(1-N_t) - \mu_t = 0 \tag{6.9}$$

$$N_t \quad : \quad -\hat{\beta}^t \frac{c_t^{1-\sigma} v'(1-N_t)}{1-\sigma} + \mu_t \frac{\partial F}{\partial N_t} = 0 \tag{6.10}$$

$$k_{t+1} \quad : \quad \mu_{t+1} [\frac{\partial F}{\partial k_{t+1}} + (1-\delta)] - \mu_t \gamma = 0 \tag{6.11}$$

$$\mu_t \quad : \quad F(k_t, N_t) - c_t - \gamma k_{t+1} + (1-\delta) k_t = 0 \tag{6.12}$$

Combining (6.9) and (6.11), we have the usual Euler equation

$$\gamma c_t^{-\sigma} v(1-N_t) = \hat{\beta} c_{t+1}^{-\sigma} v(1-N_{t+1}) [\frac{\partial F}{\partial k_{t+1}} + (1-\delta)]. \tag{6.13}$$

Combining (6.9) and (6.10), we obtain

$$\frac{\partial F}{\partial N_t} = \frac{1}{(1-\sigma)} \frac{c_t v'(1-N_t)}{v(1-N_t)} \tag{6.14}$$

It says the marginal product of labor, which is also the real wage, should be equal to the MRS between consumption and leisure.

We also have the following transversality condition (TC)

$$\lim_{t\to\infty} \mu_t k_{t+1} = \lim_{t\to\infty} \hat{\beta}^t c_t^{-\sigma} v(1-N_t) k_{t+1} = 0. \tag{6.15}$$

Optimal quantities for this economy are sequences of consumption $\{c_t\}_{t=0}^{\infty}$, working hours $\{N_t\}_{t=0}^{\infty}$, and capital stock $\{k_t\}_{t=0}^{\infty}$ that satisfy the FOCs (6.12)-(6.14) and the Transversality Condition (6.15). Under our assumptions about preferences and production possibilities, conditions (6.12)-(6.15) are necessary and sufficient for an optimum.

Note that FOCs (6.12)-(6.14) can be further reduced to a non-linear second-order equation in k. With two boundary conditions: initial boundary condition $k_0 > 0$ and the terminal boundary condition as in TC (6.15), we can solve the sequence of $\{k_t\}_{t=0}^{\infty}$ and hence other quantities.

Finally, once we solve completely this transformed model, we can transform back to the original model easily by reversing the transformation as in (6.8).

In conclusion, with the stated assumptions on preferences and on technology, the current model converges to a balanced growth path, in which all variables grow at rate γ. This rate is *exogenously* determined; it is a parameter in the model. That is the reason why it is called "exogenous" growth model.

6.3 Endogenous Growth Model

The exogenous growth framework analyzed before (Solow and Ramsey model) has a serious shortfall: growth is not truly a result in such model, it is an assumption. However, we have reasons (from data) to suspect that growth must be a rather more complex phenomenon than this long term productivity shift in γ that we have treated as somehow intrinsic to economic activity. In particular, as we showed before, output growth rates have been very different across countries for long periods; trying to explain this fact as merely the result of different γ's is not a very insightful and convincing approach. We would prefer our model to produce γ as a result. Therefore, we look for endogenous growth models.

The key assumption driving the result, that is, absent technological progress the economy will converge to a no-growth steady state, is the assumption of diminishing marginal product to the production factor that is accumulated, namely capital. As economies grow they accumulate more and more capital, which, with decreasing marginal products, yields lower and lower returns. Absent technological progress forces the economy to the steady state. Hence the key to derive sustained growth without assuming it being created by exogenous technological progress is to pose production technologies in which marginal products to accumulable factors are not driven down as the economy accumulates these factors.

We will start our discussion of these models with a stylized version of the so called AK model, then turn to models with externalities as in Romer (1986) and Lucas (1988), and finally look at Rebelo's (1991) two-sector model of endogenous growth.

6.3.1 The AK Model

Let us recall the usual assumptions on the neoclassical production technology in the exogenous growth model

$$\begin{aligned}
F(0,n) &= F(k,0)=0, F_k(k,n)>0, F_n(k,n)>0, \forall k,n>0 \\
\lambda F(k,n) &= f(\lambda k, \lambda n) \quad \text{(CRS)} \\
\lim_{k\to 0} F_k(k,n) &= \infty, \lim_{k\to\infty} F_k(k,n)=0, \quad \text{(Inada Condition)}
\end{aligned}$$

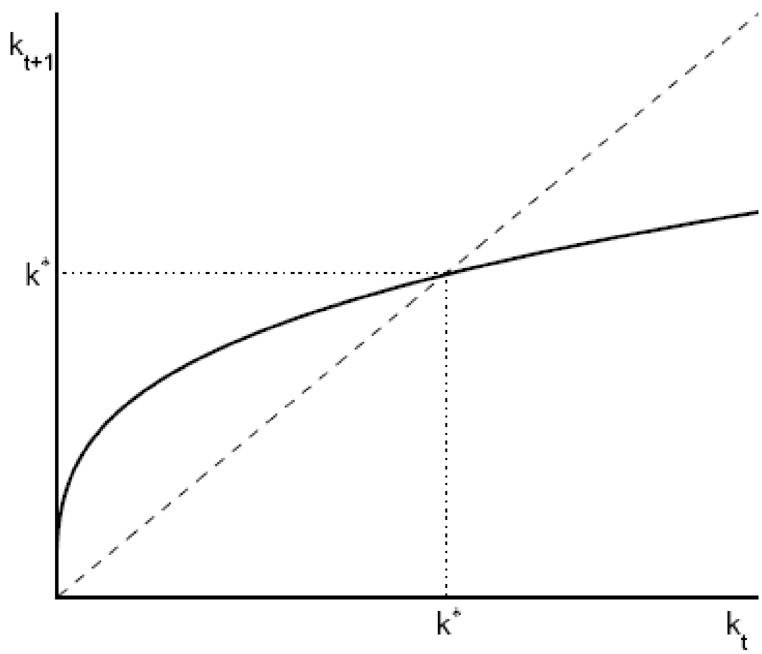
In Solow model, we have

$$\begin{aligned} k_{t+1} &= i_t + (1-\delta)k_t \\ &= sF(k_t, n_t) + (1-\delta)k_t \end{aligned}$$

Therefore we have

$$\begin{aligned} \frac{\partial k_{t+1}}{\partial k_t} &= sF_k(k,n) + (1-\delta) > 0 \\ \frac{\partial^2 k_{t+1}}{\partial k_t^2} &= sF_{kk}(k,n) < 0 \end{aligned}$$

The Figure below shows the dynamics of capital accumulation in Solow and Solow-alike model.



Capital accumulation in Solow model

where k^* is the steady state (SS) of the model in which

$$k_{t+1} = k_t = k^*, \forall t$$

We can solve k^* below from the following equation

$$sF(k^*, 1) = \delta k^*$$

In Solow model, without any exogenously given technical change, the economy will eventually converge to the SS where the growth rate is zero. Long-run growth thus is not feasible. In order to allow long-run growth, we need the introduce at least some change to the production function: We must dispose of the assumption that $\lim_{k\to\infty} F_k(k,n) = 0$. What we basically want is that F does not cross the 45^o line. i.e., we want to escape from diminishing returns to scale.

The simplest way is to assume the following production technology

$$Y = F(K) = AK, A > 1$$

Under this technology, we have

$$g_Y \equiv \frac{Y_{t+1}}{Y_t} = \frac{AK_{t+1}}{AK_t} = g_K$$

Divide the both sides of capital accumulation equation by K_t, we obtain

$$g_K \equiv \frac{K_{t+1}}{K_t} = \frac{I_t + (1-\delta)K_t}{K_t} = \frac{I_t}{K_t} + (1-\delta) \tag{6.16}$$

which implies

$$g_K = g_I$$

Similarly from resource constraint, we know

$$g_C = g_Y$$

Therefore, BGP is feasible in this model with

$$g_Y = g_K = g_C = g_I = g$$

Now let's follow Solow model to assume

$$I = sY = sAK$$

Substituting into (6.16), we have

$$g_K = sA + 1 - \delta$$

Therefore, as long as $A > \delta/s$, we will have $g_K = g > 1$. We achieve perpetual growth even without exogenous technological change. The reason is in AK model, we have $\lim_{k\to\infty} F_k(k) = A > 0$.

The next question is whether the consumer will choose this balanced growth, and if so, how fast (i.e, what is g?). We will answer this question assuming a CRRA utility function (needed for BGP, see section 6.2.4) with non-valued leisure. The planner's problem then is:

$$\begin{aligned} & \max_{\{c_t,k_{t+1}\}_{t=0}^{\infty}} \{\sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\sigma}}{1-\sigma}\} \\ & s.t. \\ c_t + k_{t+1} - (1-\delta)k_t &= Ak_t \end{aligned}$$

The Euler equation is

$$c_t^{-\sigma} = \beta c_{t+1}^{-\sigma}[A + 1 - \delta]$$

$\Rightarrow$

$$\begin{aligned} (\frac{c_{t+1}}{c_t})^{\sigma} &= \beta[A + 1 - \delta] \\ g &= g_c \equiv \frac{c_{t+1}}{c_t} = [\beta(A + 1 - \delta)]^{1/\sigma} \end{aligned}$$

As long as

$$[\beta(A + 1 - \delta)]^{1/\sigma} > 1, \tag{6.17}$$

we achieve the long-run growth. Since $\sigma > 0$, this requirement is also equivalent to the following condition

$$\beta(A + 1 - \delta) > 1.$$

The growth rate of consumption thus is a function of all the parameters in the utility function and the production function. Notice that this implies that the growth rate is constant as from $t = 0$. There are no transitional dynamics in this model. The economy is in the BGP from the beginning.

There is another thing we want to make sure: is utility bounded? Note that

$$c_t = c_0 \frac{c_1}{c_0} \frac{c_2}{c_1} \cdots \frac{c_t}{c_{t-1}} = c_0 g_c^t$$

Therefore

$$\begin{aligned} U &= \{\sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\sigma}}{1-\sigma}\} \\ &= \sum_{t=0}^{\infty} \left[\beta \left[[\beta(A + 1 - \delta)]^{1/\sigma}\right]^{1-\sigma}\right]^t \frac{c_0^{1-\sigma}}{1-\sigma}. \end{aligned}$$

So the sufficient condition for boundness is

$$\beta \left[[\beta(A + 1 - \delta)]^{1/\sigma}\right]^{1-\sigma} < 1. \tag{6.18}$$

The two conditions (6.17) and (6.18) must simultaneously hold for us to obtain a balanced growth path.

Notice that in the Solow and Ramsey model the growth rate of the economy was given by $g_Y = g_C = g_I = g_K = \gamma$, the growth rate of technological progress. In particular, savings rates (s), depreciation and the subjective time discount rate (δ, β)affect per capita income levels, but not growth rates. In contrast, in the basic AK model the growth rate of the economy is affected positively by the parameter governing the productivity of capital, A and negatively by parameters reducing the willingness to save, namely the depreciation rate δ and the degree of impatience $1/\beta$. Any policy affecting these parameters in the Solow or Ramsey model has only level, but no growth rate effects. However, it has growth rate effects in the AK model. Hence the former models are sometimes referred to as "*income level models*" whereas the others are referred to as "*growth rate models*".

In AK model, The growth has become a function of underlying parameters in the economy. And we do not have convergence. Could the dispersion in cross-country growth rates be explained by differences in these parameters? Country i's Euler Equation would be:

$$g_i = (\frac{c_{t+1}}{c_t})_i = [\beta_i(A_i + 1 - \delta_i)]^{1/\sigma}$$

Theoretically speaking, we have

$$\frac{\partial g_i}{\partial \beta_i} > 0, \frac{\partial g_i}{\partial A_i} > 0, \frac{\partial g_i}{\partial \delta_i} < 0$$

i.e., a country with higher impatience, lower TFP level, and higher depreciation rate should experience lower output growth rate. But the problem with the AK model is that, if parameters are calibrated to mimic the data's dispersion in growth rates, the simulation results in too much divergence in output level. The dispersion in 1960-1990 growth rates would result in a difference in output levels wider than the actual data.

6.3.2 Romer's Externality Model

The main assumption generating sustained growth in the last subsection was the presence of constant returns to scale with respect to production factors that are, in contrast to raw labor, accumulable. Otherwise eventually decreasing marginal products set in and bring the growth process to a halt. One obvious unsatisfactory element of the previous model was that labor was not needed for production and that therefore the capital share equals one. Even if one interprets capital broadly as including human capital, this assumption may be rather unrealistic. We are facing the following dilemma: on the one hand we want constant returns to scale to accumulable factors, on the other hand we want labor to claim a share of income, on the third hand we can't deal with increasing returns to scale on the firm level as this destroys existence of competitive equilibrium. (At least) two ways out of this problem have been proposed: a) there may be increasing returns to scale on the firm level, but the firm does not perceive it this way because part of its inputs come from positive externalities beyond the control of the firm; b) a departure from perfect competition towards monopolistic competition. We will discuss the main contributions in the first proposed resolution here.

The intellectual precedent to this model is Arrow (1962). The basic idea is that there are externalities to capital accumulation, so that individual savers do not realize the full return on their investment. Each individual firm operates the following production function

$$F(K, N, \bar{K}) = AK^{\alpha}N^{1-\alpha}\bar{K}^{\rho}$$

where K and N are capital and labor inputs for the firm, and $\bar{K}$ is the average capital stock in the economy. Individual firms view this average capital stock as given, although their input behavior will indeed affect $\bar{K}$. This is where the externality comes from. We assume that $\rho = 1 - \alpha$, so that in fact a social planner faces an AK model decision problem. Note that if we assumed that $\alpha + \rho > 1$, then balanced growth path would not be possible.

From the firm's problem, we know

$$w_t = (1-\alpha)AK^{\alpha}N^{-\alpha}\bar{K}^{1-\alpha}$$

Let us assume that leisure is not valued and normalize the labor endowment N_t to one in every t. Assume that there is a measure one of representative firms (therefore in equilibrium, $K = \bar{K}$), so that the equilibrium wage must satisfy

$$\begin{aligned} w_t &= (1-\alpha)AK_t^{\alpha}\bar{K}_t^{1-\alpha} \\ &= (1-\alpha)A\bar{K}_t \end{aligned}$$

and rental rate of capital is

$$r_t = \alpha A$$

The consumer's decentralized EE is (assume CRRA utility)

$$\begin{aligned} g_c^{CE} &\equiv \frac{c_{t+1}}{c_t} = [\beta(r_{t+1} + 1 - \delta)]^{1/\sigma} \\ &= [\beta(\alpha A + 1 - \delta)]^{1/\sigma} \end{aligned}$$

As long as

$$[\beta(\alpha A + 1 - \delta)]^{1/\sigma} > 1$$

We obtain endogenous long-run growth in decentralized CE.

How about the social planner's problem? Since the planner can internalize externalities, therefore, from social planner's viewpoint, the production technology is (normailze the labor input to be one)

$$\begin{aligned} F(K, N) &= AKN^{1-\alpha} \\ &= AK \end{aligned}$$

The planner's problem is

$$\begin{aligned} & \max \sum_{t=0}^{\infty} \beta^t u(c_t) \\ & s.t. \\ c_t + i_t &= AK_t \\ K_{t+1} &= (1-\delta)K_t + i_t \end{aligned}$$

From the AK model above, we already now that the growth rate of consumption for this problem is

$$g_c^{SP} = [\beta(A + 1 - \delta)]^{1/\sigma}.$$

Obviously $g_c^{SP} > g_c^{CE}$. This is due to the fact that competitive firms do not internalize the productivity-enhancing effect of higher average capital and hence *underinvest* capital, compared to the social optimum. Put it in another way, the private returns to investment (saving) are too low, giving rise to underinvestment and slow capital accumulation. Compared to the competitive equilibrium, the planner chooses lower period zero consumption and higher investment, which generates a higher growth rate. Obviously welfare is higher in the socially optimal allocation than under the competitive equilibrium allocation (since the planner can always choose the competitive equilibrium allocation, but does not find it optimal in general to do so).

Romer's externality model overcomes the shortfall of the AK model: labor is irrelevant. But the model still leads to a large divergence in output levels just as AK model.

6.3.3 Lucas' Endogenous Growth Model with Human Capital

We showed above that it was possible to generate long-run output growth without exogenous technological change if the returns to capital were constant asymptotically. Romer use the externality to escape the diminishing returns to scale, while Lucas (1988) use human capital to achieve the goal.

In the Lucas' model, plain labor in the production function is replaced by human capital. Human capital can be accumulated, so the technology does not run into decreasing marginal returns. Here we first present a simplified one-sector version of the Lucas' original two-sector model called $A(K, H)$ model, then we present Lucas' two-sector model.

$A(K, H)$ Model

The production function is Cobb-Douglas in physical capital K and human capital H[3]

$$F(K, H) = AK^{\alpha}H^{1-\alpha}. \tag{6.19}$$

The laws of motion for two types of capital are

$$\begin{aligned} K_{t+1} &= (1-\delta_K)K_t + I_{K,t} \\ H_{t+1} &= (1-\delta_H)H_t + I_{H,t}. \end{aligned}$$

The resource constraint in the economy is

$$c_t + I_{H,t} + I_{K,t} = Y_t = AK_t^{\alpha}H_t^{1-\alpha}$$

[3] That's the reason it is called $A(K, H)$ model.

Assuming a standard CRRA utility function, the FOCs in the planner's problem are

$$\begin{aligned} c_t &: \beta^t c_t^{-\sigma} = \mu_t \\ K_{t+1} &: \mu_t = \mu_{t+1}[F_K(K_{t+1}, H_{t+1}) + 1 - \delta_K] \\ N_{t+1} &: \mu_t = \mu_{t+1}[F_N(K_{t+1}, H_{t+1}) + 1 - \delta_N] \end{aligned}$$

which lead to two EEs

$$\begin{aligned} \frac{c_{t+1}}{c_t} &= \left(\beta\left[F_K(\frac{K_{t+1}}{H_{t+1}}, 1) + 1 - \delta_K\right]\right)^{\frac{1}{\sigma}} \\ \frac{c_{t+1}}{c_t} &= \left(\beta\left[F_N(\frac{K_{t+1}}{H_{t+1}}, 1) + 1 - \delta_N\right]\right)^{\frac{1}{\sigma}} \end{aligned}$$

Since along BGP, $g_c = \frac{c_{t+1}}{c_t}$, this requires that $\frac{K_{t+1}}{H_{t+1}}$ =constant $\forall t$, and we have

$$F_K(\frac{K_{t+1}}{H_{t+1}}, 1) + 1 - \delta_K = F_N(\frac{K_{t+1}}{H_{t+1}}, 1) + 1 - \delta_N \tag{6.20}$$

Basically this is just a non-arbitrage condition. It says the return on both capital must be equal in equilibrium.

If we further assume $\delta_H = \delta_K = \delta$, this non-arbitrage condition leads to

$$\frac{K_t}{H_t} = \frac{\alpha}{1-\alpha} \tag{6.21}$$

And we have

$$g_c = \left(\beta\left[A\alpha^\alpha(1-\alpha)^{1-\alpha} + 1 - \delta\right]\right)^{\frac{1}{\sigma}}$$

Since $\frac{K_{t+1}}{H_{t+1}}$=constant, we know $g_K = g_H$. Then it is easy to show that along BGP, we will have

$$g_Y = g_K = g_H = g_c = g_{I_k} = g_{I_H} = g$$

If $\left(\beta\left[A\alpha^\alpha(1-\alpha)^{1-\alpha} + 1 - \delta\right]\right)^{\frac{1}{\sigma}} > 1$, we obtain long-run economic growth in this model.

If we substitute equation (6.21) into production function (6.19), we have

$$\begin{aligned} Y &= AK(\frac{1-\alpha}{\alpha})^{1-\alpha} \\ &= BK \end{aligned}$$

where $B = A(\frac{1-\alpha}{\alpha})^{1-\alpha}$ is a constant. Thus, this model is equivalent to the AK model in essence. Besides, since there does not exist distortion such as externalities, in this economy, FWT and SWT hold, the growth rates for CE and SP are the same.

Uzawa-Lucas Two-Sector Model

Consider the following social palnner's problem:

$$
\begin{aligned}
& & \max \sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\sigma}}{1-\sigma} \\
& & s.t. \\
C_t + I_{K,t} & \leq & AK_t^{\alpha}(u_t H_t)^{1-\alpha} \\
I_{H,t} & \leq & B(1-u_t)H_t \\
K_{t+1} & = & (1-\delta)K_t + I_{K,t} \\
H_{t+1} & = & (1-\delta)H_t + I_{H,t} \\
& & K_0, H_0 \text{ given}
\end{aligned}
$$

There are two sectors in the economy. One sector (called C firm) uses capital and human to produce consumption and capital investment good, another sector (called H firm) uses only human capital to produce human capital investment. $u_t \in [0,1]$ is the fraction of human capital used in the consumption good sector.

Along the BGP, we have to obtain that $u_t = u^*, \forall t$. Thus from the law of motion of human capital, we have

$$g_H \equiv \frac{H_{t+1}}{H_t} = B(1-u^*) + 1 - \delta.$$

Unlike the previous one-sector $A(K,H)$ model, now in this model, there is no non-arbitrage condition that guarantees $\frac{K_t}{H_t}$=constant, hence K and H can grow at different rate g_K and g_H. Given this, it is easy to show that consumption and output also grow at different rates, and we have

$$g_Y = g_K^{\alpha} g_H^{1-\alpha}.$$

Set up the Lagrangian for the planner's problem:

$$
\begin{aligned}
\mathcal{L} &= \sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\sigma}}{1-\sigma} + \lambda_t \{AK_t^{\alpha}(u_t H_t)^{1-\alpha} - C_t - K_{t+1} + (1-\delta)K_t\} \\
& \quad + \mu_t \{B(1-u_t)H_t - H_{t+1} + (1-\delta)H_t\}
\end{aligned}
$$

FOCs are

$$
\begin{aligned}
C_t &: \beta^t C_t^{-\sigma} = \lambda_t \\
u_t &: \lambda_t (1-\alpha) K_t^{\alpha} H_t^{1-\alpha} u_t^{-\alpha} = \mu_t B H_t \\
K_{t+1} &: \lambda_t = \lambda_{t+1}(A\alpha K_{t+1}^{\alpha-1}(u_{t+1}H_{t+1})^{1-\alpha} + 1 - \delta) \\
H_{t+1} &: \mu_t = \mu_{t+1}(B(1-u_{t+1}) + 1 - \delta)
\end{aligned}
$$

Combining the FOCs, we have the EE for capital

$$(\frac{C_{t+1}}{C_t})^{\sigma} = \beta \frac{\lambda_t}{\lambda_{t+1}} = \beta \left(A\alpha K_{t+1}^{\alpha-1}(u_{t+1}H_{t+1})^{1-\alpha} + 1 - \delta \right)$$

Notice that $\frac{C_{t+1}}{C_t} \equiv g_C$ is constant. Hence to guarantee the equation above to hold on BGP, we should have $\frac{K_{t+1}}{H_{t+1}}$=constant. Hence along BGP

$$g_Y = g_K = g_H = g_C = g_{I_k} = g_{I_H} = g$$

As long as

$$g = B(1-u^*) + 1 - \delta > 1,$$

long-run endogeneous growth is achieved in this economy.

6.3.4 Rebelo's Two-Sector Endogenous Growth Model

In Lucas' model above, both production factors (K and H) are reproducible. But actually it is not necessary that all factors be producible in order to experience sustained growth through factor accumulation in the neoclassical framework. Instead, Rebelo (1991) shows that the critical requirement for perpetual growth is the existence of a "core" of capital goods that is produced with constant returns technologies and without the direct or indirect use of nonreproducible factors. Here we will study the simplest version of his model with a single capital good that is reproduced without any input of the economy's constant labor endowment.

The technologies follow

$$c_t = L^{1-\alpha}(\phi_t K_t)^\alpha \tag{6.22}$$

$$I_t = A(1-\phi_t)K_t \tag{6.23}$$

$$K_{t+1} = (1-\delta)K_t + I_t \tag{6.24}$$

where $\phi_t \in [0,1]$ is the fraction of capital used in the consumption good sector. The consumption goods sector uses labor (nonreproducible factor) and part of capital, the investment goods sector uses the remaining part of capital input. Notice that in the investment goods sector, the technology follows an AK type. We will see later that this is the source of the endogenous growth in the model.

Denote p_t as the relative price of capital in terms of consumption goods, non-arbitrage condition shows that the rental price of capital r_t (in terms of consumption goods) is equal to the marginal product of capital, which has to be the same across the two sectors.

$$r_t = \alpha L^{1-\alpha}(\phi_t K_t)^{\alpha-1} = p_t A \tag{6.25}$$

Think about why is not like $r_t = \alpha N_t^{1-\alpha}\phi_t^\alpha {K_t}^{\alpha-1} = p_t A(1-\phi_t)$?

Along BGP, obviously $\phi_t = \phi, \forall t$. Dividing equation (6.24) by K_t on both sides, we obtain

$$\begin{aligned} g_K &\equiv \frac{K_{t+1}}{K_t} = (1-\delta) + A(1-\phi) \\ &= 1 + (A - \delta - A\phi) \\ &\equiv 1 + \rho(\phi). \end{aligned}$$

Using (6.25), we also have (along BGP)

$$\frac{r_{t+1}}{r_t} = \frac{p_{t+1}}{p_t} = (\frac{K_{t+1}}{K_t})^{\alpha-1} = (1+\rho(\phi))^{\alpha-1}. \tag{6.26}$$

From (6.22), we obtain

$$g_c \equiv \frac{c_{t+1}}{c_t} = (\frac{K_{t+1}}{K_t})^{\alpha} = (1+\rho(\phi))^{\alpha}. \tag{6.27}$$

And it is also easy to show

$$g_I = g_K = 1 + \rho(\phi).$$

In a decentralized competitive equilibrium, given standard CRRA utility function, subject to the following BC

$$c_t + p_t K_{t+1} = r_t K_t + (1-\delta) p_t K_t + w_t L$$

We will have following EE

$$(\frac{c_{t+1}}{c_t})^{\sigma} = \beta \frac{(1-\delta)p_{t+1} + r_{t+1}}{p_t} \tag{6.28}$$

Note that RHS of the EE is the present value of the capital gain for the investment. After substituting $r_{t+1} = p_{t+1}A$ and equations (6.27) and (6.26) into (6.28), we arrive at the following equilibrium condition

$$[1+\rho(\phi)]^{1-\alpha(1-\sigma)} = \beta(1-\delta+A). \tag{6.29}$$

Therefore, the growth rate of capital and therefore the growth rate of consumption are positive as long as

$$\beta(1-\delta+A) > 1 \tag{6.30}$$

We achieve the same requirement for long-run growth as in the AK model above. But this result should not be so shocking because again Rebelo's model is essentially an alternative version of AK model. This result also shows that we can allow for any mix in the consumption sector as long as the capital production is "all capital" in its inputs.

Also note that in contrast to the results obtained in the AK model, consumption and capital stock here do not need to grow at the same rate. In fact, when $\alpha = 1$, consumption and investment goods sector share the same AK technology, we can combine two sectors into just one sector and obtain $g_c = g_K = g_I$. Therefore, in this sense AK model is a special case of Rebelo's model.

Moreover, the maintained assumption that parameters are such that derived growth rates yield finite lifetime utility, $\beta(\frac{c_{t+1}}{c_t})^{1-\sigma} < 1$, imposes here a parameter restriction

$$\beta[\beta(1-\delta+A)]^{\frac{\alpha(1-\sigma)}{1-\alpha(1-\sigma)}} < 1$$

which can be simplified to

$$\beta[\beta(1-\delta+A)]^{\alpha(1-\sigma)} < 1. \tag{6.31}$$

Given these conditions on parameters are satisfied, there is a unique equilibrium value of ϕ because LHS of equation (6.29) is monotonically decreasing in $\phi \in [0,1]$ and it is strictly greater (smaller) than the RHS for $\phi = 0$ ($\phi = 1$). The CE outcome is also PO in this economy since there is no distortion in this model.

Chapter 7

Asset Pricing

The objective of this chapter is to introduce the famous asset pricing formula developed by Lucas (1978). We will study the pricing of assets that is consistent with the neoclassical growth model.

In the first section we go over basic Lucas' tree model. He works out his formula using an endowment economy inhabited by one agent. The reason for doing so is that in such an environment the allocation problem is trivial; therefore, only the prices that support a no-trade general equilibrium need to be solved for. In the second section, we study the application of the Lucas pricing formula by Mehra and Prescott (1985). The authors utilized the tools developed by Lucas (1978) to determine the asset prices that would prevail in an economy whose endowment process mimicked the consumption pattern of the United States economy during the last century. They then compared the theoretical results with real data. Their findings were striking and are referred to as the "equity premium puzzle". We will also talk about several popular explanations to solve the puzzle in the third section.

7.1 Lucas' Tree Model

7.1.1 Lucas' Tree Economy

It is a representative agent model. Suppose that there is a tree which produces random amount of fruits every period. We can view these fruits as dividends and use d_t to denote it. Further, we assume d_t is a finite-order Markov process.

$$d_t \sim \Gamma(d_{t+1} = d_i \mid d_t = d_j) = \Gamma_{ij}$$

Let z^t to be the history of realization of dividends, i.e., $z^t = (d_0, d_1, ..., d_t)$. Probability that certain history z^t occurs is $\pi(z^t)$.

HHs in the economy consume the only good here which is fruit. There is no storage technology.

The social planner's problem is

$$\begin{aligned} & \max \sum_{t=0}^{\infty} \beta^t \pi(z^t) u(c_t(z^t)) \\ & s.t. \\ c_t(z^t) & \leq d_t(z^t), \forall t \end{aligned}$$

Obviously the solution to SP would be

$$c_t = d_t, \forall t$$

What about HH's problem? We write it in the way of sequential market.

$$\begin{aligned} & \max_{\{c_t, s_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \pi(z^t) u(c_t(z^t)) \\ & s.t. \\ c_t(z^t) + p_t(z^t) s_{t+1}(z^t) & \leq (p_t(z^t) + d_t(z^t)) s_t(z^{t-1}), \forall t \\ c_t(z^t) & \geq 0, s_{t+1}(z^t) \geq 0 \text{ (short sale constraint)} \end{aligned}$$

where s_t is the number of shares in the tree, and p_t is the price per share of the tree.

Surely we can also write it equivalently in terms of date-0 Arrow-Debreu equilibrium.

$$\begin{aligned} & \max_{\{c_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \pi(z^t) u(c_t(z^t)) \\ & s.t. \\ \sum_t \sum_{z^t} p_t(z^t) c_t(z^t) & \leq \sum_t \sum_{z^t} p_t(z^t) d_t(z^t) \end{aligned}$$

Let's use SME and ignore the argument of z^t for all variables. FOCs w.r.t. s_{t+1} yields

$$\begin{aligned} c_t &: \beta^t \pi u'(c_t) = \mu_t \\ s_{t+1} &: -\mu_t p_t + \mu_{t+1}(p_{t+1} + d_{t+1}) = 0 \end{aligned}$$

Combining these two equations, we have

$$\begin{aligned} \pi(z^t) u'(c_t(z^t)) p_t(z^t) &= \beta \pi(z^{t+1}) u'(c_{t+1}(z^{t+1}))(p_{t+1}(z^{t+1}) + d_{t+1}(z^t)) \\ p_t(z^t) &= \beta \frac{\pi(z^{t+1})}{\pi(z^t)} \frac{u'(c_{t+1}(z^{t+1}))}{u'(c_t(z^t))}(p_{t+1}(z^{t+1}) + d_{t+1}(z^t)) \end{aligned}$$

Notice that $\pi(z^{t+1} \mid z^t) = \frac{\pi(z^{t+1})}{\pi(z^t)}$, therefore, we have

$$\begin{aligned} p_t &= \beta \pi(z^{t+1} \mid z^t) \frac{u'(c_{t+1})}{u'(c_t)}(p_{t+1} + d_{t+1}) \\ &= \beta \sum_j \Gamma_{ij} \frac{u'(c_{t+1})}{u'(c_t)}(p_{t+1} + d_{t+1}) \end{aligned}$$

In equilibrium, $c_t = d_t$, thus we have the asset pricing formula in the Lucas' tree economy.

$$p_t = \beta \sum_j \Gamma_{ij} \frac{u'(d_{t+1})}{u'(d_t)} (p_{t+1} + d_{t+1}) \tag{7.1}$$

Forwardly iterating (7.1) and using the law of iterated expectations ($E_t E_{t+1} E_{t+2} ... = E_t$)[1]

$$\begin{aligned}
p_t &= \beta E_t \{ \frac{u'(d_{t+1})}{u'(d_t)} (p_{t+1} + d_{t+1}) \} \\
&= \beta E_t \{ \frac{u'(d_{t+1})}{u'(d_t)} [\beta E_{t+1} \{ \frac{u'(d_{t+2})}{u'(d_{t+1})} (p_{t+2} + d_{t+2}) \} + d_{t+1}] \} \\
&= \beta E_t \{ \frac{u'(d_{t+1})}{u'(d_t)} d_{t+1} \} + \beta^2 E_t [\frac{u'(d_{t+1})}{u'(d_t)}] E_{t+1} [\frac{u'(d_{t+2})}{u'(d_{t+1})} d_{t+2}] + \beta^2 E_t [\frac{u'(d_{t+1})}{u'(d_t)}] E_{t+1} [\frac{u'(d_{t+2})}{u'(d_{t+1})} p_{t+2}] \\
&= \beta E_t [\frac{u'(d_{t+1})}{u'(d_t)} d_{t+1}] + \beta^2 E_t [\frac{u'(d_{t+1})}{u'(d_t)} \frac{u'(d_{t+2})}{u'(d_{t+1})} d_{t+2}] + \beta^2 E_t [\frac{u'(d_{t+1})}{u'(d_t)} \frac{u'(d_{t+2})}{u'(d_{t+1})} p_{t+2}] \\
&= ...... \\
&= E_t \sum_{j=1}^{\infty} \beta^j \left\{ \prod_{s=0}^{j-1} \frac{u'(d_{t+s+1})}{u'(d_{t+s})} \right\} d_{t+j} \\
&= E_t \sum_{j=1}^{\infty} \underbrace{\beta^j \frac{u'(d_{t+j})}{u'(d_t)}}_{\text{stochastic discount factor (SDF)}} d_{t+j}
\end{aligned}$$

This equation says that the share price is an expected discounted stream of future dividends.

Example 42 *(Logarithmic preferences) when $u(c) = \ln c$, we have*

$$\begin{aligned}
p_t &= E_t \sum_{j=1}^{\infty} \beta^j \frac{d_t}{d_{t+j}} d_{t+j} \\
&= E_t \sum_{j=1}^{\infty} \beta^j d_t \\
&= \sum_{j=1}^{\infty} \beta^j d_t = d_t \sum_{j=1}^{\infty} \beta^j \\
&= \frac{\beta}{1-\beta} d_t
\end{aligned}$$

This asset-pricing function maps the state of the economy at t, d_t, into the price of a Lucas tree at t. When dividend d_t is higher, the price of share of the tree is also higher.

[1] Here we assume there is no bubble, i.e., the term $E_t \{ \lim_{j \to \infty} \beta^j \frac{u'(c_{t+j})}{u'(c_t)} p_{t+j} \} = 0$.

Example 43 *(A finite-state version) Mehra and Prescott (1985) consider a discrete-state version of Lucas' tree model. Let's assume the dividend process has n possible distinct values $[\sigma_1, \sigma_2, ..., \sigma_n]$ and it evolves according to a Markov chain with the following transition probability*

$$\Pr\{d_{t+1} = \sigma_j \mid d_t = \sigma_i\} = \Gamma_{ij} > 0, \forall i, j \in [1, 2, ..., n].$$

The $n \times n$ matrix Γ with a generic element Γ_{ij} is called a stochastic matrix. And we have $\sum_{j=1}^{n} \Gamma_{ij} = 1$, $\forall i \in [1, 2, ..., n]$. Recall the Euler equation in the Lucas' tree model

$$p_t u'(d_t) = \beta \sum_{j=1}^{n} \Gamma_{ij} u'(d_{t+1})(p_{t+1} + d_{t+1})$$

where $d_t = \sigma_i$ and $d_{t+1} = \sigma_j$. Therefore, more sepcifically we have

$$p(\sigma_i)u'(\sigma_i) = \beta \sum_{j=1}^{n} \Gamma_{ij} p(\sigma_j) u'(\sigma_j) + \beta \sum_{j=1}^{n} \Gamma_{ij} u'(\sigma_j)\sigma_j$$

If we define $v_i = p(\sigma_i)u'(\sigma_i)$ and $\alpha_j = \beta \sum_{j=1}^{n} \Gamma_{ij} u'(\sigma_j)\sigma_j$, we can rewrite the above equation as

$$v_i = \alpha_j + \beta \sum_{j=1}^{n} \Gamma_{ij} v_j$$

Or in matrix terms

$$\underbrace{v}_{n\times 1} = \underbrace{\alpha}_{n\times 1} + \beta \underbrace{P}_{n\times n} \underbrace{v}_{n\times 1}$$

Therefore, we obtain

$$(I - \beta P)v = \alpha$$

where I is a $n \times n$ identity matrix. This equation has a unique solution given by

$$v = (I - \beta P)^{-1}\alpha.$$

The price of the asset in state σ_i-$p(\sigma_i)$-can then be found from $p(\sigma_i) = \frac{v_i}{u'(\sigma_i)}$. Notice that the equation above can be represented as

$$v = (I + \beta P + \beta^2 P^2 + ...)\alpha$$

Or

$$p(\sigma_i) = p_i = \sum_{j=1}^{n} (I + \beta P + \beta^2 P^2 + ...)_{ij} \frac{\alpha_j}{u'(\sigma_i)}$$

Example 44 *(Lucas Tree with growth) Assume a Lucas' tree in a pure endowment economy with $c_t = d_t$ and $d_{t+1} = \lambda_{t+1} d_t$, where λ_t is a Markov process* *with transition matrix P. Let p_t be the dividend price of the Lucas' tree. Assume CRRA utility function $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$. Under this utility function, EE is*

$$\begin{aligned} p_t &= E_t\left[\beta(\frac{c_{t+1}}{c_t})^{-\gamma}(p_{t+1}+d_{t+1})\right] \\ &= E_t\left[\beta(\frac{d_{t+1}}{d_t})^{-\gamma}(p_{t+1}+d_{t+1})\right] \\ &= E_t\left[\beta(\lambda_{t+1})^{-\gamma}(p_{t+1}+d_{t+1})\right] \end{aligned}$$

Dividing both sides by d_t and rearranging terms

$$\begin{aligned} \frac{p_t}{d_t} &= E_t\left[\beta(\lambda_{t+1})^{-\gamma}(\frac{p_{t+1}}{d_{t+1}}\frac{d_{t+1}}{d_t}+\frac{d_{t+1}}{d_t})\right] \\ &= E_t\left[\beta(\lambda_{t+1})^{-\gamma}(\frac{p_{t+1}}{d_{t+1}}\lambda_{t+1}+\lambda_{t+1})\right] \\ &= E_t\left[\beta(\lambda_{t+1})^{1-\gamma}(\frac{p_{t+1}}{d_{t+1}}+1)\right] \end{aligned}$$

Or in the finite-state version

$$w_i = \beta\sum_{j=1}^{n} P_{ij}\lambda_j^{1-\gamma}(w_j+1)$$

where w_i expresses the price-dividend ratio at state σ_i. Equation above was used by Mehra and Prescott (1985) to compute equilibrium prices.

7.1.2 Basic Asset Pricing Formula

More generally, every consumption-based (means we maximize a utility function over a BC) model will give you the same FOC as the one in Lucas' tree model–equation (7.1)

$$p_t = E_t[\beta\frac{u'(c_{t+1})}{u'(c_t)}x_{t+1}] \tag{7.2}$$

Here x_{t+1} is the *payoff* at time $t+1$. In Lucas' tree economy, the payoff of the stock on tree is $p_{t+1}+d_{t+1}$. Define stochastic discount factor (SDF) m_{t+1} as

$$m_{t+1} \equiv \beta\frac{u'(c_{t+1})}{u'(c_t)}$$

It is also called *pricing kernal*. Then the basic asset pricing formula becomes

$$p_t = E_t[m_{t+1}x_{t+1}]$$

Or in matrix terms

$$p = E[mx]$$

7.2 The Equity Premium Puzzle

The equity premium is the name of an empirical regularity observed in the United States asset markets during the last century. It consists of the difference between the returns on stocks (risky asset) and on government bonds (risk-free asset). Investors who had always maintained a portfolio of shares with the same composition as Standard and Poor's SP500 index would have obtained, if patient enough, a return around 6% higher than those investing all their money in government bonds. Since shares are riskier than bonds, this fact should be explained by the *representative agent*'s dislike for risk. In the usual CRRA utility function, the degree of risk aversion (but notice that also the intertemporal elasticity of substitution!) is captured by the parameter σ.

Mehra and Prescott's exercise was intended to confront the theory with the observations. They computed statistics of the realization of (de-trended) aggregate consumption in the United States, and used those statistics to generate an endowment process in their model economy. That is, their endowment economy mimics the United States economy for a single agent.

Using parameters consistent with microeconomic behavior (drawn from microeconomics, labor, other literature, ...), they calibrated their model to simulate the response of a representative agent to the assumed endowment process. Their results were striking in that the model predicts an equity premium that is significantly lower than the actual one observed in the United States. This incompatibility could be interpreted as evidence against the neoclassical growth model (and related traditions) in general, or as a signal that some of the assumptions used by Mehra and Prescott (profusely utilized in the literature) need to be revised. It is a "puzzle" that the actual behavior differs so much from the predicted behavior, because we believe that the microfoundations tradition is essentially correct and should provide accurate predictions.

7.2.1 The Model

There is a single representative HH with preferences

$$E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\sigma} - 1}{1-\sigma} \right\}$$

Preferences involve two parameters, β and σ, the values of which need to be calibrated. β measures the time impatience of agents. And σ is the coefficient of relative risk aversion (CRRA). This parameter measures two (distinct) effects: the willingness to substitute consumption over time, and also across states of nature. The higher σ, the less variability the agent wants his consumption pattern to show, i.e., people want the consumption smoother.

HH's budget constraint is

$$c_t + p_t^s (s_{t+1} - s_t) + p_t^b b_{t+1} \leq d_t s_t + b_t$$

where d_t is the per-share, per-capita dividend, s_t is the number of shares held (risky asset), and b_t is the per-capita holdings of bonds that pay one unit of consumption good for sure in one period (risk-free asset). The control variables are c_t, s_{t+1}, and b_{t+1}. The state variables are s_t and b_t. HHs take processes for the prices (p_t^s and p_t^b) and dividends (d_t) as given when solving their problem. In equilibrium, consumption must equal dividends.

FOCs are

$$\begin{aligned} c_t &: c_t^{-\sigma} = \mu_t \\ s_{t+1} &: \mu_t p_t^s = \beta E_t \mu_{t+1}[d_{t+1} + p_{t+1}^s] \\ b_{t+1} &: \mu_t p_t^b = \beta E_t \mu_{t+1} \end{aligned}$$

Two EEs

$$\begin{aligned} p_t^s &= \beta E_t (\frac{c_{t+1}}{c_t})^{-\sigma}[p_{t+1}^s + d_{t+1}] \\ p_t^b &= \beta E_t (\frac{c_{t+1}}{c_t})^{-\sigma} \end{aligned}$$

Before we move to the detailed part of how to compute these two asset prices. Let's talk about the risk-free interest rate for a while. The (gross) risk-free rate $R_t^b = 1 + r_t^b = \frac{1}{p_t^b}$. Thus we have

$$R_t^b = \beta E_t (\frac{c_{t+1}}{c_t})^{\sigma}. \tag{7.3}$$

From this equation, we can see three effects right away:

1. Real interest rates are high when people are impatient, i.e., when β is low. If everyone wants to consume now, it takes a high interest rate to convince them to save.

2. Real interest rates are high when consumption growth rate is high. Recall that the relative price of consumption tomorrow relative to consumption today is $\frac{1}{1+r_t^b}$. Thus high r_t^b means the relative price of consumption tomorrow is low. Investors thus want to consume less now, invest more, and consume more in the future. Hence high interest rates raise the consumption growth rate from today to tomorrow.

3. Real interest rates are more sensitive to consumption growth if the CRRA parameter σ is large. It the curvature of the utility function is high, i.e., σ is high, the investor is highly risk averse, therefore he cares more about maintaining a consumption profile that is smooth over time. And he is less willing to rearrange consumption over time in response to interest rate incentives. Thus it takes a larger interest rate change to induce him to a given consumption growth.[2]

[2] Let's ignore the expectation at this moment, rewrite equation (7.3) as

$$\frac{1}{1 + r_t^b} = \beta (\frac{c_{t+1}}{c_t})^{-\sigma}$$

As we showed before, if we recurse forward for p^s, we have

$$p_t^s = E_t\left[\sum_{\tau=t+1}^{\infty}\beta^{\tau-t}(\frac{c_\tau}{c_t})^{-\sigma}d_\tau\right]$$

In the equilibrium, we will always have $c_t = d_t$ (think about why?). Therefore

$$p_t^s = E_t\left[\sum_{\tau=t+1}^{\infty}\beta^{\tau-t}(\frac{d_\tau}{d_t})^{-\sigma}d_\tau\right] \quad (7.4)$$

$$p_t^b = E_t\left[\beta(\frac{d_{t+1}}{d_t})^{-\sigma}\right] \quad (7.5)$$

Mehra and Prescott (1985) assume that the growth rate of dividends is governed by a Markov Chain. Let $x_{t+1} = \frac{d_{t+1}}{d_t}$. This growth rate can take on n possible values $\{\lambda_1, \lambda_2, ..., \lambda_n\}$. Mehra and Prescott then guess a linear solution to equation (7.4)

$$p^s(d,i) = w_i d$$

which shows the stock price at time t is a function of dividend at time t $d_t = d$ and the value of state x_t takes λ_i.

As we showed before, we can use FOCs to obtain a set of equations for w_i

$$p^s(d,i) = \beta\sum_{j=1}^{n}\pi_{ij}(\frac{\lambda_j d}{d})^{-\sigma}[p^s(\lambda_j d, j) + \lambda_j d]$$

Or

$$w_i d = \beta\sum_{j=1}^{n}\pi_{ij}(\frac{\lambda_j d}{d})^{-\sigma}[w_j\lambda_j d + \lambda_j d]$$

$\Rightarrow$

$$w_i = \beta\sum_{j=1}^{n}\pi_{ij}(\lambda_j)^{1-\sigma}[w_j + 1]$$

This is a system of n equations (one for each w_i) in n unknowns $\{w_i\}_{i=1}^n$. Note that the w_i's are implicit functions of preference parameters β and σ, the transition probabilities π_{ij}, and the growth rates of dividends λ.

Take log on both sides of the equation above, we have

$$\log(\frac{1}{1+r_t^b}) = -\sigma\log(\frac{c_{t+1}}{c_t}) + \log\beta$$

Thus we have

$$-\frac{\partial\log(\frac{c_{t+1}}{c_t})}{\partial\log(\frac{1}{1+r_t^b})} = \frac{1}{\sigma}.$$

Recall that from section 6.2.5, the LHS is called Intertemporal Elasticity of Substitution (IES), it measure the sensitivity of consumption growth w.r.t. the change in the relative price of consumption. Given the consumption growth, from the equation above, it is obvious that higher σ implies larger change in real interest rate $1 + r_t^b$.

Similarly, we also obtain the price for one-period bond

$$p^b(d,i) = \beta \sum_{j=1}^{n} \pi_{ij} (\lambda_j)^{-\sigma}.$$

We define one-period net returns for stocks and bonds as following:

$$\begin{aligned} r^s_{t,t+1} &= \frac{d_{t+1} + p^s_{t+1}}{p^s_t} - 1 \\ r^b_t &= \frac{1}{p^b_t} - 1. \end{aligned}$$

Once we have computed the w_i's, we can write the one-period return for stocks conditional on state i today and state j tomorrow as

$$\begin{aligned} r^s_{i,j} &= \frac{\lambda_j d + p^s(\lambda_j d, j)}{p^s(d,i)} - 1 \\ &= \frac{\lambda_j d + w_j(\lambda_j d)}{w_i d} - 1 \\ &= \frac{\lambda_j + w_j \lambda_j}{w_i} - 1 \end{aligned}$$

Expected returns for stocks and bonds conditional on state i today are

$$\begin{aligned} R^s_i &= \sum_{j=1}^{n} \pi_{ij} r^s_{i,j} = \sum_{j=1}^{n} \pi_{ij} [\frac{\lambda_j + w_j \lambda_j}{w_i} - 1] \\ R^b_i &= \sum_{j=1}^{n} \pi_{ij} r^b_{i,j} = \sum_{j=1}^{n} \pi_{ij} (\frac{1}{p^b(j)} - 1) \\ &= \left[\beta \sum_{j=1}^{n} \pi_{ij} (\lambda_j)^{-\sigma} \right]^{-1} - 1 \end{aligned}$$

The unconditional returns are then given by

$$\begin{aligned} R^s &= \sum_{i=1}^{n} \pi_i R^s_i \\ R^b &= \sum_{i=1}^{n} \pi_i R^b_i. \end{aligned}$$

The equity premium is defined by

$$EP = R^s - R^b$$

7.2.2 The Calibration

The US data show the equity premium for the period 1889-1978 is

$$R^s - R^b = 6.98\% - 0.80\% = 6.18\%,$$

where R^s is the average return on the S&P 500 from 1889 to 1978, and R^b is the average yield on government bonds throught that period. Mehra and Prescott tried to ask a question how large is the predicted difference in returns on a risky asset and a riskless asset in the standard Lucas' tree model? They choose parameters of the Markov chain to mimic properties of US consumption data for the period 1889-1978. More specifically, they use a two-state Markov chain with two possible values

$$\begin{aligned} \lambda_1 &= 1 + \mu + \delta \\ \lambda_2 &= 1 + \mu - \delta \end{aligned}$$

where μ is the average growth rate of consumption $\frac{c_{t+1}-c_t}{c_t}$. In data $\mu = 0.018$. δ is the variation in the growth rate. The transition matrix is assumed to be $\pi_{11} = \pi_{22} = \phi$, i.e.,

$$P = \begin{pmatrix} \phi & 1-\phi \\ 1-\phi & \phi \end{pmatrix}$$

Then ϕ and δ are picked to match

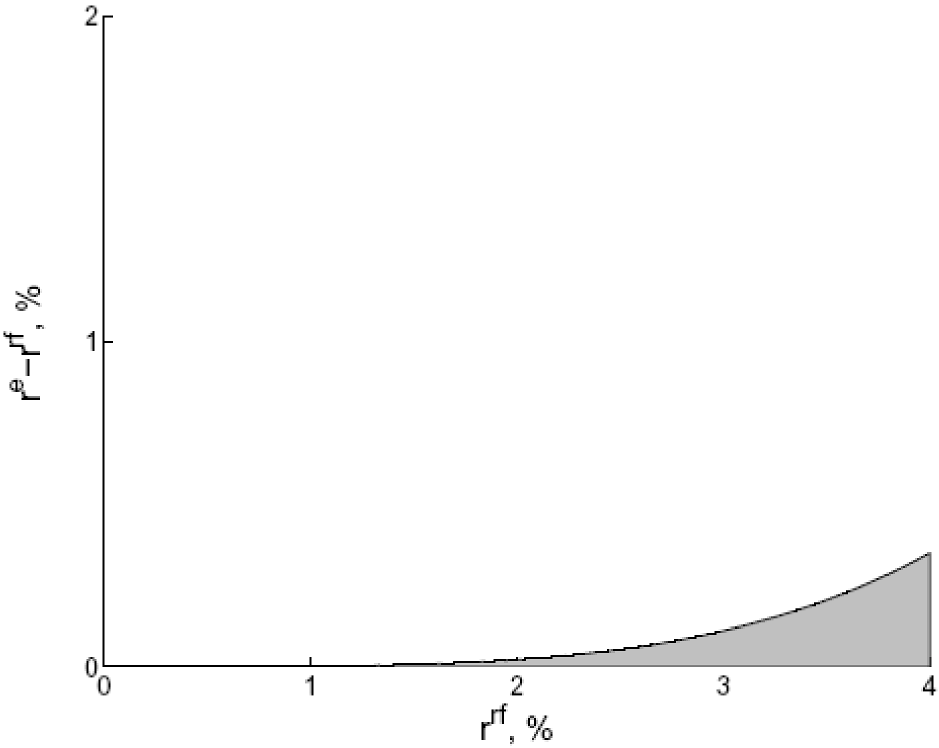
- the standard deviation of comsumption growth rate $\frac{c_{t+1}-c_t}{c_t}$ in the data is 0.036.

- the first-order serial correlation of $\frac{c_{t+1}-c_t}{c_t}$ is -0.14.

The resulting parameter values are: $\delta = 0.036$, and $\phi = 0.43$.

The remaining parameters are preference parameters β and σ. A priori, by economists believe that β must lie in the interval $(0, 1)$. With respect to σ, Mehra and Prescott cite several different studies and opinions on its likely values. Most micro literature suggests that σ must be approximately equal to 1 (this is the logarithmic preference), However, some economists also believe that it could take values as high as 2 or even 4. Certainly, there seems to be consensus that σ has to be lower than 10.

Then instead of picking values for β and σ, Mehra and Prescott plotted the level of equity premium that the model would predict for different, *reasonable* combinations of values. Figure 7.2.2 below shows approximately what they have obtained (it is a reproduction of Figure 4 of the original paper).



Set of admissible equity premia and risk-free returns

Under reasonable parameters, the model can only produce the largest premium value like 0.35%, which is far away from the data. The model can only generate the equity premium observed in actual data at the expense of a very high risk-free interest rate (e.g., $\sigma > 20$), or highly unreasonable parameter values (such as $\beta > 1$). Actually only for $\beta = 1.08$ and $\sigma = 18$ that we can replicate the equity premium as in the data.[3] Therefore, when compared to actual data, the risk premium is too low in the model. In addition, the risk-free rate is only 0.80% in the data, but the minimum rate that the model can generate is still way above it. The risk-free rate is too high in the model. In this sense, in fact, these are actually two puzzles. One is *"too low" equity premium puzzle.* Another one is *"too-high" risk-free rate puzzle.*

What is the intuition of "Equity Premium Puzzle"? Combining two Euler equations w.r.t. risky and risk-free assets in the model, we have

$$E_t \left\{ \beta(\frac{c_{t+1}}{c_t})^{-\sigma}(R^s_{t,t+1} - R^b_t) \right\} = 0$$

[3] See Kacherlakota (1996).

where $R^s_{t,t+1} = \frac{d_{t+1}+p^s_{t+1}}{p^s_t}$. Denote $m_{t+1} = \beta(\frac{c_{t+1}}{c_t})^{-\sigma}$ and $EP_{t+1} = R^s_{t,t+1} - R^b_t$, we have

$$E_t \{m_{t+1} \cdot EP_{t+1}\} = E_t(m_{t+1}) \cdot E_t(EP_{t+1}) + cov(m_{t+1}, EP_{t+1}) = 0$$

$\Rightarrow$

$$E_t(EP_{t+1}) = E_t(R^s_{t,t+1}) - R^b_t = -\frac{cov(m_{t+1}, EP_{t+1})}{E_t(m_{t+1})}$$

Recall $R^b_t = 1/p^b_t = 1/E_t[\beta(\frac{c_{t+1}}{c_t})^{-\sigma}] = 1/E_t(m_{t+1})$, thus we have

$$\text{expected equity premium} = -R^b_t cov(m_{t+1}, EP_{t+1})$$

Assets whose returns covary negatively with consumption growth ($cov(m_{t+1}, EP_{t+1}) < 0$), i.e., pays a lot when consumption is high, hence $\frac{c_{t+1}}{c_t} \uparrow$ and marginal rate of substitution is low. It makes consumption more volatile, and so must promise higher expected returns to persuade investors to hold them. Conversely, assets that covary positively with consumption growth (pay off more when the consumption is low hence MRS is high), such as insurance, helps to smooth consumption, therefore it is more valuable than its expected payoff might indicate. In other words, it can offer expected rates of return that are even lower than the risk-free rate because you do not need to persuade people to hold it. The problem is, aggregate consumption in the data is not volatile enough, thus the covariance between consumption growth and equity premium is not negative enough to make the expected equity premium high enough under the "reasonable" CRRA coefficient.

Note that

$$\begin{aligned} cov(m_{t+1}, EP_{t+1}) &= cov(m_{t+1}, R^s_{t,t+1}) \\ &= std(m_{t+1}) std(R^s_{t,t+1}) corr(m_{t+1}, R^s_{t,t+1}) \end{aligned}$$

Therefore,

$$\begin{aligned} E_t(R^s_{t,t+1}) - R^b_t &= -R^b_t std(m_{t+1}) std(R^s_{t,t+1}) corr(m_{t+1}, R^s_{t,t+1}) \\ &= -corr(m_{t+1}, R^s_{t,t+1}) \left(\frac{std(m_{t+1})}{E_t(m_{t+1})}\right) std(R^s_{t,t+1}) \end{aligned}$$

Since correlation $-1 \leq corr(m_{t+1}, R^s_{t,t+1}) \leq 1$, we have

$$\left|\frac{ER^s - R^b}{std(R^s)}\right| \leq \frac{std(m)}{Em} \tag{7.6}$$

The LHS is called *Sharpe Ratio.* The RHS is called the *market price of risk.* This inequality is also referred to as the "Hansen-Jagannathan lower bound on the pricing kernal". It indicates the lower bound of the market price of risk.

Under the lognormal assumption, after some tedious algebra, we can derive

$$\frac{std(m)}{Em} \approx \sigma std(\triangle \ln c)$$

If consumption is more volatile ($std(\triangle \ln c) \uparrow$), or CRRA coefficient is higher ($\sigma \uparrow$), then the market price of risk is higher, so does the equity premium. Over the last 50 years in the US, aggregate nondurable and services consumption growth has a mean and standard deviation of about 1%. The historical annual market Sharpe Ratio has been about 0.5. This implies to make equation (7.6) hold, the CRRA coeffcient should be at least 50! That's another way to view the "Equity Premium puzzle".

7.3 Solving "Equity Premium Puzzle"

In this section we will briefly discuss the solutions that have been proposed in the literature.

7.3.1 Epstein-Zin Recursive Utility

One of the shortcomings of the CRRA utility function is that for this kind of preference, the coefficient of relative risk aversion and the elasticity of intertemporal substitution are represented in just one parameter: σ. In some sense, this is consistent with the way the risk is modelled in expected utility framework: remember that uncertainty is just the expansion of the decision making scenario to a multiplicity of "states of nature". Total utility is just the expected value of optimal decision making in each of these states. You may notice there is no difference between "time" and "states of nature". "Time" is just another subindex to identify states of the world.

However, people seem to regard time and uncertainty as essentially different phenomena. It is natural then to seek a representation of preferences that can treat these two components of reality separately. This has been addressed by Epstein and Zin (1989), who axiomatically worked on non-expected utility and came up with the following (non-expected) utility function representation for a preference relation that considers time and states of nature as more than just two indices of the state of the world:

$$U_t = \left[c_t^{1-\rho} + \beta (E_t [U_{t+1}^{1-\sigma}]^{\frac{1-\rho}{1-\sigma}} \right]^{\frac{1}{1-\rho}}, 1 \neq \rho > 0, 1 \neq \sigma > 0$$

where $1/\rho$ measures intertemporal elasticity of substitution, and σ captures risk aversion. Utility function today is a CES function w.r.t consumption today and discounted expected utility tomorrow. Note that if $\rho = \sigma$, we go back to original CRRA utility by solving forward the recursive structure.

With this recursive preference, we can show that EEs now are

$$\begin{aligned} \beta E_t \left\{ (E_t U_{t+1}^{1-\sigma})^{\frac{\sigma-\rho}{1-\rho}} U_{t+1}^{\rho-\sigma} (\frac{c_{t+1}}{c_t})^{-\rho} R_t^b \right\} &= 1 \\ E_t \left\{ (\frac{c_{t+1}}{c_t})^{-\rho} U_{t+1}^{\rho-\sigma} (R_{t,t+1}^s - R_t^b) \right\} &= 0 \end{aligned}$$

Note that when $\rho = \sigma$, once again the EE goes back to the one as in the original Mehra-Prescott model.The additional term $U_{t+1}^{\rho-\sigma}$ here plays an important role in reconciling the equity premium puzzle. But the problem is it is unobservable at time t.

This proposed solution is able to account for the risk-free rate puzzle. For example, with Epstein-Zin preference, when we have parameters $\beta = 0.99$, $\sigma = 2.00$, and $\frac{1}{\rho} = 8.126$, we can resolve the risk-free rate puzzle to replicate the data. However, to match the equity premium it still requires an unreasonably high σ.

The intuition behind these result is simple. The main benefit of this generalized expected utility (GEU) is that individual attitudes towards (static) risk and (dynamic) growth are no longer governed by the same parameter. To make the risk-free rate low enough to match data, we have to make intertemporal elasticity of substitution (IES) high enough (recall the relative price of consumption tomorrow to consumption today is $1/(1 + r^b)$, when r^b is lower, relative price of consumption tomorrow is higher, people want to consume more today, smoothing consumption motive is lower, hence IES $\frac{1}{\rho}$ is higher). The standard CRRA preference cannot do it because in order to match equity premium which mostly is governed by risk aversion, risk aversion σ has to be high. But since under CRRA preference, IES is just the reciprocal of risk aversion, therefore IES has to be low, induces much higher risk-free rate in the model than data. Esptein-Zin utility solves this puzzle by allowing IES and risk aversion to be high simultaneously.

7.3.2 Habit Persistence

Campbell and Cochrane modify preferences of the standard Lucas' tree model to allow for a time-varying level of subsistence (or "habit"). Adding in "habit" will change the form of SDF. If the consumption falls toward this subsistence level, curvature of utility function increases. In this way, they tie risk aversion to the level of consumption.

The Model

HH's preference

$$\max E_0 \sum_{t=0}^{\infty} \beta^t \frac{(C_t - X_t)^{1-\gamma} - 1}{1-\gamma}$$

When $X_t = 0$, we go back to the standard CRRA utility function. Here X_t represents the level of "habit", and it is a weighted average of past consumption.

$$x_t = \lambda \sum_{t=0}^{\infty} \phi^j c_{t-j}$$

Or equivalently

$$x_t = \phi x_{t-1} + \lambda c_t.$$

Here small letters denote the log of capital letters, e.g., $x_t = \ln X_t$. We can check that with this preference, CRRA is

$$\eta_t \equiv -C_t \frac{U_{cc}(C_t, X_t)}{U_c(C_t, X_t)} = \frac{\gamma}{S_t}$$

where S_t is the "surplus"

$$S_t = \frac{C_t - X_t}{C_t}$$

Therefore, when C_t is closer to the subsistence (or habit) level, S_t is smaller and the CRRA gets bigger. HH is becoming more risk-averse.

The economy is an endowment economy with i.i.d. consumption growth:

$$\ln(\frac{C_{t+1}}{C_t}) = c_{t+1} - c_t = g + v_{t+1},\ v_{t+1} \sim i.i.d.\mathcal{N}(0, \sigma^2).$$

The surplus is governed by an $AR(1)$ process

$$\begin{aligned} s_{t+1} &= (1-\phi)\bar{s} + \phi s_t + \lambda(s_t) v_{t+1} \\ &= (1-\phi)\bar{s} + \phi s_t + \lambda(s_t)(c_{t+1} - c_t - g) \end{aligned}$$

We also allow consumption to affect habit differently in different states by specifying a square root type process rather than a simple $AR(1)$.

$$\begin{aligned} \lambda(s_t) &= \frac{1}{\bar{S}} \sqrt{1 - 2(s_t - \bar{s})} - 1 \\ \bar{S} &= \sigma \sqrt{\frac{\gamma}{1-\phi}} \end{aligned}$$

The specification of $\lambda(s_t)$ guarantees that (1). $C - X > 0, \forall t$; (2). X and C move in the same direction; and (3). X reacts with delay to changes in consumption.

S_t now is the only state variable in this model. Time-varying expected returns, price/dividend ratio, etc. are all functrions of this variable.

The marginal utility is

$$U_c(C_t, X_t) = (C_t - X_t)^{-\gamma} = S_t^{-\gamma} C_t^{-\gamma}.$$

This model assumes *external* habit in the sense that each individual's habit is determined by everyone else's consumption, i.e., X_t is independent with C_t. This simplification allows us to ignore terms by which current consumption will affect future habits.

SDF is

$$M_{t+1} = \beta \frac{U_c(C_{t+1}, X_{t+1})}{U_c(C_t, X_t)} = \beta(\frac{S_{t+1}}{S_t} \frac{C_{t+1}}{C_t})^{-\gamma}.$$

The EE for risk-free asset is

$$E_t\{M_{t+1} R_t^b\} = 1.$$

Since we have a stochastic process for S and C, and each is lognormal, we can (through tedious algebra) obtain

$$\begin{aligned} r_t^b &= -\ln E_t(M_{t+1}) = -\ln\beta + \gamma g - \frac{1}{2}\gamma(1-\phi) \\ &= -\ln\beta + \gamma g - \frac{1}{2}(\frac{\gamma}{\bar{S}})^2\sigma^2. \end{aligned}$$

The EE for risky asset is

$$P_t(s_t) = E_t\{M_{t+1}(P_{t+1}(s_t) + C_{t+1}(s_t))\}$$

Thus the price/dividend ratio (as same as price/consumption ratio) is given by

$$\begin{aligned} \frac{P_t}{C_t}(s_t) &= E_t\{M_{t+1}\frac{C_{t+1}}{C_t}(1+\frac{P_{t+1}}{C_{t+1}}(s_{t+1}))\} \\ &= \beta E_t\{e^{-\gamma(s_{t+1}-s_t)-\gamma(c_{t+1}-c_t)}e^{c_{t+1}-c_t}(1+\frac{P_{t+1}}{C_{t+1}}(s_{t+1}))\} \\ &= \beta E_t\{e^{-\gamma(s_{t+1}-s_t)+(1-\gamma)(c_{t+1}-c_t)}(1+\frac{P_{t+1}}{C_{t+1}}(s_{t+1}))\} \\ &= \beta E_t\{e^{-\gamma((\phi-1)(s_t-\bar{s})+\lambda(s_t)v_{t+1})+(1-\gamma)(g+v_{t+1})}[1+\frac{P_{t+1}}{C_{t+1}}(s_{t+1})]\}. \end{aligned}$$

This equation can be solved numerically. Use price/dividend ratio, we can in turn calculate returns and expected returns.

The Calibration

Campbell and Cochrane choose the free parameters of the model to match certain moments of the US postwar data. There are some parameters whose values are directly taken from data. For example, the mean and standard deviation of log consumption growth g and σ are 1.89% and 1.50% directly from the data. The log of risk-free rate r^b is 0.94% from the postwar data. Then they choose persistence coefficient ϕ to match the autocorrelation of the log price/dividend ratio. they choose subjective discount factor β to match risk-free rate with the average real returns on T-bills. And they search for a value of γ to match the postwar Sharpe Ratio. The resulting parameter values are $\phi = 0.87$, $\beta = 0.89$, and $\gamma = 2.00$ (reasonable!).

Given these parameters, Campbell and Cochrane simulate 100000 annually artificial data points to calculate population values of a variety of statistics. Table (7.1) below reports means and standard deviations predicted by the model and compare with the data.

The model accounts for expected equity premium, standard deviation of equity premium, and standard deviation of price/dividend ratio very well. It also does fairly well in terms of mean p/d ratio.

How does this model work? Recall that with habit, the risk aversion coefficient is now time-varying

$$\eta_t \equiv \frac{\gamma}{S_t}$$

Statistic	**Model (%)**	**Data (%)**
Sharpe Ratio	0.50 (construction)	0.50
$E(R^s - R^b)$	6.64	6.69
$\sigma(R^s - R^b)$	15.2	15.7
mean p/d ratio	18.3	24.7
std (p/d ratio)	0.27	0.26

Table 7.1: Means and standard deviations of simulated and historical data

As consumption falls towards habit, people become much less willing to tolerate further falls in consumption, i.e., they become very risk averse. Therefore, although we can have a low CRRA coeffcient γ, but we can still obtain a high, time-varying curvature. Recall our fundamental equation for Sharpe Ratio, we have

$$\frac{ER^s - R^b}{std(R^s)} \approx \eta_t \cdot std(\triangle \ln c) \cdot corr(\triangle c, R^s)$$

Now even given constant consumption volatility $std(\triangle \ln c)$ and correlation $corr(\triangle c, R^s)$, since η_t can be really high, this means that the model has a opportunity to explain high equity premium as in the data. An obvious implication of the model is the equity premium is counter-cyclical (high in recessions and low in booms).

What about risk-free rate puzzle? Right now looks like the model assumes that right number of r^b. Actually this model does have a flavor to solve risk-free rate puzzle too. Here is the intuition: Suppose we are in a bad time in which consumption is low relative to habit. People want to borrow against future (potential) higher consumption, i.e., they want to consume more and save less today, this force will drive up the interest rate (the demand for bond decreases, drive the price of bond down, the bond price is $1/(1+r^b)$, therefore, r^b is higher). But in this external habit model, people are also much more risk averse when consumption is low. This consideration induces them to save more, in order to build up assets against the event that might be even worse tomorrow. This "*precautionary saving*" motive helps to lower down the interest rate. When the latter dominates, the model can resolve the risk-free rate puzzle.

7.3.3 Alternative Habit Model: Catching Up with the Joneses

Abel (1990) studies the Lucas' tree model with following preference

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

where v_t is a preference parameter specified as

$$v_t \equiv [c_{t-1}^D C_{t-1}^{1-D}]^\gamma, \gamma \geq 0 \text{ and } D \geq 0 \tag{7.7}$$

where c_{t-1} is the consumer's own consumption in period $t-1$ and C_{t-1} is aggregate consumption per capita in period $t-1$. If $\gamma = 0$, then $v_t \equiv 1$ and the utility function is standard time-separable. If $\gamma > 0$ and $D = 0$, the parameter v_t depends only on the *lagged* level of aggregate consumption per capita. This formation is the relative consumption or "Catching up with the Joneses" model. If $\gamma > 0$ and $D = 1$, $u(c_t, v_t) = u(c_t, c_{t-1})$, this is the habit formation model.

Suppose the period utility function has the CRRA form

$$u(c_t, v_t) = \frac{[c_t/v_t]^{1-\sigma}}{1-\sigma}, \sigma > 0$$

Marginal utility is

$$MU_t = \frac{\partial U_t}{\partial c_t} = [1 - \beta\gamma D(\frac{c_{t+1}}{c_t})^{1-\sigma}(\frac{v_{t+1}}{v_t})^{1-\sigma}](\frac{c_t}{v_t})^{1-\sigma}(\frac{1}{c_t}). \tag{7.8}$$

In equilibrium, since all HHs are identical, we have $c_t = C_t = y_t$ in every period. Denote $x_{t+1} = \frac{y_{t+1}}{y_t} = \frac{c_{t+1}}{c_t} = \frac{C_{t+1}}{C_t}$, we have $\frac{v_{t+1}}{v_t} = x_t^{\gamma}$. Hence we can rewrite equation (7.8) as

$$\begin{aligned} MU_t &= H_{t+1} c_t^{-\sigma} v_t^{\sigma-1} \\ \text{where } H_{t+1} &\equiv 1 - \beta\gamma D x_{t+1}^{1-\sigma} x_t^{-\gamma(1-\sigma)}. \end{aligned}$$

Note that if $\gamma D = 0$ then $H_{t+1} \equiv 1$, which is the case for both time-separable and relative consumption preferences.

EE for one-period bond is

$$E_t\{\beta \frac{MU_{t+1}}{E_t(MU_t)} R_{t+1}^b\} = 1$$

SDF is

$$\beta \frac{MU_{t+1}}{E_t(MU_t)} = [\frac{H_{t+2}}{E_t(H_{t+1})}] x_t^{\gamma(\sigma-1)} x_{t+1}^{-\sigma}.$$

Similarly, EE for risky asset is

$$E_t\{\beta \frac{MU_{t+1}}{E_t(MU_t)} \frac{p_{t+1}^s + y_{t+1}}{p_t^s}\} = 1$$

Suppose that consumption growth x_{t+1} is i.i.d. over time, we can obtain explicit solutions for stock and bond price.

Using two-point Markov process for consumption growth with $E(x_t) = 1.018$, $Var(x_t) = 0.036^2$, assume $\beta = 0.99$, Abel reports the unconditional expected returns as in Table below (it is also Table 1 in Abel (1990)).

Panel A in Table (7.2) displays the equity premium puzzle. When risk aversion is as high as 10, the equity premium is still as low as 1.37%. Panel B shows that relative consumption model does a much better job. With $\sigma = 6$, the model predicts equity premium 4.63%. But the problem is it also produces too volatile conditional expected rate of return. Panel C demonstrates that the expected rate of returns are extremely sensitive to the value of σ.

$\boldsymbol{\sigma}$	**Stocks (%)**	**Bonds(%)**
A. Time-separable preferences ($\gamma = 0$)		
0.5	1.93	1.87
1.0	2.83	2.70
6.0	10.34	9.52
10.00	14.22	12.85
B. Relative consumption ($\gamma = 1, D = 0$)		
0.5	2.80	2.76
1.0	2.83	2.70
6.0	6.70	2.07
10.0	14.73	1.59
C. Habit Formation ($\gamma = 1, D = 1$)		
0.86	33.56	4.53
0.94	6.83	3.48
1.00	2.83	2.70
1.06	8.43	1.93
1.14	38.28	0.93

Table 7.2: Unconditional expected returns in the alternative models

Chapter 8

Real Business Cycles and Calibration Exercise

This chapter is heavily based on Krusell's "Lecture Notes for Macroeconomics" and the Chapter 1 in Cooley and Prescott (1995). The purpose of this chapter is to introduce the study of business cycles. By business cycles we mean fluctuations of output around its long term growth trend. In this sense, this chapter complements growth theory to provide a thorough explanation of the behavior of economic aggregates: first, output grows secularly; second, it fluctuates around its long term trend. We have already analyzed the former phenomenon in Chapter 6. The latter is our topic now.

We will first overview the empirical facts that constitute our object of study, and the history of the attempts at explaining these facts. After that, we will study the theory that has developed. We could separate the evolution of this theory in three phases: (i) Pre-Real Business Cycles; (ii) Real Business Cycles; (iii) Current trends.

8.1 Introduction

Business cycle theory in historical perspective:

1. Earlier than Keynes's *General Theory*, economists such as Wesley Mitchell, Simon Kuznets and Frederick Mills established business cycle theory as an important part of twentieth-century economics. Mitchell was primarily concerned with documenting the simultaneity of movement (comovement) of variables over the cycle. F. Mills was concerned with documenting the behavior of prices, and in particular the comovement of prices and quantities over economic expansions and contractions. S. Kuznets studied patterns of both growth and fluctuations.

2. 1930s was a very active period for business cycle research. NBER continued its program of empirically documenting the features of cycles. They

found that the business cycle fluctuations are a recurrent event with many similarities over time and across countries. This finding prompted many attempts to explain the cycle as a natural property or outcome of an economic system.

3. The so-called Keynesian Revolution that followed the publication of the *General Theory* turned attention away from thinking about business cycles to the determination of the level of output at a point of time, i.e., the focus shifts from dynamic to static.

4. Concurrent with the emergence of Keynesian macroeconomics was a renewed interest in the problem of understanding the long-run laws of motion of modern economies. Harrod and Domar initiated the research agenda of modern growth theory. The growth theory also evolved from the observation of empirical regularities, as had the business cycle theory. Kaldor's "stylized facts" (see Chapter 6) of economic growth suggested economic laws at work that can be captured in formal models.

5. But for a long time, study of short-term economic fluctuations and study of long-term growth were divorced. The generally accepted view was that we needed one theory for economic growth and one for business cycle. Several important developments in growth theory established the foundation that made it possible to think about growth theory and business cycles within the same theoretical framework. One is Brock and Mirman (1972)'s characterization of optimal growth path in an economy with stochastic productivity shock. A second was the introduction of the labor-leisure choice into the basic neoclassical model.

6. Kydland and Prescott (1982) is the pioneering work of real business cycles (RBC) theory. It uses modified equilibrium growth model with exogenous productivity shock to explain the cyclical variances of a set of economic time series, the covariances between real output and the other series, and the autocovariance of output. It provides methodological contribution to integrate growth and business cycle theory. According to their view, it is the technology shock that causes the cycles in the economy. The source of the shock is real, and the propagation mechanism is real as well: it is a consequence of the intertemporal substitution of labor that optimizing decision makers choose whenever confronted with such a technology shock.

8.2 Stylized Facts about Business Cycles

In this section we are interested in presenting the main "facts" that business cycle theory seeks to explain.

8.2.1 Early Facts

Burns and Mitchell (1946)'s main findings are:

1. Output in different sectors of the economy have positive covariance.
2. Both investment and consumption of durable goods exhibit high variability.
3. Company profits are very pro-cyclical and variable.
4. Prices are pro-cyclical as well. (This is not true for the post-war years, but it was for the sample analyzed by Burns and Mitchell.)
5. The amount of outstanding money balances and the velocity of circulation are pro-cyclical.
6. Short term interest rates are pro-cyclical.
7. Long term interest rates are counter-cyclical.
8. Business cycles are "all alike" across countries and across time.

But the main weaknesses of Burns and Mitchell's work are: first of all, the work was not carefully done, and was hard to replicate. secondly, there was no solid underlying statistical theory. Relevant issues were not addressed altogether, such as the statistical significance of the assertions. Hence what needed in the profession is a set of for credible (maybe just properly presented!) stylized facts, and then a suitable theoretical framework to explain them. This is exactly what Kydland and Prescott did in their pathbreaking 1982 paper (They won Nobel Prize partly due to this contribution).

8.2.2 Kydland and Prescott (1982): How to Convert Raw Data into Facts

What are the business cycles? Business cycles are the recurrent fluctuations of output about the trend and the co-movements among other aggregate time series. In the first place, since the phenomenon to be studied is the short-run fluctuations of output around its long-term growth, these fluctuations need to be pinned down with precision. Raw data need to be get rid of the secular growth component before the cycle can be identified. This is done by filtering the data, using the method developed by Hodrick and Prescott (the so-called "H-P filter").

H-P Filter

We consider an observed time series (say log of output) y_t as the sum of a cyclical component y_t^c and a growth component y_t^g.

$$y_t = y_t^c + y_t^g.$$

Now we must find a way to *detrend*, means extract the cycle from its long-run growth trend. The idea of H-P filter is to minimize the sum of squared deviations from a given time series $\{y_t\}$ subject to the constraint that the sum of the squared second differences not be too big. That is

$$\begin{aligned} &\min_{\{y_t^g\}_{t=1}^T} \sum_{t=1}^{T} (y_t - y_t^g)^2 \\ &s.t. \\ \sum_{t=2}^{T-1} [(y_{t+1}^g - y_t^g) - (y_t^g - y_{t-1}^g)]^2 \quad \leq \quad &\mu \end{aligned}$$

The smaller is μ, the smoother is the trend path. If $\mu = 0$, the least squares linear time trend results. We can write this problem into a Lagrangian function and let λ be the multiplier for the constraint, then H-P filtering problem is to choose the growth component y_t^g, to minimize the loss function

$$\sum_{t=1}^{T} (y_t - y_t^g)^2 + \lambda \sum_{t=2}^{T-1} [(y_{t+1}^g - y_t^g) - (y_t^g - y_{t-1}^g)]^2$$

The nature of this optimization problem is to trade off the extent to which the growth component tracks the actual data series. For $\lambda = 0$, obviously $y_t = y_t^g, \forall t$, the growth component is simply the series. As $\lambda \to \infty$, the growth component approaches a linear trend. For quarterly data, Hodrick and Prescott chose $\lambda = 1600$, and $\lambda = 400$ for annual data. Once the problem is solved, the object of study is the resulting y_t^c series. With this in hand, "facts" in business cycles research are a series of relevant statistics computed from de-trended data.

K-P Facts

There are three types of *statistics* we look at for the pattern of business cycles. First, we measure the *standard deviation* of variables to understand volatilities.[1]

Volatilities

1. $\sigma_c < \sigma_y$. Consumption (nondurables and services) volatility is about 50% of output volatility. (Think about why? Smoothing consumption).

2. $\sigma_{c_d} >> \sigma_y$. Consumption of durables is more volatile (2.9 times) than output. This implies there is a huge difference between consumption behavior for nondurables and durables. Durables are more like investment goods.

3. $\sigma_i >> \sigma_y$. Investment is more volatile (4.8 times) than output.

4. $\sigma_{TB} > \sigma_y$. Trade balance is more volatile than output.

[1] The standard deviation of random variable X is $\sigma_X = \sqrt{E[X - E(X)]^2}$.

5. $\sigma_N \approx \sigma_y$. The magnitude of fluctuations in output and aggregate hours worked are nearly equal. It is well known that the business cycle is most clearly manifested in the labor market. This fact confirms that.

6. $\sigma_E \approx \sigma_N \approx \sigma_y$. $E =$ employment.

7. $\sigma_{\frac{N}{week}} < \sigma_y$. Average weekly hours fluctuates less than output. This fact together with fact 6 suggests that most fluctuations in total hours represent movements into and out of the labor force (extensive margin) rather than adjustments in average hours of work (intensive margin).

8. $\sigma_K << \sigma_y$. The capital stock fluctuates much less than output.

9. $\sigma_w < \sigma_{\frac{y}{N}}$. Here w is the real wage and y/N is the output per worked hour, i.e., labor productivity. Wage vary less than productivity. We can view this as sort of evidence of "sticky wage".

Correlations Second, we measure the correlation of aggregate variables with real output to capture the extent to which variables are procyclical (positively correlated) or countercyclical (negatively correlated).[2]

1. $\rho(c, y) = 0.80 > 0$. Both consumption of durables and nondurables are strongly procyclical.

2. $\rho(i, y) = 0.76 > 0$. Investment is procyclical.

3. $\rho(\frac{y}{N}, y) = 0.34 > 0$. Productivity is slightly procyclical.

4. $\rho(w, y) = 0.03 \approx 0$.

5. $\rho(K, y) \approx 0$.

6. $\rho(G, y) = 0.04 \approx 0$. G is the government expenditure.

7. $\rho(IM, y) = 0.72 > \rho(EX, y) = 0.37$. Imports are more strongly procyclical than exports.

8. $\rho(P, y) < 0$ (in post-war period). P is the price level.

Lead / Lag Finally we measure the cross-correlation over time to indicate whether there is any evidence that variables lead or lag one another, especially lead or lag output.

We find for most of variables such as consumption, investment, average labor productivity, and price level, they are coincident with y.

Let's summarize the basic business cycle stylized facts in Table 8.1.

[2] The correlation between two random variables X and Y is defined by $\rho(X, Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-E(X))(Y-E(Y))]}{\sigma_X \sigma_Y}$.

Variable	**Cyclicality**	**Lead/Lag**	**Volatility relative to y**
c (nondurables)	procyclical	coincident	less
c (durables)	procyclical	coincident	more
i	procyclical	coincident	more
K	acyclical	lagging	less
G	acyclical	n/a	more
N	procyclical	coincident	almost same
y/N	procyclical	coincident	less
w	acyclical	n/a	less
E	procyclical	lagging	almost same

Table 8.1: Summary of Business Cycle Facts

8.3 Theory After Facts

Once the facts are well established, a theory to account for them can be developed. The research on this was initiated by Kydland and Prescott (1982) and Long and Plosser (1983). The framework is the stochastic neoclassical growth model. And, remember: this project is *quantitative*. Everything is in numbers. The success of a real business cycle model is measured by its ability to numerically replicate the "facts".

8.3.1 The Benchmark Model

The paradigm here is stochastic neoclassical growth model with labor-leisure choice. The economy is populated by infinitely many identical HHs that will exist forever. The HH's preference is defined as

$$E\{\sum_{t=0}^{\infty}\beta^t u(c_t, 1-n_t)\}$$

Aggreagte technology is defined by

$$Y_t = e^{z_t}F(K_t, N_t)$$

The productivity shock z_t is the source of uncertainty in the economy. It evolves according to the law of motion

$$z_{t+1} = \rho z_t + \varepsilon_{t+1}, \rho \in (0,1), \varepsilon \sim N(0, \sigma_\varepsilon^2).$$

Aggregate capital stock follows law of motion

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

Firm's problem at time t is

$$\max_{K_t, N_t} p_t[e^{z_t}F(K_t, N_t) - r_tK_t - w_tN_t], \forall t \tag{8.1}$$

FOCs are

$$r_t = e^{z_t} F_K(K_t, N_t) \tag{8.2}$$

$$w_t = e^{z_t} F_N(K_t, N_t) \tag{8.3}$$

The HH's problem

$$\begin{aligned} & \max_{c,i,n} E\{\sum_{t=0}^{\infty} \beta^t u(c_t, 1-n_t)\} \\ & s.t. \\ c_t + i_t &\leq w_t n_t + r_t k_t \\ k_{t+1} &= (1-\delta)k_t + i_t \end{aligned}$$

Note that HH's state variables are $s_t = (z_t, k_t, K_t)$, but the economy-wide state variables are $S_t = (z_t, K_t)$. Thus, we can also write HH's problem as a DP problem:

$$\begin{aligned} v(z,k,K) &= \max_{c,i,n}\{u(c, 1-n) + \beta E[v(z', k', K')]\} \\ & s.t. \\ c + i &\leq r(z,K)k + w(z,K)n \\ k' &= (1-\delta)k + i \\ K' &= (1-\delta)K + I(z,K) \\ z' &= \rho z + \varepsilon' \\ c &\geq 0, 0 \leq n \leq 1 \end{aligned} \tag{8.4}$$

Definition 45 *A Recursive Competitive Equilibrium (RCE) is a collection of value function $v(z,k,K)$; a set of decision rules $c(z,k,K)$, $n(z,k,K)$, and $i(z,k,K)$ for the HH; a corresponding set of aggregate per capita decison rules $C(z,K)$, $N(z,K)$, and $I(z,K)$; and factor price equations $w(z,K)$ and $r(z,K)$, such that:*

(i). Given prices, $v(z,k,K)$, $c(z,k,K)$, $n(z,k,K)$, and $i(z,k,K)$ solve HH's problem (8.4).

(ii). Factor price function $w(z,K)$ and $r(z,K)$ are the solution to the firm's problem (8.1), hence they satisfy (8.2)-(8.3).

(iii). The individual and aggregate decisions are consistent, i.e.,

$$\begin{aligned} c(z,k,K) &= C(z,K) \\ n(z,k,K) &= N(z,K) \\ i(z,k,K) &= I(z,K) \\ k &= K \end{aligned}$$

(iv). Aggregate resource constraint is

$$C(z,K) + I(z,K) = Y(z,K).$$

The central methodological issue is how to pick the parameters in the utility and production functions. In this sense, the work of Kydland and Prescott has also a dialectic dimension. The authors are advocates of the technique known as "*calibration*". This is more than merely how to pick values for parameters to solve a numerical problem. It is a way of contrasting models against data as opposed to traditional econometrics. Calibration, sometimes also called "back-of-the-envelope calculations", requires that values for parameters be picked from sources independent of the phenomenon under study. The discipline advocated by Kydland and Prescott bans "curve fitting" practices. For example, admissible sources of parameter values are:

1. Household data on consumption, hours worked, and other microeconomic evidence, for individual preference parameters.

2. Long run trend data for the factor shares in production (namely α in the Cobb-Douglas case).

8.3.2 How to Measure z

In order to measure technology shock z_t, we will take inspiration from the growth accounting technique developed by Solow. In his framework, there are two inputs, capital and labor, and a total factor productivity (TFP) shock. Hence the output takes the form:

$$y_t = F_t(K_t, N_t) = z_t F(K_t, N_t).$$

The issue that we have to address is what z is or more precisely, *what the counterpart in the data to the theoretical variable z_t is.* To this effect, we will assume that time is continuous, and differentiate the production function:

$$dy_t = F(K_t, N_t)dz_t + z_t F_K(K_t, N_t)dK_t + z_t F_N(K_t, N_t)dN_t$$

Then divide both sides by total output y_t

$$\begin{aligned}\frac{dy_t}{y_t} &= \frac{z_t F(K_t, N_t)}{y_t}\frac{dz_t}{z_t} + \frac{z_t F_K(K_t, N_t)K_t}{y_t}\frac{dK_t}{K_t} + \frac{z_t F_N(K_t, N_t)N_t}{y_t}\frac{dN_t}{N_t} \\ &= \frac{dz_t}{z_t} + \frac{z_t F_K(K_t, N_t)K_t}{y_t}\frac{dK_t}{K_t} + \frac{z_t F_N(K_t, N_t)N_t}{y_t}\frac{dN_t}{N_t}\end{aligned}$$

We assume the production technology exhibits CRS and the market is perfectly competitive. Under these assumptions, FOCs of the firm's problem show

$$\begin{aligned}r_t &= z_t F_K(K_t, N_t) \\ w_t &= z_t F_N(K_t, N_t)\end{aligned}$$

Hence we have

$$\underbrace{\frac{dy_t}{y_t}}_{\text{output growth rate}} = \underbrace{\frac{dz_t}{z_t}}_{\text{Technological change}} + \underbrace{\frac{r_t K_t}{y_t}}_{\text{capital share in income}} \frac{dK_t}{K_t} + \underbrace{\frac{w_t N_t}{y_t}}_{\text{labor share in income}} \frac{dN_t}{N_t}$$

Changes in y	**Secular Growth**	**Business Cycles**
Due to changes in K	1/3	0
Due to changes in N	0	2/3
Due to changes in z	2/3	1/3

Table 8.2: Growth and Business Cycle Accounting of Postwar US Economy

Growth accounting exercise thus attributes the economic growth to the three different sources:

1. due to changes in productivity $\frac{dz_t}{z_t}$, this term is often called "Solow Residual".

2. due to changes in capital input $\frac{dK_t}{K_t}$.

3. due to changes in labor input $\frac{dN_t}{N_t}$.

Define capital share $\alpha_t = \frac{r_t K_t}{y_t}$. In the postwar US economy, this share is around one-third. Therefore, labor share $\frac{w_t N_t}{y_t} = 1 - \alpha_t$ is about two-thirds. We know the US postwar economy data about y, K, and N. Thus we can decompose the US postwar economic growth by applying growth accounting exercise. We find that since there is no trend in the average hours worked per worker in the postwar period, so variations in the labor input do not contribute to secular growth. Changes in capital inputs account for approximately 1/3 of the growth in output per worker. This suggests that the remaining 2/3 of the secular growth in output is attributable to improvements in productivity (Solow Residual).

We can do the similar things for business cycles. But the decomposition of output fluctuations for the business cycle reveals that the sources of business cycle fluctuations are quite different. We know for the facts that $\sigma_K << \sigma_y$ and $\rho(K, y) \approx 0$. Therefore capital fluctuations do not acount for much of busienss cycles. About 2/3 of the fluctuations in aggreagte output are attributable to fluctuations in the labor input. Since capital does not vary much, the remaining 1/3 is due to fluctuations in z_t.

So how to measure z? z is measured by the part of the output that cannot be attributable to the production inputs. More specifically, if we assume neoclassical production function takes Cobb-Douglas form, the TFP shock (or Solow Residual) can be measured as

$$z_t = \frac{y_t}{K_t^{\alpha} N_t^{1-\alpha}}.$$

8.3.3 Calibration

We already lay down the theoretical foundation of the RBC, which is the stochastic neoclassic growth model. But the question is: does a model designed to be consistent with long-run economic growth produce the sort of fluctuations that we associate with business cycle? This is indeed a *quantitative* question because we want to ask does the model will produce the results that similar to the quantitative facts we found in the data.

To go from the general theoretical framework to quantitative assessment is a five-step process. In what follows, let me describe these steps briefly in a cookbook procedure fashion.

Restricting the Model Economy to a Parametric Class

The basic observations about economic growth suggest that capital and labor shares have been approximately constant over time even while the relative prices of these inputs have changed. This suggests a Cobb-Douglas production function as following

$$Y_t = e^{z_t} K_t^{\theta} N_t^{1-\theta}.$$

θ is the capital income share. $1-\theta$ is the labor income share. The C-D assumption defines a parametric class of technologies for this model economy.

Speaking preferences, we know from the data that per capita leisure has been approximately constant over time in the postwar period. Meantime, real wages have increased steadily. Taken together, these two observations imply that the income effect and the substitution effect of increases in real wages on leisure exactly offset each other, therefore, the elasticity of substitution between consumption and leisure should be near unity. This suggests the general parametric class of preferences of the form

$$u(c_t, l_t) = \frac{(c_t^{1-\alpha} l_t^{\alpha})^{1-\sigma} - 1}{1-\sigma},$$

where $\frac{1}{\sigma}$ is intertemporal elasticity of substitution (IES) and α is the share parameter for leisure in the composite commodity.

As described earlier, the productivity shock z_t is the source of uncertainty in the economy. It evolves according to the law of motion

$$z_{t+1} = \rho z_t + \varepsilon_{t+1}, \rho \in (0,1), \varepsilon \sim N(0, \sigma_\varepsilon^2).$$

Aggregate capital stock follows law of motion

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

We further assume that the population size of the economy is growing at rate η. And the production technology is subject to a labor-augumenting technological change at rate γ.

The social planner's problem is

$$\begin{aligned}
& \max \sum_{t=0}^{\infty} \beta^t u(C_t, L_t) \\
& s.t. \\
C_t + I_t &\leq e^{z_t} K_t^{\theta} [(1+\gamma)^t N_t]^{1-\theta} \\
K_{t+1} &= (1-\delta) K_t + I_t \\
z_{t+1} &= \rho z_t + \varepsilon_{t+1} \\
C_t, K_{t+1} &\geq 0
\end{aligned}$$

All the capital letters denote aggregate variables.

Suppose the population size of this economy is H_t. We try to convert the agregate variables to per capita variables by the transformation as following[3]

$$\begin{aligned}
\beta^t u(C_t, L_t) &= \beta^t H_t u(\frac{C_t}{H_t}, \frac{L_t}{H_t}) \\
&= \beta^t (\frac{H_t}{H_0}) u(c_t, l_t)
\end{aligned}$$

We normalize H_0 to be one, hence $H_t = (1+\eta)^t$. Small letters stand for per capita variables. We have

$$\beta^t u(C_t, L_t) = \beta^t (1+\eta)^t u(c_t, l_t).$$

Resource constraint changes to

$$\begin{aligned}
\frac{C_t}{H_t} + \frac{I_t}{H_t} &= e^{z_t} (1+\gamma)^{t(1-\theta)} (\frac{K_t}{H_t})^{\theta} (\frac{N_t}{H_t})^{1-\theta} \\
&\Rightarrow \\
c_t + i_t &= e^{z_t} (1+\gamma)^{t(1-\theta)} k_t^{\theta} n_t^{1-\theta}.
\end{aligned}$$

Assume labor force participation rate is 100% and no unemployment, we have additional constraint

$$L_t + N_t = H_t$$

Dividing both sides by H_t, we have

$$l_t + n_t = 1$$

Transforming the capital accumulation equation:

$$\begin{aligned}
\frac{K_{t+1}}{H_{t+1}} \frac{H_{t+1}}{H_t} &= (1-\delta) \frac{K_t}{H_t} + \frac{I_t}{H_t} \\
&\Rightarrow \\
(1+\eta) k_{t+1} &= (1-\delta) k_t + i_t.
\end{aligned}$$

[3] More precisely, this transformation is true only under the Cobb-Douglas preferences (i.e., $\sigma = 0$). But since preferences preserve the ranking under the positive transformation, we are also fine by expressing the following transformation.

Therefore, we transform the original preoblem into the one in terms of per capita variables

$$\begin{aligned} & \max_{\{c_t,n_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t (1+\eta)^t u(c_t, 1-n_t) \\ & s.t. \\ c_t + i_t &= e^{z_t}(1+\gamma)^{t(1-\theta)} k_t^{\theta} n_t^{1-\theta} \\ (1+\eta)k_{t+1} &= (1-\delta)k_t + i_t \\ z_{t+1} &= \rho z_t + \varepsilon_{t+1} \\ c_t, k_{t+1} &\geq 0, 0 \leq n_t \leq 1 \end{aligned}$$

Therefore, there are nine parameters in the model economy: technology parameter θ; preferences parameters α, σ, β; depreciation rate of capital δ; shock parameters ρ and σ_ε^2; growth rate of population and technological change η and γ.

Construct Measurements Consistent with the Parametric Class

Calibrating the parametric class of economies chosen requires that we consider the *correspondence* between the model economy and the measurements that are taken from the US economy, i.e., we try to find a data counterpart of our model variables.

Our data counterpart for the whole economy is US National Income and Product Accounts (NIPA). Let's take a quick look at Table 8.3.

Notice that our model economy is very abstract: it contains no government, no household (home production) sector, no foreign trade and no explicit treatment of inventories. Accordingly, capital K in the model includes capital used in all of these sectors plus the stock of inventories. Similarly output Y also includes the output produced by all of this capital. But NIPA are somewhat inconsistent in their treatment of these issues. For example, the NIPA does not provide a consistent treatment of the household sector. The accounts do include the imputed flow of services from the owner-occupied housing as part of GDP. But they do not attempt to impute the flow of services from the stock of consumer durables. The NIPA treats consumer durables with consumption rather than treating them as investment. When we deal with the measured data, in consistent with the model, we will add the HH's capital stock–the stock of residential structures and the stock of consumer durables–to producer's equipment and structures. To be consistent, we will also have to impute the flow of serivices from durables and add that to measured output.

The measurement issues discussed above are central to the task of calibrating any model economy because a consistent set of measurements is necessary to align the model economy with the data. For example, in order to estimate the capital income share parameter θ in the production function, it is crucial that we make sure measure all forms of capital and include measures of all forms of output. Similarly, when we treat aggregate investment it is necessary to include

Line	**Variables**	**$**
1	Gross domestic product (GDP)	10480.8
2	Personal consumption expenditures	7385.3
3	Durable goods	911.3
4	Nondurable goods	2086
5	Services	4388
6	Gross private domestic investment	1589.2
7	Fixed investment	1583.9
8	Nonresidential	1080.2
9	Structures	266.3
10	Equipment and software	813.9
11	Residential	503.7
12	Change in private inventories	5.4
13	Net exports of goods and services	-426.3
14	Exports	1006.8
15	Goods	697.8
16	Services	309.1
17	Imports	1433.1
18	Goods	1190.3
19	Services	242.7
20	Government consumption expenditures and gross investment	1932.5
21	Federal	679.5
22	Natinal defense	438.3
23	Nondefense	241.2
24	State and local	1253.1

Table 8.3: 2002 US Gross Domestic Product: Billions of dollars

Variables	**NIPA Accounting Category**	**Theory Category**
Personal consumption expenditures:	C	
Durables		I
Nondurables		C
Services		C
Gross private domestic investment	I	I
Net exports	NE	I
Government purchases:	G	
Public consumption		C
Public investment		I

Table 8.4: NIPA categoroes and their model counterparts

in investment additions to all forms of capital stock. For the current model economy, the concept of investment that corresponds to the aggregate capital stock includes government investment, "consumption" of consumer durables, change in inventories, gross fixed investment, and net exports[4]. Making sure that the conceptual framework of the model economy and the conceptual framework of the measured data are consistent, is a crucial step in the process of calibration.

Here in Table 8.4 we summarize the category in NIPA and the counterpart of this category in the theory (model economy).

We will use economic theory to guide us to impute the flow of services from government capital and consumer durables. We know that income from capital is related to the stock of capital in the following way

$$\begin{aligned} Y_{KP} &= (r + \delta_{KP})K_P \\ K_P &: \text{fixed private capital stock} \\ Y_{KP} &: \text{income on fixed private capital} \\ \delta_{KP} &: \text{depreciation rate of capital stock } K_P \\ r &: \text{return on capital.} \end{aligned}$$

K_P will be measured from data:

$$\begin{aligned} K_P = &\text{ net stock of fixed reproducible private capital (not including the stock of consumer durables)} \\ &+ \text{the stock of inventories (from NIPA)} \\ &+ \text{stock of land (from Flow of Funds Accounts)} \end{aligned}$$

The measurement of capital income Y_{KP} is taken from NIPA. There is some concern involved in defining this because of ambiguity about how much of Proprietors' income and some other smaller categories (NNP-NI) should be treated as capital income. We define Y_{KP} as following way. First of all, define unam-

[4]We include net exports into investment because we view net exports representing additions to or claims on the domestic capital stock in this closed model economy.

bigous capital income as

$$\begin{aligned} \text{Unambigous Capital Income} &= \text{Rental income} + \text{Corporate Profits} \\ &\quad + \text{Net interest}, \end{aligned}$$

with all three incomes from NIPA. Next, we want to allocate the ambiguous components of income according to the share of capital income in measured GNP. Let θ_P denote the share of capital in measured GNP. Further, notice that the measured value of $\delta_{KP} K_P$ is called "Consumption of Fixed Capital" (GNP-NNP) in the NIPA.

$$DEP = \delta_{KP} K_P.$$

Now we can define Y_{KP} as follows

$$\begin{aligned} Y_{KP} &= \text{Unambigous Capital Income} \\ &\quad + \theta_P(\text{Proprietors' Income} + \text{NNP-NI}) \\ &\quad + DEP \\ &= \theta_P \cdot GNP \end{aligned}$$

If we ignore the difference between NNP and NI, we can solve for θ_P

$$\theta_P = \frac{\text{Unambigous Capital Income} + DEP}{(GNP - \text{Unambigous Capital Income})}.$$

Once we know θ_P, since GNP is known, we can determine Y_{KP}. Then the return to capital r is determined by

$$r = \frac{Y_{KP} - DEP}{K_p}.$$

For the time period 1954-1992, this yields an average interest rate of 6.9%.

To estimate the flow of services from the stock of consumer durables and the stock of government capital, we adopt the following procedure. First, we need to obtain the depreciation rates for both capital stocks.

Divide the law of motions of each capital stock by Y_t, we have

$$\frac{Y_{t+1}}{Y_t}\frac{K_{t+1}}{Y_{t+1}} = (1-\delta)\frac{K_t}{Y_t} + \frac{I_t}{Y_t}.$$

On a BGP, $\frac{K_{t+1}}{Y_{t+1}} = \frac{K_t}{Y_t}$, and we do have data about $\frac{Y_{t+1}}{Y_t}$, $\frac{K_t}{Y_t}$ and $\frac{I_t}{Y_t}$ from NIPA. Therefore, equation above provides the basis for measuring the depreciation rates for consumer durables and government capital. For the sample period 1954-1992, we end up with

$$\begin{aligned} \delta_D &= 21\% \\ \delta_G &= 5\%. \end{aligned}$$

Once we know the depreciation rates, the service flows are then estimated as

$$\begin{aligned} Y_D &= (r + \delta_D)K_D \\ Y_G &= (r + \delta_G)K_G \end{aligned}$$

Assign Values to the Parameters

Since we construct our data counterpart of the model economy according to the theory, capital now includes private capital, government capital, and consumer durables; and output also includes flows of services from government capital and consumer durables. Therefore, capital income share θ is determined by

$$\theta = \frac{Y_{KP} + Y_D + Y_G}{GNP + Y_D + Y_G} = 0.40.$$

The population growth rate η and growth rate of real per capita output γ are also determined from the data

$$\eta = 0.012, \gamma = 0.0156.$$

We pick CRRA coefficient $\sigma = 1$ (therefore IES=1) which leads to logarithmic utility function

$$(1-\alpha)\log c_t + \alpha \log(1-n_t)$$

Before we move on to show how to calibrate the remaining parameters, we would like to transform the original per capita economy into a stationary one. From Section 6.2, we already know that in this model economy, all variables grow at a rate γ. We then define the de-trended variables:

$$\tilde{c}_t = \frac{c_t}{(1+\gamma)^t}, \tilde{i}_t = \frac{i_t}{(1+\gamma)^t}, \tilde{k}_t = \frac{k_t}{(1+\gamma)^t},$$

With these transformation, preference changes to

$$\beta^t(1+\eta)^t[(1-\alpha)\log\tilde{c}_t + \alpha\log(1-n_t)] + \beta^t(1-\alpha)(1+\eta)^t\log(1+\gamma)^t.$$

Since the second term does not include any choice variable, hence it is irrelevant for our utility maximization.

The law of motion for capital stock changes to

$$\begin{aligned} (1+\eta)(1+\gamma)^{t+1}\tilde{k}_{t+1} &= (1-\delta)(1+\gamma)^t\tilde{k}_t + (1+\gamma)^t\tilde{i}_t \\ &\Rightarrow \\ (1+\eta)(1+\gamma)\tilde{k}_{t+1} &= (1-\delta)\tilde{k}_t + \tilde{i}_t. \end{aligned}$$

And the resource constraint becomes

$$\begin{aligned} (1+\gamma)^t\tilde{c}_t + (1+\gamma)^t\tilde{i}_t &= e^{z_t}(1+\gamma)^t\tilde{k}_t^\theta n_t^{1-\theta} \\ &\Rightarrow \\ \tilde{c}_t + \tilde{i}_t &= e^{z_t}\tilde{k}_t^\theta n_t^{1-\theta}. \end{aligned}$$

Therefore, our new stationary economy is

$$
\begin{aligned}
& \max_{\{c_t, n_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t (1+\eta)^t [(1-\alpha)\log \tilde{c}_t + \alpha \log(1-n_t)] \\
& s.t. \\
\tilde{c}_t + \tilde{\imath}_t &= e^{z_t} \tilde{k}_t^{\theta} n_t^{1-\theta} \\
(1+\eta)(1+\gamma)\tilde{k}_{t+1} &= (1-\delta)\tilde{k}_t + \tilde{\imath}_t \\
z_{t+1} &= \rho z_t + \varepsilon_{t+1} \\
\tilde{c}_t, \tilde{k}_{t+1} &\geq 0, 0 \leq n_t \leq 1
\end{aligned}
$$

We calibrate the remaining parameters $(\delta, \beta, \rho, \sigma_{\varepsilon}, \alpha)$ by choosing them so that *the deterministic balanced-growth path (BGP) of our model economy matches certain long-run features of the measured US economy.* In other words, we pick the parameter values so that it more or less replicates the US economy "*on average*". Therefore, for this step, we ignore the stochastic part of the model economy and focus on its "growth" part. The principle in our mind is: *when modifying the standard growth model to address a business cycle question, the modification should continue to display the growth facts.* (See Prescott (1998)).

EE for this stationary economy is

$$\frac{(1+\gamma)(1+\eta)}{\tilde{c}_t} = \frac{\beta(1+\eta)[\theta \tilde{k}_{t+1}^{\theta-1} n_{t+1}^{1-\theta} + 1 - \delta]}{\tilde{c}_{t+1}}$$

Along BGP, notice that $\tilde{c}_{t+1} = \tilde{c}_t = \tilde{c}$, $\tilde{k}_{t+1} = \tilde{k}_t = \tilde{k}$, $\tilde{\imath}_{t+1} = \tilde{\imath}_t = \tilde{\imath}$, $n_{t+1} = n$, thus we have

$$\frac{1+\gamma}{\beta} + \delta - 1 = \theta \frac{y}{k}. \tag{8.5}$$

FOC w.r.t. hours n_t is

$$\frac{\alpha}{1-\alpha} \frac{\tilde{c}_t}{1-n_t} = (1-\theta)\tilde{k}_t^{\theta} n_t^{-\theta} = (1-\theta)\frac{\tilde{y}_t}{n_t}.$$

Along BGP, this implies

$$\frac{\alpha}{1-\alpha} \frac{n}{1-n} = (1-\theta)\frac{y}{c}. \tag{8.6}$$

Finally, along BGP, the law of motion for the capital stock implies

$$
\begin{aligned}
(1+\eta)(1+\gamma)\frac{k}{y} &= (1-\delta)\frac{k}{y} + \frac{i}{y} \\
&\Rightarrow \\
\delta &= \frac{i}{k} + 1 - (1+\eta)(1+\gamma). \qquad (8.7)
\end{aligned}
$$

From the US data, the steady state investment/capital ratio is 0.076. Given the value of η and γ, equation (8.7) can be used to calibrate depreciation rate δ. The resulting value for $\delta = 0.048$.

Technology					**Preferences**			
θ	δ	ρ	σ_ε	γ	β	σ	α	η
0.40	0.012	0.95	0.007	0.0156	0.987	1	0.64	0.012

Table 8.5: Calibrated parameters

Once δ is known, equation (8.5) can be used to calibrate β. Given values for γ, δ, and θ, β is chosen to match the steady state output/capital ratio. Under our broad definition of capital stock and output, in data $\frac{k}{y} = 3.32$. This implies $\beta = 0.947$ or quarterly 0.987.

Equation (8.6) can be used to calibrate α. First notice that according to resource constraint

$$\begin{aligned} c_t + i_t &= y_t \\ &\Rightarrow \\ \frac{c_t}{k_t} + \frac{i_t}{k_t} &= \frac{y_t}{k_t} \\ \frac{c_t}{k_t} &= \frac{y_t}{k_t} - \frac{i_t}{k_t} \end{aligned}$$

We know the value for $\frac{y}{k}$ and $\frac{i}{k}$, hence we know the steady state $\frac{c}{k}$ ratio. We also have

$$\frac{y}{c} = \frac{y}{k}\frac{k}{c}$$

Since we know $\frac{y}{k}$ and $\frac{k}{c}$, we can obtain the value for $\frac{y}{c} = 1.33$. The next issue is what a reasonable estimate n is. We have to look at microeconomic evidence. From micro literature, researchers have found that HHs allocate about one-third of their discretinary time–i.e., time not spent sleeping or in personal maintenance–to market activities. Therefore, we pick value $n = 0.31$. This yields the value for $\alpha = 0.64$.

This leaves only two remaining parameters to be calibrated: ρ, σ_ε. Notice that

$$z_t = \log Y_t - \theta \log K_t - (1-\theta)\log N_t \tag{8.8}$$

We use a quarterly hours series based on the Established Survey for the labor input N_t. Given value of θ, we calculate Solow Residual z_t according to equation (8.8). And then we use a $AR(1)$ process to estimate ρ and σ_ε. We obtain

$$\rho = 0.95, \sigma_\varepsilon = 0.007.$$

We summarize the calibrated (quarterly) parameters in Table 8.5.

Solve the Model Numerically

Once we pick all the parameter values for the model economy, we can put the shock back and solve this model numerically. More generally, this kind of problem can be written in the following DP problem

$$\begin{aligned} v(s,z) &= \max\{r(z,s,d)+\beta E[v(s',z') \mid z]\} \\ & s.t. \\ s' &= B(s,z,d) \\ z' &= A(z)+\varepsilon'. \end{aligned}$$

Here s is the vector of state variable for this economy, d is the vector of decision (control) variables. z is the stochastic shock. $r(z,s,d)$ is the return function. In our case here, $s=(k,z), d=(k',n)$. We can use different computation methods to solve this DP problem (value function iteration, policy function iteration, transform it into a LQ problem,...). Most often, this will be done using linearization methods. In order to do this, recall that given an $AR(1)$ process for the stochastic chock

$$z' = \rho z + \varepsilon,$$

we can guess the policy functions are linear in the state variables $s=(k,z)$

$$k' = a_k + b_k k + c_k z \tag{8.9}$$

$$n = a_n + b_n k + c_n z. \tag{8.10}$$

The task of solving the model is then to determine the parameters $a_k, b_k, c_k, a_n, b_n, c_n$.

Simulate the Model economy and Compare with Data

Once we have discretized the state space $s=(k,z)$ and obtain the linearized decision rules as in (8.9)-(8.10), we can use a random number generator to simulate a realization (history) of the stochastic shock. This gives rise to a time series in each of the variables. These series are the researcher's artificial "dataset". Sample moments of the variables (in general, second moments) are computed and compared to actual data. The performance of the model is based on this comparison. In other words, we calibrate the model based on the consistency with growth facts. But we evaluate model based on its performance on business cycle facts.

How is our benchmark model doing in terms of replicating those stylized business cycle facts? Table 8.6 below (taken from Cooley and Prescott (1995) Table 1.2) presents the results of simulating the stochastic growth model using the parameter values discussed above.

Conclusion:

1. Technology shock can account for about 79% of the variations in output.

2. The model captures the feature that $\sigma_i >> \sigma_y$, but does not capture the magnitude.

3. The model captures the feature $\sigma_c < \sigma_y$, but misses the magnitude.

Variable	**SD (%)**	$\boldsymbol{\rho}$**(variable, y)**	**Data: SD (%)**	**Data:** $\boldsymbol{\rho}$**(variable, y)**
y	1.351	1.0	1.72	1.0
c	0.329	0.843	0.86	0.77
i	5.954	0.992	8.24	0.91
N	0.769	0.986	1.69	0.92
y/N	0.606	0.978	0.73	0.34

Table 8.6: Cyclical Behavior of the Artificial Economy: Model vs. Data

4. The model fails to capture $\sigma_N \approx \sigma_y$. This suggests that some important feature of the labor market is missing here in the model.

5. Consumption, investment, and hours are all strongly procyclical as in the data.

6. Failure in capturing $\rho(\frac{y}{N}, y)$.

What do we learn from this exercise? The broad features of the model economy suggest that it makes sense to claim that technology shock is the major driving force behind the business cycles. But the failures of the model tell us that there are important margins along which decisions are made that have not been captured in this simple world (especially in the labor market).

Chapter 9

Overlapping Generations Model

We have learned enough about the a major workhorse model of modern dynamic macroeconomics: infinite horizon growth (IHG) model. Here in this chapter we are going to study the second major workhorse of modern macroeconomics: the Overlapping Generations (OLG) model.

In terms of model setup, now the big differences bwtween OLG and IHG are:

1. HHs have *finite* horizon. Every period there is a new cohort born. And they do not live forever.

2. At any point of time, there are *heterogeneous* HHs in terms of age and allocations.

Therefore, a natural question can be raised is: Do these differences affect all the basic properties of the competitive equilibrium? The answer is yes. We can highlight some of the big differences here:

1. CE may no longer be PO.

2. There may exist a continuum of equilibria.

3. (Outside) money may have positive value.

What is the motivation for OLG model?

1. To make the model closer to the real world. Apparently HHs do not live forever and there is birth and death in every period.

2. We need models where agents undergo a life-cycle with low-income youth, high income middle ages, and retirement where labor income drops to zero. We want to analyze issues like social security, the effect of taxes

on retirement decisions, the distributive effects of taxes vs. government deficits, the effects of life-cycle saving on capital accumulation, educational policies, etc.

3. Because it has many interesting theoretical properties as we summarized some above, it is also an area of intense study among economic theorists.

9.1 A Pure Exchange Overlapping Generations Model

9.1.1 Basic Setup of the Model

Here we describe the basic setup of a pure exchange OLG model.

1. Time is discrete, $t = 0, 1, 2, ...$
2. In each period there is a single, nonstorable consumption good.
3. In each period t a new generation (of measure 1) is born, which we index by its date of birth and call it cohort t.
4. People live for two periods and then die. Therefore, the total population at each period is measure 1.

With this demographic structure, we have to distinguish the calendar time and the cohort age. Since now, we use subscript for the calendar time, and use superscript for the cohort index. For example, e_t^t is the time t endowment of the cohort born at time t, e_{t+1}^t is the time $t+1$ endowment of the cohort t. Similarly, (c_t^t, c_{t+1}^t) denotes the consumption allocation for the cohort t.

In period 1 there is an initial old generation 0 that has endowment e_1^0 and consumes c_1^0. In some of our applications we will endow the initial generation with an amount of outside money m. We will NOT assume $m \geq 0$. If $m \geq 0$, then m can be interpreted straightforwardly as fiat money, if $m < 0$ one should view this as that the initial old people having borrowed from some institution (which is, however, outside the model) and m is the amount to be repaid.

In Table 9.1 we demonstrate the demographic structure of this model economy. Note that there are both an infinite number of periods as well as well as an infinite number of agents in this economy. This "double infinity" has been cited to be the major source of the theoretical peculiarities of the OLG model.

Preferences of cohort t are assumed to be representable by an additively separable utility function of the form

$$U_t(c) = u(c_t^t) + \beta u(c_{t+1}^t).$$

The preference for the initial old cohort is

$$U_0(c) = u(c_1^0).$$

	Calendar Time								
	0	1	2	3	...	...	t	$t+1$	...
C	0	(c_1^0, e_1^0)							
o	1	(c_1^1, e_1^1)	(c_2^1, e_2^1)						
h	2		(c_2^2, e_2^2)	(c_3^2, e_3^2)					
o	3			(c_3^3, e_3^3)					
r	...				...	...			
t	...				...	...	(c_t^{t-1}, e_t^{t-1})		
	t						(c_t^t, e_t^t)	(c_{t+1}^t, e_{t+1}^t)	
	$t+1$							$(c_{t+1}^{t+1}, e_{t+1}^{t+1})$	
	...								...

Table 9.1: Demographic structure of OLG model

We assume that period utility function u is strictly increasing, strictly concave, and twice continuously differentiable. We also assume Inada condition holds, i.e., HHs want to consume positive amount of consumption in both periods. This completes the description of the economy.

We have following definition.

Definition 46 *An* ***allocation*** *in this pure exchange OLG economy is a sequence* $\{c_t^t, c_{t+1}^t\}_{t=1}^{\infty} \cup c_1^0$.

Definition 47 *An allocation is* ***feasible*** *if* $c_t^{t-1}, c_t^t \geq 0$ *for all* $t \geq 1$ *and*

$$c_t^{t-1} + c_t^t = e_t^{t-1} + e_t^t, \forall t \geq 1.$$

Definition 48 *An allocation is* ***Pareto Optimal*** *(PO) if it is feasible and if there is no other feasible allocation* $\{\hat{c}_t^t, \hat{c}_{t+1}^t\}_{t=1}^{\infty} \cup \hat{c}_1^0$ *such that*

$$\begin{aligned} U_t(\hat{c}_t^t, \hat{c}_{t+1}^t) &\geq U_t(c_t^t, c_{t+1}^t), \forall t \geq 1 \\ u(\hat{c}_1^0) &\geq u(c_1^0) \end{aligned}$$

with strict inequality for at least one $t \geq 0$.

Let p_t be the price of one unit of the consumption good at period t. In the presence of money (i.e., $m \neq 0$) we will take money to be the numeraire.

9.1.2 Definition of Competitive Equilibrium

Now we can define a CE for this economy.

Definition 49 *Given money* m*, an ADE for this economy is an allocation* $\{c_t^t, c_{t+1}^t\}_{t=1}^{\infty} \cup c_1^0$ *and prices* $\{p_t\}_{t=1}^{\infty}$ *such that* *(i). Given prices* $\{p_t\}_{t=1}^{\infty}$*, for each* $t \geq 1$*,* $\{c_t^t, c_{t+1}^t\}_{t=1}^{\infty}$ *solves cohort* t*'s problem*

$$\begin{aligned} & \max_{\{c_t^t, c_{t+1}^t\}} u(c_t^t) + \beta u(c_{t+1}^t) \\ & s.t. \\ p_t c_t^t + p_{t+1} c_{t+1}^t & \leq p_t e_t^t + p_{t+1} e_{t+1}^t \\ c_t^t, c_{t+1}^t & \geq 0 \end{aligned}$$

(ii). Given p_1*,* c_1^0 *solves the initial old's problem*

$$\begin{aligned} & \max_{c_1^0} u(c_1^0) \\ & s.t. \\ p_1 c_1^0 & \leq p_1 e_1^0 + m \\ c_1^0 & \geq 0. \end{aligned}$$

(iii). Market clears for every $t \geq 1$*.*

$$c_t^{t-1} + c_t^t = e_t^{t-1} + e_t^t, \forall t \geq 1.$$

As in the infinite horizon model case, we can also define CE in sequential market setting for OLG economy. We assume that contracts between agents specifying one-period loans are enforceable. We difine s_t^t be the loans of the cohort t from period t to period $t+1$. And we let r_{t+1} denote the gross interest rate for loans granted at period t and maturing at $t+1$.

We define a SME as follows.

Definition 50 *Given money* m*, an SME for this economy is an allocation* $\{c_t^t, c_{t+1}^t, s_t^t\}_{t=1}^{\infty} \cup c_1^0$ *and interest rates* $\{r_t\}_{t=1}^{\infty}$ *such that*

(i). Given prices $\{r_t\}_{t=1}^{\infty}$*, for each* $t \geq 1$*,* $\{c_t^t, c_{t+1}^t, s_t^t\}_{t=1}^{\infty}$ *solves cohort* t*'s problem*

$$\begin{aligned} & \max_{\{c_t^t, c_{t+1}^t, s_t^t\}} u(c_t^t) + \beta u(c_{t+1}^t) \\ & s.t. \\ c_t^t + s_t^t & \leq e_t^t \quad (9.1) \\ c_{t+1}^t & \leq e_{t+1}^t + (1 + r_{t+1}) s_t^t \quad (9.2) \\ c_t^t, c_{t+1}^t & \geq 0. \end{aligned}$$

(ii). Given r_1*,* c_1^0 *solves the initial old's problem*

$$\begin{aligned} & \max_{c_1^0} u(c_1^0) \\ & s.t. \\ c_1^0 & \leq e_1^0 + (1 + r_1) m \quad (9.3) \\ c_1^0 & \geq 0. \end{aligned}$$

(iii). Market clears for every $t \geq 1$.

$$c_t^{t-1} + c_t^t = e_t^{t-1} + e_t^t, \forall t \geq 1. \tag{9.4}$$

Remark 51 *Unlike IHG model, here we do not need to impose no-Ponzi-game restriction. In IHG model, HHs could keep on building up debt forever (because they never die), unless they are prohibited to do so. But now in OLG setting, the non-negativity constraint of consumption actually imposes an implicit restriction on borrowing $s_t^t \geq -\frac{e_{t+1}^t}{1+r_{t+1}}$. That is, they have to repay their loans before the death.*

Remark 52 *Due to the setup of the model, HHs are assumed to only live up to two periods. Because of this, there is no inter-generational borrowing and lending in the model.*

9.1.3 Characterization of the Equilibrium Solutions

Equivalence between ADE and SME

Let's solve the ADE first. FOCs to the cohort t's peoblem are

$$\begin{aligned} c_t^t &: \quad u'(c_t^t) = \lambda p_t \\ c_{t+1}^t &: \quad \beta u'(c_{t+1}^t) = \lambda p_{t+1} \end{aligned}$$

Combining these two FOCs, we have

$$\frac{p_{t+1}}{p_t} = \frac{\beta u'(c_{t+1}^t)}{u'(c_t^t)}.$$

Next, let's solve SME. FOCs to the cohort t's problem are

$$\begin{aligned} c_t^t &: \quad u'(c_t^t) = \lambda_t \\ c_{t+1}^t &: \quad \beta u'(c_{t+1}^t) = \lambda_{t+1} \\ s_t^t &: \quad \lambda_{t+1}(1 + r_{t+1}) = \lambda_t \end{aligned}$$

Combining these three FOCs, we have

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

If we define the A-D price sequence $\{p_t\}_{t=1}^{\infty}$ in ADE in following way

$$p_1 = \frac{1}{1 + r_1} \tag{9.5}$$

$$p_{t+1} = (\frac{1}{1 + r_{t+1}})p_t \tag{9.6}$$

Then ADE and SME end up with the identical EE.

Next we want to show that given this price sequence, the BC in SME and ADE are identical.

From equation (9.2), we know

$$s_t^t = \frac{1}{1+r_{t+1}}(c_{t+1}^t - e_{t+1}^t) \tag{9.7}$$

Substituting this expression into equation (9.1), we have

$$c_t^t + \frac{1}{1+r_{t+1}}c_{t+1}^t \leq e_t^t + \frac{1}{1+r_{t+1}}e_{t+1}^t$$

Multiply p_t on both sides, and using (9.6), we have

$$p_t c_t^t + p_{t+1}c_{t+1}^t \leq p_t e_t^t + p_{t+1}e_{t+1}^t$$

This is exactly is the BC in ADE for cohort $t \geq 1$.

For initial old, multiply p_1 on both sides of the BC for initial old (9.3), we obtain

$$p_1 c_1^0 \leq p_1 e_1^0 + p_1(1+r_1)m$$

Now recall (9.5), $p_1(1+r_1) = 1$, we end up with the same BC for initial old in ADE.

Since ADE and SME share the identical preferences, budget constraints, and first-order conditions, it is straightforward to show that the solutions to ADE and SME are identical. The equivalence between ADE and SME also holds true in the OLG setting.

Remark 53 *The equivalence of the two equilibrium definitions requires that the amount of loans that can be drawn is unrestricted (that is, that agents face no borrowing constraints). The reason is that we can switch from BC in SME to BC in ADE is we do not restrict the sign of s_t^t in equation (9.7). Suppose instead we impose a borrowing constraint*

$$s_t^t \geq b$$

where b is some number such that $b > -\frac{e_{t+1}^t}{1+r_{t+1}}$[1]. In this case, equation (9.7) becomes

$$s_t^t = \max\{\frac{1}{1+r_{t+1}}(c_{t+1}^t - e_{t+1}^t), b\}$$

We cannot have identical BCs any more.

---

[1] b cannot be smaller than $-\frac{e_{t+1}^t}{1+r_{t+1}}$ because if this is the case, it violates the non-negativity constraint for the consumption. Similarly, we cannot have $b = -\frac{e_{t+1}^t}{1+r_{t+1}}$ because if it is the case, we end up with $c_{t+1}^t = 0$. But Inada condition excludes thsi possibility.

We want to further characterize the allocation for savings in SME. The BC for cohort t at time t is

$$c_t^t + s_t^t = e_t^t.$$

And the BC for cohort $t-1$ at time t is

$$c_t^{t-1} = e_t^{t-1} + (1 + r_t)s_{t-1}^{t-1}.$$

Combining these two BCs

$$c_t^t + c_t^{t-1} + s_t^t = e_t^t + e_t^{t-1} + (1 + r_t)s_{t-1}^{t-1}$$

Using resource constraint (9.4), we obtain

$$s_t^t = (1 + r_t)s_{t-1}^{t-1}$$

This is the market clearing condition for loans.

For the initial old, we have $s_0^0 = m$. Thus we obtain $s_1^1 = (1+r_1)m$. Staring from this initial savings, after repeated substitution, we will obtain

$$s_t^t = \prod_{\tau=1}^{t}(1 + r_\tau)m \tag{9.8}$$

the amount of saving (in terms of the period t consumption good) has to equal the value of the outside supply of assets $\prod_{\tau=1}^{t}(1 + r_\tau)m$. Strictly speaking one should also include market clearing condition (9.8) in the definition of SME. However, by Walras' law, either the asset market or the good market clearing condition is redundant.

Notice that by the equivalence between SME and ADE, we have

$$\frac{p_t}{p_{t+1}} = 1 + r_{t+1}$$

Thus

$$\prod_{\tau=1}^{t}(1 + r_\tau) = \frac{p_0}{p_1} \cdot \frac{p_1}{p_2} \cdot ... \cdot \frac{p_{t-1}}{p_t}$$

If we normailze $p_0 = 1$, we have

$$\prod_{\tau=1}^{t}(1 + r_\tau) = \frac{1}{p_t}$$

Hence we obtain

$$\begin{aligned} s_t^t &= \prod_{\tau=1}^{t}(1 + r_\tau)m = \frac{m}{p_t} \\ &\text{or} \\ p_t s_t^t &= m. \end{aligned}$$

The present value of savings (in terms of time 1 consumption) should be equal to the outside supply of assets m.

An Example Equilibrium

Let's consider the following OLG economy. For ecvery cohort t, the preference is represented by utility function (we ignore discount factor β for simplicity):

$$\log c_t^t + \log c_{t+1}^t.$$

The endowment is given by

$$\begin{aligned} e_t^t &= \omega_y \\ e_{t+1}^t &= \omega_o \end{aligned}$$

for all t. Cohort t now sloves the problem

$$\begin{aligned} & \max_{c_t^t, c_{t+1}^t} \log c_t^t + \log c_{t+1}^t \\ & s.t. \\ c_t^t + \frac{1}{1+r_{t+1}} c_{t+1}^t \leq & \; e_t^t + \frac{1}{1+r_{t+1}} e_{t+1}^t \end{aligned}$$

We can substitute for c_{t+1}^t to transform the cohort's problem into an unconstrained one

$$\max_{c_t^t} \log c_t^t + \log[(e_t^t + \frac{1}{1+r_{t+1}} e_{t+1}^t - c_t^t)(1+r_{t+1})]$$

FOC is

$$\frac{1}{c_t^t} = \frac{1+r_{t+1}}{(1+r_{t+1})e_t^t + e_{t+1}^t - (1+r_{t+1})c_t^t}$$

This implies

$$c_t^t = \frac{1}{2}(\omega_y + \frac{1}{1+r_{t+1}}\omega_o) \tag{9.9}$$

$$c_{t+1}^t = \frac{1}{2}((1+r_{t+1})\omega_y + \omega_o). \tag{9.10}$$

The initial old's problem is

$$\begin{aligned} & \max \log c_1^0 \\ & s.t. \\ c_1^0 \leq & \; \omega_o \end{aligned}$$

Obviously the solution for the initial old is $c_1^0 = \omega_o$. Recall the resource constraint of period 1 is

$$c_1^0 + c_1^1 = \omega_y + \omega_o$$

Thsi implies $c_1^1 = \omega_y$. Repeatedly using the market clearing condition (or resource constraint), we have

$$c_t^t = \omega_y, c_{t+1}^t = \omega_0$$

Given this allocation, we go back to solution equation (9.9) and (9.10) to find out the price r_{t+1} that supports the allocation as a solution in CE. We obtain

$$r_{t+1} = \frac{\omega_0}{\omega_y} - 1.$$

Equivalently in ADE, the AD price will be

$$p_t = \frac{1}{\prod_{\tau=1}^{t}(1 + r_\tau)} = (\frac{\omega_y}{\omega_0})^t.$$

This constant price sequence supports the equilibrium where HHs do not trade: they just consume their initial endowments. We call this equilibrium "autarky".

9.1.4 Analysis of the Model Using Offer Curves

The Offer Curves

But actually the example economy above may have other equilibria. To construct and analyze all equilibria, we summarize preferences and consumption decisions in terms of an *offer curve.* We will use a graphical apparatus proposed by David Gale (1973). As in the example above, in thsi section, we consider the stationary economy with $e_t^t = \omega_y, e_{t+1}^t = \omega_o$.

What is an offer curve? We have following definition.

Definition 54 *The HH's offer curve (OC) is the locus of* (c_t^t, c_{t+1}^t) $\forall t \geq 1$ *that solves the cohort t's problem*

$$\begin{aligned} & \max_{\{c_t^t, c_{t+1}^t\}} u(c_t^t) + \beta u(c_{t+1}^t) \\ & s.t. \\ c_t^t + \frac{p_{t+1}}{p_t} c_{t+1}^t \quad \leq \quad & e_t^t + \frac{p_{t+1}}{p_t} e_{t+1}^t \end{aligned}$$

Evidently, the OC solves the following pair of equations

$$\begin{aligned} c_t^t + \frac{p_{t+1}}{p_t} c_{t+1}^t &= e_t^t + \frac{p_{t+1}}{p_t} e_{t+1}^t \quad \text{(BC)} \\ \frac{p_{t+1}}{p_t} &= \frac{\beta u'(c_{t+1}^t)}{u'(c_t^t)} \quad \text{(FOC)} \end{aligned}$$

for $\frac{p_{t+1}}{p_t} > 0$. We denote the OC by

$$\psi(c_t^t, c_{t+1}^t) = 0$$

Apparently, OC solves allocation (c_t^t, c_{t+1}^t) as a function of relative A-D price ratio $\frac{p_{t+1}}{p_t}$ (it is also the reciprocal of the one-period gross rate of return from period t to $t+1$). Therefore, if we draw a graph with c_t^t on the x-axis and

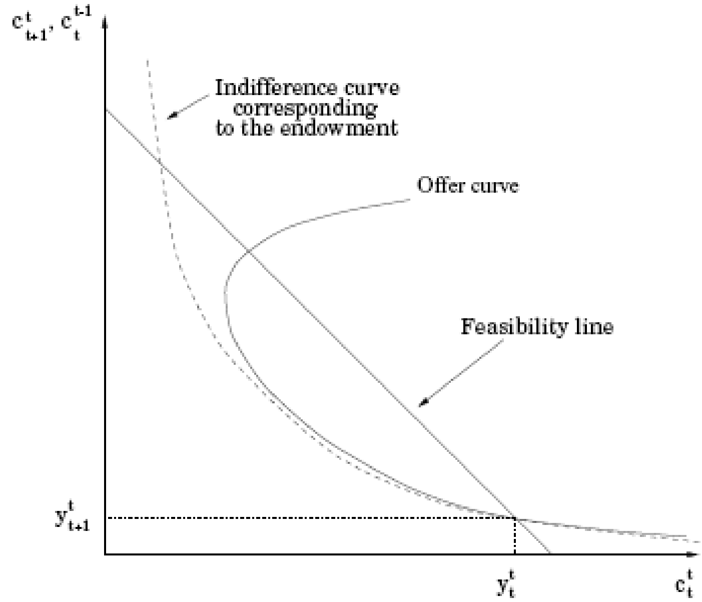


Figure 9.1: The offer curve and feasibility line

c_{t+1}^t on the y-axis and vary $\frac{p_{t+1}}{p_t}$, we will have a resutling locus for c_t^t and c_{t+1}^t, that is the OC. Every point on OC corresponds to a tangent point between the indifference curve and the BC line.

Following Gale (1973), we can use the offer curve and a straight line depicting feasibility (resource constraint) in the (c_t^t, c_{t+1}^t) plane to construct a machine for computing equilibrium allocations and prices. In particular, we can use the following pair of difference equations to solve for an equilibrium allocation $\forall t \geq 1$

$$\begin{aligned} \psi(c_t^t, c_{t+1}^t) &= 0 \\ c_t^{t-1} + c_t^t &= e_t^{t-1} + e_t^t \end{aligned}$$

See Figure 9.1 for the graph. Note that the slope of the feasibility line is -1. By the definition of CE, the intersection between the OC and feasibility line is an equilibrium. Also note that the endowment point (e_t^t, e_{t+1}^t) is on the feasibility line (it is feasible) and the OC (as we showed above, autarky is an equilibrium). For any point other than endowment point, draw a straight line through both points, the slope of that line equals $-\frac{p_t}{p_{t+1}}$, and this line is the BC given this price ratio $\frac{p_{t+1}}{p_t}$.

Computing Equilibria using Offer Curves

A procedure for finding an equilibrium is illustrated in Figure 9.2, which reproduces a version of a graph of David Gale (1973). Let's start with a proposed initial old consumption c_1^0.[2] Then go horizontally to the feasibility line to find the maximal feasible value for c_1^1 on x-axis, the time 1 allocation to the cohort born at time 1. The candidate time 1 allocation is thus feasible, but the time 1 young will choose c_2^1 only if the price $\frac{p_2}{p_1}$ is such that (c_1^1, c_2^1) lies on the offer curve. Therefore, we choose c_2^1 from the point on the offer curve that cuts a vertical line through c_1^1. Then we proceed to find c_2^2 from the intersection of a horizontal line through c_2^1 and the feasibility line. We continue recursively in this way, choosing c_t^t as the intersection of the feasibility line with a horizontal line through c_t^{t-1}, then choosing c_{t+1}^t as the intersection of a vertical line through c_t^t and the offer curve. In this way, the entire equilibrium consumption allocation path can be constructed. We can also construct a sequence of price $\frac{p_{t+1}}{p_t}$ from the slope of a straight line through the endowment point and the sequence of (c_t^t, c_{t+1}^t) pairs that lie on the offer curve. Thus using this procedure, we construct an entire equilibrium graphically.

If the OC has the shape drawn in Figure 9.2, any initial old consumption c_1^0 between the upper and lower intersections of the OC and the feasibility line is an equilibrium setting of c_1^0. Each such c_1^0 is associated with a distinct allocation (c_t^t, c_{t+1}^t) and a price sequence, all but one (the one on the upper intersection of OC and feasibility line) of them converging to the autarky equilibrium allocation and interest rate. Thus this endowment economy has a continuum of equilibria.

[2] Readers may notice that the initial old consumption we chose is more than the endowment for initial old $e_1^0 = \omega_o$. this is because initial old is endowed with some fiat money $\frac{m}{p_1} > 0$.

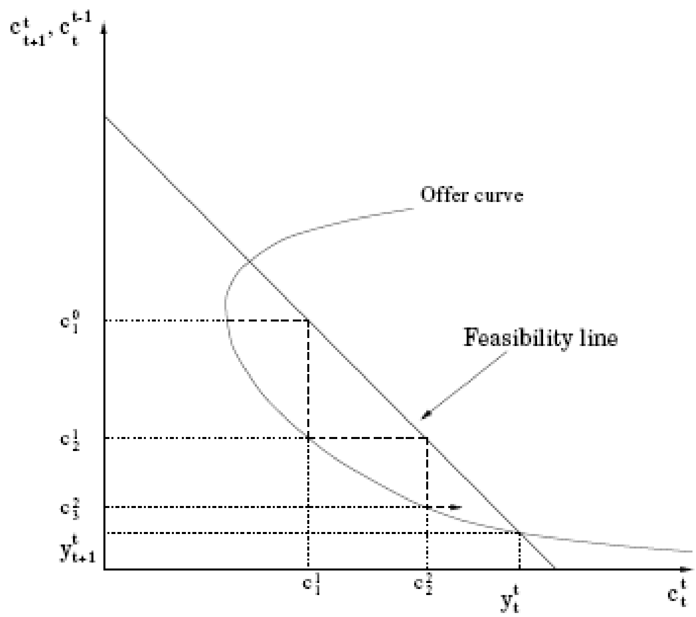


Figure 9.2: Computing an equilibrium using Offer Curve

Remark 55 *For the economy without money ($m = 0$), the autarkic equilibrium is the unique equilibrium for this economy. This is easily seen. Since the initial old generation has no money, only its endowments ω_o, there is no way for them to consume more than their endowments. Obviously they can always assure to consume at least their endowments by not trading, and that is what they do for any $p_1 > 0$ (obviously $p_1 \leq 0$ is not possible in equilibrium). But then from the resource constraint it follows that the first young generation must consume their endowments when young. Since they haven't saved anything, the best they can do when old is to consume their endowment again. But then the next young generation is forced to consume their endowments and so forth. Trade breaks down completely. For this allocation to be an equilibrium prices must be such that at these prices all generations actually find it optimal not to trade, which yields the price sequence $\frac{p_{t+1}}{p_t} = \frac{\beta u'(\omega_o)}{u'(\omega_y)}$.*

Remark 56 *In general, the price ratio supporting the autarkic equilibrium satisfies*

$$\frac{p_{t+1}}{p_t} = \frac{\beta u'(e^t_{t+1})}{u'(e^t_t)} = \frac{\beta u'(\omega_o)}{u'(\omega_y)}.$$

It is also the slope of the OC at the endowment point. Accordingly, the autarkic interest rate (recall the equivalence result we derived above) is determined by

$$\frac{1}{1 + r_{AUT}} = \frac{\beta u'(\omega_o)}{u'(\omega_y)}.$$

*We cal this equation "**Condition of Efficiency**". When $\beta = 1$, this condition reduces to*

$$\frac{1}{1 + r_{AUT}} = \frac{u'(\omega_o)}{u'(\omega_y)}.$$

Gale (1973) has invented the following terminology: If $\omega_o \geq \omega_y$, $\frac{1}{1+r_{AUT}} < 1$ (recall the utility function is strictly concave), hence $r_{AUT} \geq 0$, he calls it "Classical Case"; otherwise if $\omega_o < \omega_y$, $\frac{1}{1+r_{AUT}} > 1 \Rightarrow r_{AUT} < 0$, he calls it "Samuelson Case". As it will turn out and will be demonstrated below autarkic equilibria are not Pareto optimal in the Samuelson case while they are in the classical case.

9.1.5 Pareto Optimality in OLG Model

The preceding example can also serve to demonstrate our first major feature of OLG economies that sets it apart from the standard infinite horizon model with finite number of agents: competitive equilibria may be not be Pareto optimal. In the discuss below, we restrict our attention on the OLG model without money, i.e., $m = 0$.

Definition 57 *An equilibrium is stationary if $c^t_t = c_y$, $c^t_{t+1} = c_o$, and $\frac{p_{t+1}}{p_t}$ =constant.*

Obviously, the autarky equilibrium is a stationary equilibrium. But if autarky equilibrium is PO?

Theorem 58 *Autarky Equilibrium* ($c_t^t = \omega_y, c_{t+1}^t = \omega_o$) *is efficient (PO) iff* $r_{AUT} \geq 0$.

Proof. To establish sufficiency part (if $r_{AUT} \geq 0$, the autarky equilibrium is PO), we prove by contradiction. Suppose there exists another feasible allocation $\tilde{C} = (\tilde{c}_t^t, \tilde{c}_{t+1}^t)$ that Pareto dominates the autarky equilibrium and $r_{AUT} \geq 0$. Since it is feasible, it has to satisfy the resource constraint

$$\tilde{c}_t^t + \tilde{c}_t^{t-1} = \omega_y + \omega_o.$$

Let t be the first period when this alternative allocation $\tilde{C}$ differs from the autarkic allocation. The requirement that the old generation in this period t is not made worse off, $\tilde{c}_t^{t-1} \geq \omega_o$, implies $\tilde{c}_t^t < \omega_y$ (think about why we cannot have $\tilde{c}_t^t = \omega_y$). But since time t, $\tilde{C}$ is different and Pareto dominates allocation $C = (c_t^t, c_{t+1}^t)$, therefore, we must have $u(\tilde{c}_t^t) + u(\tilde{c}_{t+1}^t) \geq u(c_t^t) + u(c_{t+1}^t)$, which implies $\tilde{c}_{t+1}^t > \omega_o$. Feasibility constraint then implies $\tilde{c}_{t+1}^{t+1} < \omega_y$. Let's define

$$\varepsilon_{t+1} \equiv \omega_y - \tilde{c}_{t+1}^{t+1} > 0 \tag{9.11}$$

Non-negativity constraint on consumption also implies $\varepsilon_{t+1} \leq \omega_y$.

Now given $\tilde{c}_{t+1}^{t+1}$, we compute the smallest number c_{t+2}^{t+1} such that

$$u(\tilde{c}_{t+1}^{t+1}) + u(c_{t+2}^{t+1}) = u(\omega_y) + u(\omega_o).$$

Let $\bar{c}_{t+2}^{t+1}$ be the solution to this problem. Since $\tilde{C}$ Pareto dominates the autarky equilibrium, we have

$$\tilde{c}_{t+2}^{t+1} \geq \bar{c}_{t+2}^{t+1}.$$

Now we want to derive a convenient expression for $\bar{c}_{t+2}^{t+1}$. Consider the indifference curve of preferences over (c_y, c_o) that yields a fixed utility level $u(\omega_y) + u(\omega_o)$ as in autrakic equilibrium

$$u(c_y) + u(c_o) = u(\omega_y) + u(\omega_o)$$

This indifference curve defines consumption when old as a function of consumption when young. $c_o = h(c_y)$ where $h' = MRS_{y,o} = -\frac{u'(c_y)}{u'(c_o)}$. Define $\theta = \frac{u'(c_y)}{u'(c_o)}$. Notice that $h'' > 0$. Therefore, applying Taylor's Theorem to h, we have

$$\begin{aligned} c_o &= h(c_y) \\ &= h(\omega_y) + (\omega_y - c_y)[-h'(\omega_y) + \frac{1}{2}h''(\omega_y)(\omega_y - c_y)] \end{aligned}$$

Define

$$f(\omega_y - c_y) = \frac{1}{2}h''(\omega_y)(\omega_y - c_y)$$

Notice that $f(\cdot)$ is strictly increasing and $f(0) = 0$. Thus we have

$$h(c_y) = h(\omega_y) + (\omega_y - c_y)[-h'(\omega_y) + f(\omega_y - c_y)]. \tag{9.12}$$

Now allocation $(\tilde{c}_{t+1}^{t+1}, \bar{c}_{t+2}^{t+1})$ and $(c_{t+1}^{t+1} = \omega_y, c_{t+2}^{t+1} = \omega_o)$ are on the same indifference curve h, we may use (9.11) and (9.12) to write

$$\begin{aligned}\bar{c}_{t+2}^{t+1} &= h(\tilde{c}_{t+1}^{t+1}) \\ &= h(\omega_y) + (\omega_y - \tilde{c}_{t+1}^{t+1})[-h'(\omega_y) + f(\omega_y - \tilde{c}_{t+1}^{t+1})] \\ &= h(\omega_y) + \varepsilon_{t+1}[-h'(\omega_y) + f(\varepsilon_{t+1})]\end{aligned}$$

Notice that $\omega_o = h(\omega_y)$ and $\theta_{AUT} = -h'(\omega_y)$, thus we have

$$\bar{c}_{t+2}^{t+1} = \omega_o + \varepsilon_{t+1}[\theta_{AUT} + f(\varepsilon_{t+1})].$$

Since $\tilde{c}_{t+2}^{t+1} \geq \bar{c}_{t+2}^{t+1}$, we have

$$\tilde{c}_{t+2}^{t+1} - \omega_o \geq \varepsilon_{t+1}[\theta_{AUT} + f(\varepsilon_{t+1})] \tag{9.13}$$

Since $\tilde{C}$ is also feasible, $(\tilde{c}_{t+2}^{t+1}, \tilde{c}_{t+2}^{t+2})$ satisfies the feasibility constraint

$$\tilde{c}_{t+2}^{t+1} + \tilde{c}_{t+2}^{t+2} = \omega_y + \omega_o.$$

Thus we have

$$\begin{aligned}\varepsilon_{t+2} &\equiv \omega_y - \tilde{c}_{t+2}^{t+2} = \tilde{c}_{t+2}^{t+1} - \omega_o \\ &\geq \varepsilon_{t+1}[\theta_{AUT} + f(\varepsilon_{t+1})]\end{aligned}$$

Since $\theta_{AUT} = \frac{u'(\omega_y)}{u'(\omega_o)} = 1 + r_{AUT} \geq 1$ and $f(\varepsilon_{t+1}) > 0$ (implied by $\varepsilon_{t+1} > 0$ and f is strictly increasing). We have

$$\varepsilon_{t+2} \geq \varepsilon_{t+1}[\theta_{AUT} + f(\varepsilon_{t+1})] > \varepsilon_{t+1}$$

Continuing these computations of successive values of ε_{t+k} yields

$$\varepsilon_{t+k} \geq \varepsilon_{t+1} \prod_{j=1}^{k-1} (\theta_{AUT} + f(\varepsilon_{t+j})) > \varepsilon_{t+1}[\theta_{AUT} + f(\varepsilon_{t+1})]^{k-1}, \forall k > 2,$$

where the strict inequality comes from the fact that $\{\varepsilon_{t+j}\}_{j=1}^{k}$ is a strictly increasing sequence and f is strictly increasing. Since $\theta_{AUT} + f(\varepsilon_{t+1}) > 1$, the conclusion is that the ε sequence is bounded below by a strictly increasing exponential and hence is *unbounded*. This is a contradiction because we already know that ε cannot exceed the endowment ω_y. Done with suffcicency.

To establish necessity, we will prove that when $r_{AUT} < 0$, we can construct an alternative Pareto superior allocation. Deviating from autarky equilibrium, think about the following alternative allocation

$$\tilde{c}_t^t = \omega_y - \varepsilon, \tilde{c}_{t+1}^t = \omega_o + \varepsilon, \forall t \geq 1$$

where ε is an arbitrarily small positive number. First of all, we can check this alternative allocation is feasible since

$$\tilde{c}_t^t + \tilde{c}_{t+1}^t = \omega_y + \omega_o$$

Second, we can show that it makes every cohort $t \geq 1$ strictly better off.

$$\begin{aligned} & [u(\tilde{c}_t^t) + u(\tilde{c}_{t+1}^t)] - [u(\omega_y) + u(\omega_o)] \\ = & [u(\omega_y - \varepsilon) - u(\omega_y)] + [u(\omega_o + \varepsilon) - u(\omega_o)] \\ = & -u'(\omega_y)\varepsilon + u'(\omega_o)\varepsilon \\ = & \varepsilon[u'(\omega_o) - u'(\omega_y)] \end{aligned}$$

Notice that $\frac{1}{1+r_{AUT}} = \frac{u'(\omega_o)}{u'(\omega_y)}$. If $r_{AUT} < 0$, $u'(\omega_o) > u'(\omega_y)$, which implies

$$u(\tilde{c}_t^t) + u(\tilde{c}_{t+1}^t) > u(\omega_y) + u(\omega_o), \forall t \geq 1$$

We can also make $\tilde{c}_1^0 = c_1^0 = \omega_o$, thus the initial old has the same utility as before. This alternative allocation then Pareto dominates the autarky equilibrium. ■

Why this condition $r_{AUT} \geq 0$ is important for Pareto optimality? What is the intuition behind this theorem? Note that in the autarky equilibrium

$$\frac{p_{t+1}}{p_t} = \frac{1}{1 + r_{AUT}} = \frac{u'(\omega_o)}{u'(\omega_y)}$$

If $r_{AUT} < 0$, we have $\frac{p_{t+1}}{p_t} > 1, \forall t \geq 1$. Thus the price sequence is strictly increasing and unbounded eventually, therefore we have $\sum_{t=1}^{\infty} p_t = +\infty$, this implies the value of the aggregate endowment is infinite. Therefore, the standard proof of FWT fails (if alternative allocation is better than the current one in terms of utility, then it will violate the resource constraint: value of aggregate consumption $\leq$ value of aggregate endowment. If the resource is not constrained at all, then how can we say it violates RC?), CE may not be PO.

We have following theorem which is equivalent to the theorem above.

Theorem 59 *(Balasko and Shell) A stationary autarkic equilibrium in an OLG endowment economy is PO iff*

$$\sum_{t=1}^{\infty} \frac{1}{p_t} = \sum_{t=1}^{\infty} \prod_{\tau=1}^{t} (1 + r_\tau) = +\infty.$$

9.1.6 Monetary Equilibrium in OLG Model

Now we want to demonstrate the second and third feature of OLG models that set it apart from standard Arrow-Debreu economies, namely the possibility of a continuum of equilibria and the fact that outside money may have positive value. We will see that, given the way we have defined our equilibria, these two issues are intimately linked. So now let us suppose that $m \neq 0$. In our discussion we will assume that $m > 0$, the situation for $m < 0$ is symmetric.

Although we have done it before, but let's repeat the definition of monetary equilibrium here with some modification.

Definition 60 *Given money m, an ADE for this economy is an allocation $\{c_t^t, c_{t+1}^t\}_{t=1}^{\infty} \cup c_1^0$ and prices $\{p_t, q_t\}_{t=1}^{\infty}$ such that*

(i). Given prices $\{p_t\}_{t=1}^{\infty}$, for each $t \geq 1$, $\{c_t^t, c_{t+1}^t\}_{t=1}^{\infty}$ solves cohort t's problem

$$\begin{aligned} & \max_{\{c_t^t, c_{t+1}^t\}} u(c_t^t) + \beta u(c_{t+1}^t) \\ & s.t. \\ p_t c_t^t + p_{t+1} c_{t+1}^t & \leq p_t e_t^t + p_{t+1} e_{t+1}^t \\ c_t^t, c_{t+1}^t & \geq 0 \end{aligned}$$

(ii). Given p_1 and q_1,, c_1^0 solves the initial old's problem

$$\begin{aligned} & \max_{c_1^0} u(c_1^0) \\ & s.t. \\ p_1 c_1^0 & \leq p_1 e_1^0 + q_1 m \\ c_1^0 & \geq 0. \end{aligned}$$

(iii). Markets clear for every $t \geq 1$.

$$\begin{aligned} c_t^{t-1} + c_t^t &= e_t^{t-1} + e_t^t \quad \textit{(goods market)} \\ q_t m &= p_t(e_t^t - c_t^t). \quad \textit{(money market)} \end{aligned}$$

We can also define a SME as follows. Let's denote p_{mt} be the value of a unit of money at time t in terms of consumption good ate time t. Also let $p_t \equiv \frac{1}{p_{mt}}$ be the price level at time t, i.e., the price of one unit of consumption good in terms of money. Notice that the difference between time t price level and A-D date-0 price p_t (a little notation abuse:-)).

Definition 61 *Given money m, an SME for this economy is an allocation $\{c_t^t, c_{t+1}^t, s_t^t\}_{t=1}^{\infty} \cup c_1^0$ and price $\{p_t\}_{t=1}^{\infty}$ such that*

(i). Given prices $\{p_t\}_{t=1}^{\infty}$, for each $t \geq 1$, $\{c_t^t, c_{t+1}^t, m_t^t\}_{t=1}^{\infty}$ solves cohort t's problem

$$\begin{aligned} & \max_{\{c_t^t, c_{t+1}^t, m_t^t\}} u(c_t^t) + \beta u(c_{t+1}^t) \\ & s.t. \\ c_t^t + \frac{m_t^t}{p_t} & \leq e_t^t \qquad (9.14) \\ c_{t+1}^t & \leq e_{t+1}^t + \frac{m_t^t}{p_{t+1}} \qquad (9.15) \\ c_t^t, c_{t+1}^t, m_t^t & \geq 0. \end{aligned}$$

(ii). Given p_1, c_1^0 solves the initial old's problem

$$\begin{aligned} & \max_{c_1^0} u(c_1^0) \\ & s.t. \\ c_1^0 &\leq e_1^0 + \frac{m}{p_1} \\ c_1^0 &\geq 0. \end{aligned} \tag{9.16}$$

(iii). Market clears for every $t \geq 1$.

$$c_t^{t-1} + c_t^t = e_t^{t-1} + e_t^t, \forall t \geq 1.$$

Combining BCs (9.14) and (9.15), we have for every $t \geq 1$

$$c_t^t + \frac{c_{t+1}^t}{p_t/p_{t+1}} = e_t^t + \frac{e_{t+1}^t}{p_t/p_{t+1}}$$

We still assume

$$e_t^t = \omega_y, e_{t+1}^t = \omega_o.$$

Thus we have consolidated BC

$$c_t^t + \frac{c_{t+1}^t}{p_t/p_{t+1}} = \omega_y + \frac{\omega_o}{p_t/p_{t+1}}.$$

Furthermore, let's assume log utility and $\beta = 1$, hence the cohort t's preference is

$$\log c_t^t + \log c_{t+1}^t$$

Given this preference, subject to the consolidated BC above, it is easy to show the solution to SME is

$$\begin{aligned} c_t^t &= \frac{1}{2}\left(\omega_y + \omega_o \frac{p_{t+1}}{p_t}\right), \\ c_{t+1}^t &= \frac{1}{2}\left(\omega_y + \omega_o \frac{p_{t+1}}{p_t}\right)\frac{p_t}{p_{t+1}}. \end{aligned}$$

Therefore the real demand for money of the young at time t is

$$\frac{m_t^t}{p_t} = \omega_y - c_t^t = \frac{1}{2}\omega_y - \frac{1}{2}\omega_o \frac{p_{t+1}}{p_t}$$

Imposing money market clearing condition (money supply is fixed)

$$m_t^t = m, \forall t \geq 1$$

We can recover the law of motion for prices in this economy:

$$p_{t+1} = p_t \frac{\omega_y}{\omega_o} - \frac{2m}{\omega_o}. \tag{9.17}$$

Consider following three cases:

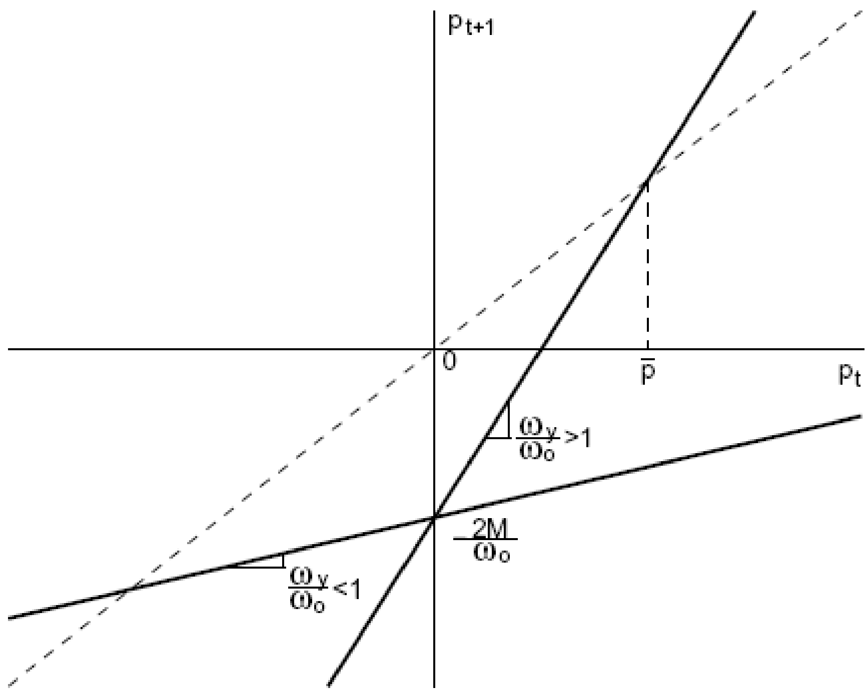


Figure 9.3: Price sequence in the monetary equilibrium

1. $\frac{\omega_y}{\omega_o} > 1$;
2. $\frac{\omega_y}{\omega_o} = 1$;
3. $\frac{\omega_y}{\omega_o} < 1$.

We can domonstrate these three cases for the first-order difference equation (9.17) in the graph below. (We show case 1 and 3, the case 2 in the graph should be a straight line goes through intercept $-\frac{2m}{\omega_o}$ and parallel to the 45 degree dash line.)

From this graph, we learn:

1. The only case consistent with positive and finite values of p_t is the first one, when $\omega_y > \omega_o$. In this case, there exists a solution $p_t = \bar{p} = \frac{2m/\omega_o}{\frac{\omega_y}{\omega_o} - 1} > 0, \forall t$. Money can have real value.

2. Money can overcome suboptimality. Recall that in the previous section, $\omega_y > \omega_o$ is called "Samuelson case". Without money in this case, the CE is not PO. By introducing money, we achieve the PO. (We have stationary equilibrium allocation $c_y = c_o = \frac{\omega_y+\omega_o}{2}$, $\frac{p_{t+1}}{p_t} = 1$, hence $MRS = 1$. By Balasko-Shell criterion, the resulting allocation is PO.)

3. There is no equilibrium with $p_1 < \bar{p}$, which means that one unit of money at $t = 1$ has value at most $\frac{1}{\bar{p}}$. It is easy to show that if $p_1 < \bar{p}$, then $c_1^0 > \frac{\omega_y+\omega_o}{2}$ which contradicts to the equilibrium allocation. (If $c_1^0 > \frac{\omega_y+\omega_o}{2}$, then according to feasibility constraint, $c_y = c_1^1 < \frac{\omega_y+\omega_o}{2}$, to make cohort 1 not worse off, we have to let $c_o = c_2^1 > \frac{\omega_y+\omega_o}{2}$, then the whole equilibrium path messes up.)

4. If $p_1 > \bar{p}$, there exists an equilibrium which the price sequence follows

$$p_{t+1} = p_t \frac{\omega_y}{\omega_o} - \frac{2m}{\omega_o}$$

with p_1 given. Equation above can be rewritten as

$$p_t = \frac{2m}{\omega_y} + \frac{\omega_o}{\omega_y} p_{t+1}$$

If we do repeated substitution forwardly, we obtain

$$\begin{aligned} p_t &= \frac{2m/\omega_o}{\frac{\omega_y}{\omega_o} - 1} + c(\frac{\omega_y}{\omega_o})^t \\ &= \bar{p} + c(\frac{\omega_y}{\omega_o})^t \end{aligned}$$

where scalar $c \geq 0$ is the terminal value of the price. $c = 0$ is the solution for the stationary equilibrium $p_t = \bar{p}$. $p_1 > \bar{p}$ implies $c > 0$. In this equilibrium, since $\omega_y > \omega_o$, when $t \to \infty$, $p_t \to \infty \Rightarrow p_{mt} \to 0$. This is an equilibrium with hyperinflation. Money loses value and goes to zero in the limit. Therfore, for any $p_1 \in (\bar{p}, +\infty)$, we can construct an equilibrium. This idea was already shown in the previous section by using offer curve technique.

5. $p_{mt} = 0$ ($p_t = +\infty$) is also an equilibrium. Money has zero value. The economy goes back to the one without money, and we all know that the only equilibrium there is autarky equilibrium.

These five points summarize the esstiential two differences between OLG and infinite horizon model.

- Money may have positive value in OLG. But this is not true in IHG.

Proposition 62 *In pure exchange economies with a finite number of infinitely lived agents there cannot be an equilibrium in which outside money is valued.*

Proof. Suppose, to the contrary, that there is an equilibrium allocation $\{(\hat{c}_t^i)_{i\in I}\}_{t=1}^{\infty}$ and $\{\hat{p}_t\}_{t=1}^{\infty}$ for initial endowments of outside money $(m^i)_{i\in I}$ such that $\sum_i m^i \neq 0$ and money as numeraire (hence $p_{m1} = 1$). Given the assumption of strict concavity of utility function, each consumer in equilibrium satisfies the A-D BC with equality

$$\sum_{t=1}^{\infty} \hat{p}_t \hat{c}_t^i = \sum_{t=1}^{\infty} \hat{p}_t e_t^i + m^i < \infty$$

Summing over all individuals $i \in I$ yields

$$\sum_{t=1}^{\infty} \hat{p}_t \sum_{i\in I} (\hat{c}_t^i - e_t^i) = \sum_{i\in I} m^i \tag{9.18}$$

But resource constraint requires for every period t

$$\sum_{i\in I} \hat{c}_t^i = \sum_{i\in I} e_t^i$$

Substituting into (9.18), we have

$$\sum_{i\in I} m^i = 0$$

Contradiction. This shows that there cannot exist an equilibrium in this type of economy in which outside money is valued in equilibrium. Note that this result applies to a much wider class of standard Arrow-Debreu economies than just the pure exchange economies considered in this section. ∎

While in OLG, when $\omega_y > \omega_o$, the initial old had endowment ω_o but consumed more than his endowment. He can do that because stock of outside money m he held is valued in equilibrium in that the old guys can exchange m pieces of intrinsically worthless paper for $\frac{m}{p_1} > 0$ units of period 1 consumption goods. The currently young generation accepts to transfer some of their endowment to the old people for pieces of paper because they *expect* (correctly so, in equilibrium) to exchange these pieces of paper against consumption goods when they are getting old, and hence to achieve an intertemporal allocation of consumption goods that dominates the autarkic allocation. Without the outside asset, again, this economy can do nothing else but remain in the possibly dismal state of autarky. The fact that money has value may be also seen as a "rational bubble"[3]: what people are willing to "pay" for money today depends on what they expect others will "pay" for it tomorrow. The role of money here is to mitigate the suboptimality present in the economy. It is the suboptimality that gives money positive value.

[3] In this equilibrium money is a "bubble". The fundamental value of an asset is the value of its dividends, evaluated at the equilibrium Arrow-Debreu prices. An asset is (or has) a bubble if its price does not equal its fundamental value. Obviously, since money doesn't pay dividends, its fundamental value is zero and the fact that it is valued positively in equilibrium makes it a bubble.

A deeper reason for this positively valued money in OLG is now we have "double infinity" in OLG, while in IHG only time is infinite, the agents are finite though. As we explain above, the "rational bubble" of money is supported by the expectation that the next cohort will always accept my money because the cohort after next cohort will do the same thing. This chain will continue forever because we have infinite agents here. But for IHG model, agents are finite, if this chain goes like the way in OLG, eventually there will be a last person who does not have another successor to accept his money. Therefore, he would not to accept the old guys' money when he is young. Go backward, no one will accept money from the previous cohort. Thus money have zero value.

- OLG can have a continuum of equilibria, while IHG only has finite number of equilibrium.

As we show above, there is an entire continuum of equilibria in OLG with positive valued money, indexed by $p_1 \in (\bar{p}, +\infty)$. These equilibria are arbitrarily close to each other, but they differ in the sense that at any finite time t, the consumption allocations and price ratios (and levels) differ across equilibria. As we showed in OC graph, eventually all nonstationary equilibria so constructed in the limit converge to the stationary autarkic equilibrium where $p_t = +\infty$. This is again in stark contrast to standard Arrow-Debreu economies where, generically, the set of equilibria is finite and all equilibria are locally unique.

Notice that when we are in the "Samuelson case", i.e., $\omega_y > \omega_o$, all these equilibria are Pareto-ranked. One can show, by similar arguments that demonstrated that the autarkic equilibrium is not Pareto optimal, that these equilibria are Pareto-ranked: let $p_1, \hat{p}_1 \in (\bar{p}, +\infty)$, then the equilibrium corresponding to $\hat{p}_1$ Pareto-dominates the equilibrium indexed by p_1. By the same token, the only Pareto optimal equilibrium allocation in this case is the nonautarkic stationary monetary equilibrium $p_t = \bar{p}$.

9.2 Government Policys in Overlapping Generations Model

We already know that from the previous section that autarky equilibrium is not PO. A natural question is can we has some welfare-improving government policy in this case? The answer is yes. The "pay-as-you-go" (PAYG) social security system is the one.

9.2.1 Pay-As-You-Go Social Security

Let's consider a simple pure exchange OLG model without money ($m = 0$) but with population growth. We assume population grows at constant rate n, so that for each old person in a given period there are $(1+n)$ young people around.

Equilibrium conditions are the same, except resource feasibility changes

$$c_t^{t-1} + (1+n)c_t^t = e_t^{t-1} + (1+n)e_t^t$$

Hence, in our offer curve diagram, the slope of the resource line is not -1 anymore, but $-(1+n)$.

Without any government intervention, we already know that the unique equilibrium is the autarkic equilibrium. Now let's introduce the government. What the government does in this economy is to introduce a PAYG social security system. Let's still assume stationary endowment $e_t^t = \omega_y, e_{t+1}^t = \omega_o$, $\forall t$. But now at each period the government imposes a social security tax $\tau \in (0, \omega_y)$ on the young and give social security benefits b to the old. We assume that the social security system balances its budget in each period, so that benefits are given by

$$b = \tau(1+n)$$

Therefore, we can define a SME for this economy as following

Definition 63 *An SME for this economy is an allocation* $\{c_t^t, c_{t+1}^t, s_t^t\}_{t=1}^{\infty} \cup c_1^0$, *a government policy* $\{\tau_t, b_t\}_{t=1}^{\infty}$, *and interest rates* $\{r_t\}_{t=1}^{\infty}$ *such that*

(i). Given prices $\{r_t\}_{t=1}^{\infty}$, *for each* $t \geq 1$, $\{c_t^t, c_{t+1}^t, s_t^t\}_{t=1}^{\infty}$ *solves cohort t's problem*

$$\begin{aligned} & \max_{\{c_t^t, c_{t+1}^t, s_t^t\}} u(c_t^t) + \beta u(c_{t+1}^t) \\ & s.t. \\ c_t^t + s_t^t & \leq e_t^t - \tau_t \\ c_{t+1}^t & \leq e_{t+1}^t + b_t + (1+r_{t+1})s_t^t \\ c_t^t, c_{t+1}^t & \geq 0. \end{aligned}$$

(ii). Given r_1, c_1^0 *solves the initial old's problem*

$$\begin{aligned} & \max_{c_1^0} u(c_1^0) \\ & s.t. \\ c_1^0 & \leq e_1^0 + b_1 \\ c_1^0 & \geq 0. \end{aligned}$$

(iii). Government budget constraint is balanced for every period

$$b_t = \tau_t(1+n)$$

(iii). Market clears for every $t \geq 1$.

$$c_t^{t-1} + (1+n)c_t^t = e_t^{t-1} + (1+n)e_t^t, \forall t \geq 1.$$

Again, we consider the stationary equilibrium with $c_t^t = c_y, c_{t+1}^t = c_o, r_t = r, b_t = b, \tau_t = \tau$. Obviously the new unique competitive equilibrium is again autarkic with endowments $(\hat{c}_t^t = \omega_y - \tau, \hat{c}_{t+1}^t = \omega_o + \tau(1+n))$ and equilibrium interest rates

$$1 + \hat{r}_{t+1} = 1 + \hat{r} = \frac{u'(\omega_y - \tau)}{\beta u'(\omega_o + \tau(1+n))}.$$

Now let's show that under what condition this new qutarkic equilibrium with PAYG is Pareto-improving to the original autarkic equilibrium ($c_t^t = \omega_y, c_{t+1}^t = \omega_o$).

Obviously for the initial old, $\hat{c}_1^0 = \omega_o + \tau(1+n) > \omega_o = c_1^0$, the initial old is strictly better off. For all other generations, define the equilibrium lifetime utility, as a function of the social security system, as

$$V(\tau) = u(\omega_y - \tau) + \beta u(\omega_o + \tau(1+n))$$

The introduction of a small social security system is welfare improving if and only if $V'(\tau)$, evaluated at $\tau = 0$, is positive. But we have

$$V'(\tau) = -u'(\omega_y - \tau) + \beta u'(\omega_o + \tau(1+n))(1+n)$$

Evaluating at $\tau = 0$, we have

$$V'(0) = -u'(\omega_y) + \beta u'(\omega_o)(1+n)$$

Hence $V'(0) > 0$ iff

$$n > \frac{u'(\omega_y)}{\beta u'(\omega_o)} - 1 = r_{AUT}$$

where r_{AUT} is the autarkic interest rate. Hence the introduction of a (marginal) PAYG social security system is welfare improving if and only if the population growth rate exceeds the equilibrium (autarkic) interest rate, or, to use our previous terminology, if we are in the Samuelson case where autarky is not a Pareto optimal allocation. Note that social security has the same function as money in our economy: it is a social institution that transfers resources between generations (backward in time) that do not trade among each other in equilibrium. In enhancing intergenerational exchange not provided by the market it may generate allocations that are Pareto superior to the autarkic allocation, in the case in which individuals private marginal rate of substitution $1+r_{AUT}$ (at the autarkic allocation) falls short of the social intertemporal rate of transformation $1+n$.

9.2.2 The Ricardian Equivalence Hypothesis

Question: Whether the timing of taxes matters?

Answer: Under some assumptions it does, and under others it does not.

The Ricardian doctrine describes assumptions under which the timing of lump taxes does not matter. In this subsection, we will study how the timing of taxes interacts with restrictions on the ability of households to borrow. We study the issue in two equivalent settings: (1) an infinite horizon economy with an infinitely lived representative agent; and (2) an infinite horizon economy with a sequence of finite-period-lived agents, each of whom cares about its immediate descendant.

Infinite Lifetime Horizon and Borrowing Constraints

The Ricardian Equivalence hypothesis is, in fact, a theorem that holds in a fairly wide class of models. It is most easily demonstrated within the Arrow-Debreu market structure of infinite horizon models. Consider the simple infinite horizon pure exchange model with N identical HHs. The index of HH is $i \in I = \{1, 2, ..., N\}$. The preference of HH i is

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

Now introduce a government that has to finance a given exogenous stream of government expenditures (in real terms) denoted by $\{G_t\}_{t=0}^{\infty}$. These government expenditures do not yield any utility to the agents (this assumption is not at all restrictive for the results to come). Let p_t denote the Arrow-Debreu price at date 0 of one unit of the consumption good delivered at period t. The government has initial outstanding real debt of B_1 that is held by the public. Let b_1^i denote the initial endowment of government bonds of agent i. Obviously we have the restriction

$$\sum_{i \in I} b_1^i = B_1$$

In order to finance the government expenditures the government levies lump-sum taxes: let τ_t^i denote the taxes that agent i pays in period t, denoted in terms of the period t consumption good. We define an Arrow-Debreu equilibrium with government as follows.

Definition 64 *Given a sequence of government spending $\{G_t\}_{t=0}^{\infty}$ and initial government debt B_1 and initial individual bond holding $(b_1^i)_{i \in I}$, an ADE is an allocation $\{(c_t^i)_{i \in I}\}_{t=0}^{\infty}$, a price system $\{p_t\}_{t=0}^{\infty}$, and a sequence of taxes $\{(\tau_t^i)_{i \in I}\}_{t=0}^{\infty}$ such that*

(i). Given the price and the sequence of taxes, $\{(c_t^i)_{i \in I}\}_{t=0}^{\infty}$ solves HH i's problem

$$\max_{\{(c_t^i)_{i \in I}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(c_t^i)$$

$$s.t.$$

$$\sum_{t=0}^{\infty} p_t(c_t^i + \tau_t^i) \leq \sum_{t=0}^{\infty} p_t e_t^i + p_1 b_1^i. \tag{9.19}$$

(ii). Given the price, government tax policy $\{(\tau_t^i)_{i \in I}\}_{t=0}^{\infty}$ satisfies

$$\sum_{t=0}^{\infty} p_t G_t + p_1 B_1 = \sum_{t=0}^{\infty} \sum_{i=1}^{N} p_t \tau_t^i \tag{9.20}$$

(iii). Market clears $\forall t \geq 0$.

$$\sum_{i=1}^{N} c_t^i + G_t = \sum_{i=1}^{N} e_t^i. \tag{9.21}$$

Remark 65 *Muitiply A-D price p_t on both sides of the resource constraint equation (9.21), use government BC () to substitute out $\sum_{t=0}^{\infty} p_t G_t$ and use $\sum_{i \in I} b_1^i = B_1$ to substitute out $p_1 B_1$, we will end up with HH's BC (9.19).*

In an Arrow-Debreu definition of equilibrium the government, as the agent, faces a single intertemporal budget (9.20) constraint which states that the total value of tax receipts is sufficient to finance the value of all government purchases plus the initial government debt. From the definition it is clear that, with respect to government tax policies, the only thing that matters is the total value of taxes $\sum_{t=0}^{\infty} \sum_{i=1}^{N} p_t \tau_t^i$ that the individual has to pay, but not the timing of taxes. It is then straightforward to prove the Ricardian Equivalence theorem for this economy.

Theorem 66 *(Ricardo Equivalence) Take as given a sequence of government spending $\{G_t\}_{t=0}^{\infty}$ and initial government debt B_1 and initial endowment of government bond $(b_1^i)_{i \in I}$, Suppose that allocation $\{(c_t^i)_{i \in I}\}_{t=0}^{\infty}$, prices $\{p_t\}_{t=0}^{\infty}$, and a sequence of taxes $\{(\tau_t^i)_{i \in I}\}_{t=0}^{\infty}$ consist an ADE. Let $\{(\hat{\tau}_t^i)_{i \in I}\}_{t=0}^{\infty}$ be an arbitrary alternative tax system satisfying*

$$\sum_{t=0}^{\infty} \sum_{i=1}^{N} p_t \tau_t^i = \sum_{t=0}^{\infty} \sum_{i=1}^{N} p_t \hat{\tau}_t^i. \tag{9.22}$$

Then allocation $\{(c_t^i)_{i \in I}\}_{t=0}^{\infty}$, prices $\{p_t\}_{t=0}^{\infty}$, and a sequence of taxes $\{(\hat{\tau}_t^i)_{i \in I}\}_{t=0}^{\infty}$ also consist an ADE.

Proof. Summing HH i's BC over i, we have

$$\sum_{t=0}^{\infty} \sum_{i=1}^{N} p_t (c_t^i + \tau_t^i) \leq \sum_{t=0}^{\infty} \sum_{i=1}^{N} p_t e_t^i + p_1 \sum_{i=1}^{N} b_1^i$$

Substituting the government BC (9.20) into the aggregate BC above, we have

$$\begin{aligned} \sum_{t=0}^{\infty} \sum_{i=1}^{N} p_t c_t^i + \sum_{t=0}^{\infty} p_t G_t + p_1 B_1 &\leq \sum_{t=0}^{\infty} \sum_{i=1}^{N} p_t e_t^i + p_1 B_1 \\ &\Rightarrow \\ \sum_{t=0}^{\infty} \sum_{i=1}^{N} p_t c_t^i + \sum_{t=0}^{\infty} p_t G_t &\leq \sum_{t=0}^{\infty} \sum_{i=1}^{N} p_t e_t^i \end{aligned}$$

As long as (9.22) holds, the difference between $\{(\hat{\tau}_t^i)_{i \in I}\}_{t=0}^{\infty}$ and $\{(\tau_t^i)_{i \in I}\}_{t=0}^{\infty}$ does not change this aggregate BC. Notice that all individuals are indentical, therefore we have $c_t^i = c_t, e_t^i = e_t$, the aggreagte BC becomes

$$\sum_{t=0}^{\infty} p_t c_t + \sum_{t=0}^{\infty} p_t \frac{G_t}{N} \leq \sum_{t=0}^{\infty} p_t e_t$$

This disaggregated BC does not change either for different tax sequence. Thus the allocation is unchanged. ■

A shortcoming of the Arrow-Debreu equilibrium definition and the preceding theorem is that it does not make explicit the substitution between current taxes and government deficits that may occur for two equivalent tax systems $\{(\hat{\tau}_t^i)_{i\in I}\}_{t=0}^{\infty}$ and $\{(\tau_t^i)_{i\in I}\}_{t=0}^{\infty}$. Therefore we will now reformulate this economy sequentially. This will also allow us to see that one of the main assumptions of the theorem, the absence of borrowing constraints is crucial for the validity of the theorem.

As usual with sequential markets we now assume that markets for the consumption good and one-period loans open every period. Let r_{t+1} denote the interest rate on one period loans from period t to period $t+1$. Given the tax system and initial bond holdings each agent i now faces a sequence of budget constraints of the form

$$c_t^i + \frac{b_{t+1}^i}{1+r_{t+1}} + \tau_t^i \leq e_t^i + b_t^i$$

with b_1^i given. In order to rule out Ponzi schemes we have to impose a no Ponzi scheme condition of the form (borrowing constraint)

$$b_{t+1}^i \geq -\alpha_t^i(r, G, \tau)$$

which shows that in general borrowing limit may depend on the sequence of interest rates as well as the endowment stream of the individual and the tax system. We will be more specific about the exact from of the constraint later. In fact, we will see that the exact specification of the borrowing constraint is crucial for the validity of Ricardian equivalence.

The government faces a similar sequence of budget constraints of the form (non-arbitrage condition guarantee that the government bond has the same interest rate as private bond)

$$G_t + B_t \leq \sum_{i\in I} \tau_t^i + \frac{B_{t+1}}{1+r_{t+1}}$$

with B_1 given. B_{t+1} is the bond that government issue at time t, while government has to pay back the bond issued one year ago B_t. We also impose a condition on the government that rules out government policies that run a Ponzi scheme

$$B_t \geq -A_t(r, G, \tau)$$

Now we can define a SME in this economy

Definition 67 *Given a sequence of government spending $\{G_t\}_{t=0}^{\infty}$ and initial government debt B_1 and initial individual bond holding $(b_1^i)_{i\in I}$, an ADE is an allocation $\{(c_t^i, b_{t+1}^i)_{i\in I}\}_{t=0}^{\infty}$, an interest rate $\{r_t\}_{t=0}^{\infty}$, and a sequence of government policies $\{(\tau_t^i)_{i\in I}, B_{t+1}\}_{t=0}^{\infty}$ such that* *(i). Given the interest rate and the sequence of policies,* $\{(c_t^i, b_{t+1}^i)_{i\in I}\}_{t=0}^\infty$ *solves HH i's problem*

$$\max_{\{(c_t^i)_{i\in I}\}_{t=0}^\infty} \sum_{t=0}^\infty \beta^t u(c_t^i)$$

$$s.t.$$

$$c_t^i + \frac{b_{t+1}^i}{1+r_{t+1}} + \tau_t^i \leq e_t^i + b_t^i, \forall t \tag{9.23}$$

$$b_{t+1}^i \geq -\alpha_t^i(r, G, \tau) \tag{9.24}$$

$$c_t^i \geq 0,$$

(ii). Given the price, government policies satisfy $\forall t \geq 0$

$$G_t + B_t \leq \sum_{i\in I} \tau_t^i + \frac{B_{t+1}}{1+r_{t+1}} \tag{9.25}$$

$$B_t \geq -A_t(r, G, \tau) \tag{9.26}$$

(iii). Market clears $\forall t \geq 0$.

$$\sum_{i\in I} c_t^i + G_t = \sum_{i\in I} e_t^i$$

$$\sum_{i\in I} b_{t+1}^i = B_{t+1}.$$

We will particularly look at two forms of borrowing constraints. The first is the so called *natural borrowing or debt limit* following Aiyagari (1994): it is that amount that, at given sequence of interest rates, the consumer can maximally repay, by setting consumption to zero in each period since now. It is given by

$$b_{t+1}^i \geq -\sum_{s=0}^\infty \frac{e_{t+s}^i - \tau_{t+s}^i}{\prod_{j=t+1}^{t+s}(1+r_{j+1})}$$

where $\prod_{j=t+1}^{t}(1+r_{j+1}) = 1$. Similarly we set the borrowing limit of the government at its natural limit:

$$B_t \geq -\sum_{s=0}^\infty \frac{\sum_{i\in I} \tau_{t+s}^i}{\prod_{j=t+1}^{t+s}(1+r_{j+1})}.$$

Another form of borrowing constraint is a commonly used one

$$b_{t+1}^i \geq 0$$

i.e., the HHs cannot borrow at all (but they can lend out). Actually notice that since there is positive supply of government bonds, such restriction does not rule out saving of individuals in equilibrium in the form of holding government bonds.

We can now state a Ricardo Equivalence theorem under the natural borrowing constraint. Let's first prove the following proposition that shows the equivalence between ADE and SME under government policies and natural debt limits.

Proposition 68 *Given a sequence of government spending* $\{G_t\}_{t=0}^{\infty}$ *and initial government debt* B_1 *and initial individual bond holding* $(b_1^i)_{i\in I}$*, Suppose an allocation* $\{(c_t^i)_{i\in I}\}_{t=0}^{\infty}$*, an interest rate* $\{p_t\}_{t=0}^{\infty}$*, and a sequence of government policies* $\{(\tau_t^i)_{i\in I}\}_{t=0}^{\infty}$ *form an ADE. Then there exists a corresponding SME with the natural debt limits* $\left\{\{(\tilde{c}_t^i, \tilde{b}_{t+1}^i)_{i\in I}\}_{t=0}^{\infty}, \{\tilde{r}_t\}_{t=0}^{\infty}, \{(\tilde{\tau}_t^i)_{i\in I}, \tilde{B}_{t+1}\}_{t=0}^{\infty}\right\}$ *such that*

$$
\begin{aligned}
c_t^i &= \tilde{c}_t^i \\
\tau_t^i &= \tilde{\tau}_t^i, \forall i, \forall t.
\end{aligned}
$$

Reversely, let $\left\{\{(\tilde{c}_t^i, \tilde{b}_{t+1}^i)_{i\in I}\}_{t=0}^{\infty}, \{\tilde{r}_t\}_{t=0}^{\infty}, \{(\tilde{\tau}_t^i)_{i\in I}, \tilde{B}_{t+1}\}_{t=0}^{\infty}\right\}$ *form a SME with natural debe limits. If we have*

$$
\begin{aligned}
\tilde{r}_{t+1} &> -1, \forall t \\
\sum_{t=0}^{\infty} \frac{e_t^i - \tau_t^i}{\prod_{j=1}^{t}(1+r_{j+1})} &< \infty, \forall i, \forall t \\
\sum_{s=0}^{\infty} \frac{\sum_{i\in I} \tau_{t+s}^i}{\prod_{j=t+1}^{t+s}(1+r_{j+1})} &< \infty, \forall t
\end{aligned}
$$

Then there exists a corresponding ADE with an allocation $\{(c_t^i)_{i\in I}\}_{t=0}^{\infty}$*, an interest rate* $\{p_t\}_{t=0}^{\infty}$*, and a sequence of government policies* $\{(\tau_t^i)_{i\in I}\}_{t=0}^{\infty}$ *such that*

$$
\begin{aligned}
c_t^i &= \tilde{c}_t^i \\
\tau_t^i &= \tilde{\tau}_t^i, \forall i, \forall t.
\end{aligned}
$$

Proof. The key to the proof is to show the equivalence of the budget sets for the Arrow-Debreu and the sequential markets structure. Normailize $p_1 = 1$ and link the interest rate $\{\tilde{r}_t\}_{t=1}^{\infty}$ and AD price $\{p_t\}_{t=0}^{\infty}$ as following

$$1 + \tilde{r}_{t+1} = \frac{p_t}{p_{t+1}} \tag{9.27}$$

From BC in SME equation (9.23), recursively substituting out bond holding b^i_{t+1}, then using equation above to substituting interest rate, we end up with

$$\sum_{t=0}^{\infty} p_t(c^i_t + \tau^i_t - e^i_t) + \lim_{T\to\infty} p_{T+1} b^i_{T+1} = b^i_1$$

This will be equivalent to the BC in ADE (9.19) if

$$\lim_{T\to\infty} p_{T+1} b^i_{T+1} \geq 0$$

But from the natural debt constraint

$$\begin{aligned} p_{T+1} b^i_{T+1} &\geq -p_{T+1} \sum_{s=0}^{\infty} \frac{e^i_{T+s} - \tau^i_{T+s}}{\prod_{j=T+1}^{T+s}(1+r_{j+1})} \\ &= -\sum_{s=T+1}^{\infty} p_s(e^i_s - \tau^i_s) \\ &= -\sum_{s=0}^{\infty} p_s(e^i_s - \tau^i_s) + \sum_{s=0}^{T} p_s(e^i_s - \tau^i_s) \end{aligned}$$

Taking limit on both sides when $T \to \infty$, we have

$$\begin{aligned} \lim_{T\to\infty} p_{T+1} b^i_{T+1} &\geq -\sum_{s=0}^{\infty} p_s(e^i_s - \tau^i_s) + \lim_{T\to\infty} \sum_{s=0}^{T} p_s(e^i_s - \tau^i_s) \\ &= 0 \end{aligned}$$

Notice that the condition of this proposition guarantee $\sum_{s=0}^{\infty} p_s(e^i_s - \tau^i_s) < \infty$, therefore the limit is well-defined.

So at equilibrium prices, with natural debt limits and the restrictions posed in the proposition a consumption allocation satisfies the Arrow-Debreu budget constraint (at equilibrium prices) if and only if it satisfies the sequence of budget constraints in sequential markets. A similar argument can be carried out for the budget constraint(s) of the government. Since the utility function is same in both equilibrium, the allocation has to be same since they are identical problems.

Notice that given an ADE consumption allocation, the corresponding bond holdings for the sequential markets formulation are

$$b^i_{t+1} = \sum_{s=0}^{\infty} \frac{c^i_{t+s} + e^i_{t+s} - \tau^i_{t+s}}{\prod_{j=t}^{t+s}(1+r_{j+1})}.$$

■

Now we can prove the Ricardian Equivalence under the natural debt limit.

Proposition 69 *Suppose that the natural debt limit prevails. Take as given a sequence of government spending* $\{G_t\}_{t=0}^{\infty}$ *and initial government debt* B_1 *and initial endowment of government bond* $(b_1^i)_{i\in I}$*, Suppose that allocation* $\{(c_t^i, b_{t+1}^i)_{i\in I}\}_{t=0}^{\infty}$*, interest rate* $\{r_t\}_{t=0}^{\infty}$*, and a sequence of government policies* $\{(\tau_t^i)_{i\in I}, B_{t+1}\}_{t=0}^{\infty}$ *consist an SME. Let* $\{(\hat{\tau}_t^i)_{i\in I}\}_{t=0}^{\infty}$ *be an arbitrary alternative tax system satisfying*

$$\sum_{t=0}^{\infty}\sum_{i=1}^{N} p_t \tau_t^i = \sum_{t=0}^{\infty}\sum_{i=1}^{N} p_t \hat{\tau}_t^i.$$

Then $\{\{(c_t^i, \hat{b}_{t+1}^i)_{i\in I}\}_{t=0}^{\infty}, \{r_t\}_{t=0}^{\infty}, \{(\hat{\tau}_t^i)_{i\in I}, \hat{B}_{t+1}\}_{t=0}^{\infty}\}$ *also consist an SME where*

$$\hat{b}_{t+1}^i = \sum_{s=0}^{\infty} \frac{c_{t+s}^i + e_{t+s}^i - \hat{\tau}_{t+s}^i}{\prod_{j=t+1}^{t+s}(1+r_{j+1})} \tag{9.28}$$

$$B_t = \sum_{s=0}^{\infty} \frac{\left(\sum_{i\in I}\tau_{t+s}^i - G_{t+s}\right)}{\prod_{j=t+1}^{t+s}(1+r_{j+1})}. \tag{9.29}$$

Proof. We only sketch the proof. According to the proposition above, under natural debt limit, HHs in SME face a single intertemporal BC as same as in ADE. This alternative tax scheme does not change this Arrow-Debreu BC. The optimal consumption plan only depends on this AD budget constraint, hence it is also unchanged. Next, the adjusted borrowing plan $\{\hat{b}_{t+1}^i)_{i\in I}\}_{t=0}^{\infty}$ is obtained by solving BC in SME (9.23)forwardly given new tax scheme. Note that the adjusted borrowing plan satisfies trivially the (adjusted) natural debt limit in every period, since the consumption plan $\{(c_t^i)_{i\in I}\}_{t=0}^{\infty}$ is a nonnegative sequence.

The second point of the proposition is that the altered government tax and borrowing plans continue to satisfy the government's budget constraint. In particular, we see that the government's budget set at time 0 does not depend on the timing of taxes, only their present value. Thus, under the altered tax plan with an unchanged present value of taxes, the government can finance the same expenditure plan $\{G_t\}_{t=0}^{\infty}$. The adjusted borrowing plan for government $\{B_{t+1}\}_{t=0}^{\infty}$ is computed in a similar way as above to arrive at (9.29). ■

It is not surprised that we still obtain Ricardian Equivalence under natural debt limit since natural debt limit is equivalent to the non-negativity constraint of consumption. The Ricardian equivalence theorem rests on several important assumptions. The first is that there are *perfect capital markets*. If consumers face binding borrowing constraints (e.g. non-borrowing constraint $b_{t+1}^i \geq 0$), or if, with uncertainty, not a full set of contingent claims is available, then Ricardian equivalence may fail. Secondly one has to require that all taxes are *lump-sum*. Non-lump sum taxes may distort relative prices (e.g. labor income taxes distort the relative price of leisure) and hence a change in the timing of taxes may have real effects. Here in this chapter, all the taxes are imposed on the endowment, hence are lump-sum. Finally a change from one to another tax system is assumed to *not redistribute wealth among agents*. This was a maintained assumption of the theorem, which required that the total tax bill that each agent faces was left unchanged by a change in the tax system. In a world with finitely lived overlapping generations this would mean that a change in the tax system is not supposed to redistribute the tax burden among different generations, which is automatically satisfied in two-period OLG setting.

Now let's move to a stronger restriction on debt limit, which we call no-borrowing constraint

$$b_{t+1}^i \geq 0, \forall i, \forall t.$$

Following example shows that with no-borrowing constraint, Ricardian equivalence may fail.

Example 70 *Consider an infinite horizon economy with 2 agents,* $u(c) = \ln c$, $\beta = 0.5$, $b_1^i = B_1 = 0$, $G_t = 0$, $\forall t$, *and the endowment follows*

$$e_t^1 \begin{cases} 2 & \text{if } t \text{ odd} \\ 1 & \text{if } t \text{ even} \end{cases}$$

$$e_t^2 \begin{cases} 1 & \text{if } t \text{ odd} \\ 2 & \text{if } t \text{ even} \end{cases}$$

First let's consider a tax scheme

$$\tau_t^1 \begin{cases} 0.5 & \text{if } t \text{ odd} \\ -0.5 & \text{if } t \text{ even} \end{cases}$$

$$\tau_t^2 \begin{cases} -0.5 & \text{if } t \text{ odd} \\ 0.5 & \text{if } t \text{ even} \end{cases}$$

Thus HH i's problem is

$$\begin{aligned} & \max \sum_{t=1}^{\infty} \beta^t \ln c_t^i \\ & s.t. \\ c_t^i + \frac{b_{t+1}^i}{1+r_{t+1}} + \tau_t^i \leq & \; e_t^i + b_t^i, \forall t \\ c_t^i, b_{t+1}^i \geq & \; 0 \end{aligned}$$

Note that this tax scheme is self-balanced. $\sum_{i=1}^2 \tau_t^i = 0$. *Obviously with no-borrowing constraint, the unique equilibrium is autarky with* $c_t^i = 1.5, \forall i, \forall t$. *FOCs of HH i's problem are (*$\lambda_t \geq 0$ *is the multiplier associated with BC at time t,* $\mu_{t+1} \geq 0$ *is the multiplier associated with no-borrowing constraint* $b_{t+1}^i \geq 0$*)*

$$\begin{aligned} c_t^i &: \beta^t \frac{1}{c_t^i} = \lambda_t \\ b_{t+1}^i &: \frac{\lambda_t}{1+r_{t+1}} = \lambda_{t+1} + \mu_{t+1} \end{aligned}$$

Combining these FOCs, we have

$$\frac{c_{t+1}^i}{\beta c_t^i} = \frac{\lambda_t}{\lambda_{t+1}} = 1 + r_{t+1} + \frac{(1+r_{t+1})\mu_{t+1}}{\lambda_{t+1}}$$

Hence we have

$$\begin{aligned} \frac{c_{t+1}^i}{\beta c_t^i} &\geq 1 + r_{t+1} \\ &\quad \text{and} \\ \frac{c_{t+1}^i}{\beta c_t^i} &= 1 + r_{t+1}, \text{ if } b_{t+1}^i > 0 \end{aligned}$$

We know in equilibrium, $\frac{c_{t+1}^i}{\beta c_t^i} = 2$, *which implies the equilibrium interest rate* $r_{t+1} \leq 1$ *and* $r_{t+1} = 1$ *if* $b_{t+1}^i > 0$. *For concreteness let's take* $r_{t+1} = 1$ *for all t.18 Then the present value of total bill of taxes for the first consumer is*

$$\begin{aligned} \sum_{t=1}^{\infty} \frac{\tau_t^1}{(1+r_{t+1})^{t-1}} &= \frac{1}{2} - \frac{1}{4} + \frac{1}{8} - \frac{1}{16} + \frac{1}{32} - .... \\ &= \frac{1}{2}(1 + \frac{1}{4} + (\frac{1}{4})^2 + ...) - \frac{1}{4}(1 + \frac{1}{4} + (\frac{1}{4})^2 + ...) \\ &= \frac{1}{3}. \end{aligned}$$

The present value of total tax bill for consumer 2 is $-\frac{1}{3}$. *Now lets consider an alternative tax system that has* $\tau_1^1 = \frac{1}{3}, \tau_1^2 = -\frac{1}{3}$ *and* $\tau_t^i = 0, \forall i, \forall t \geq 2$. *Obviously the new tax system has the same present value of total tax bill as in the original one. Hence it satisfies the government budget constraint and does not redistribute among agents. However, equilibrium allocations change to* $c_1^1 = \frac{5}{3}$, $c_1^2 = \frac{4}{3}$ *and* $c_t^i = e_t^i, \forall i, \forall t \geq 2$. *Furthermore, equilibrium interest rate change to* $r_2 = \frac{3}{2.5}$ *and* $r_t = 0, \forall t \geq 3$. *Ricardian equivalence fails in this economy.*

Under the noborrowing constraint, we require that the asset holding at time t b_{t+1}^i satisfies budget constraint and not fall below zero simultaneously. That is, under the noborrowing constraint, we have to check more than just a single intertemporal budget constraint for the HH at time 0. Changes in the timing of taxes that obey equation (9.22) evidently alter RHS of sequential BC (9.23) and can, for example, cause a previously binding borrowing constraint no longer to be binding, and vice versa. The allocation has to be changed accordingly. But if we can have an additional condition to guarantee that under the new tax system, the bond holding still satisfies no-borrowing constraint, we can have Ricardian equivalence again. See the following proposition.

Proposition 71 *Suppose that no-borrowing constraint prevails. Take as given a sequence of government spending* $\{G_t\}_{t=0}^{\infty}$ *and initial government debt* B_1 *and* *initial endowment of government bond* $(b_1^i)_{i\in I}$*, Suppose that allocation* $\{(c_t^i, b_{t+1}^i)_{i\in I}\}_{t=0}^{\infty}$*, interest rate* $\{r_t\}_{t=0}^{\infty}$*, and a sequence of government policies* $\{(\tau_t^i)_{i\in I}, B_{t+1}\}_{t=0}^{\infty}$ *consist an SME. Let* $\{(\hat{\tau}_t^i)_{i\in I}\}_{t=0}^{\infty}$ *be an arbitrary alternative tax system satisfying*

$$\begin{aligned} \sum_{t=0}^{\infty}\sum_{i=1}^{N} p_t \tau_t^i &= \sum_{t=0}^{\infty}\sum_{i=1}^{N} p_t \hat{\tau}_t^i \\ &\text{and} \\ \hat{b}_{t+1}^i &= \sum_{s=0}^{\infty} \frac{c_{t+s}^i + e_{t+s}^i - \hat{\tau}_{t+s}^i}{\prod_{j=t+1}^{t+s}(1+r_{j+1})} \geq 0 \end{aligned} \tag{9.30}$$

Then $\{\{(c_t^i, \hat{b}_{t+1}^i)_{i\in I}\}_{t=0}^{\infty}, \{r_t\}_{t=0}^{\infty}, \{(\hat{\tau}_t^i)_{i\in I}, \hat{B}_{t+1}\}_{t=0}^{\infty}\}$ *also consist an SME.*

We leave the proof of this proposition as an exercise.

Finite Horizon and Operative Bequest Motives

How about Ricardian equivalence in OLG setting? It should be clear from the above discussion that one only obtains a very limited Ricardian equivalence theorem for OLG economies. Any change in the timing of taxes that redistributes among generations is in general not neutral in the Ricardian sense. If we insist on representative agents within one generation and purely selfish, two-period lived individuals, then in fact any change in the timing of taxes can't be neutral unless it is targeted towards a particular generation, i.e. the tax change is such that it decreases taxes for the currently young only and increases them for the old next period. Hence, with sufficient generality we can say that Ricardian equivalence does not hold for OLG economies with purely selfish individuals.

Rather than to demonstrate this obvious point with another example we now briefly review Barro's (1974) argument that under certain conditions finitely lived agents will behave as if they had infinite lifetime. As a consequence, Ricardian equivalence is re-established. Barro's (1974) classic article "Are Government Bonds Net Wealth?" asks exactly the Ricardian question, namely does an increase in government debt, financed by future taxes to pay the interest on the debt increase the net wealth of the private sector? Barro identified two main sources for why future taxes are not exactly offsetting current tax cuts (increasing government deficits): a) finite lives of agents that lead to intergenerational redistribution caused by a change in the timing of taxes b) imperfect private capital markets. Barro's paper focuses on the first source of nonneutrality.

Barro's key result is the following: in OLG models finiteness of lives does not invalidate Ricardian equivalence as long as current generations are connected to future generations by a chain of operational *intergenerational, altruistically motivated transfers*. These may be transfers from old to young via bequests or from young to old via social security programs. Let us look at his formal model.

Setup of Barro's Model

1. It is the standard pure exchange OLG model with two-period lived agents.
2. No population growth.
3. Cohort t has endowment process $(e_t^t = w, e_{t+1}^t = 0)$.
4. There is a government that, for simplicity, has 0 government expenditures but initial outstanding government debt B. This debt is denominated in terms of the period 1 consumption good.
5. The initial old generation is endowed with the B units of bonds.
6. Government bonds are zero coupon bonds with maturity of one period which means one unit of government bond pays off one unit of consumption good tomorrow.
7. Government keeps its outstanding government debt constant and we assume a constant one-period real interest rate r on these bonds.
8. In order to finance the interest payments on government debt the government taxes the currently young people.

The government's BC at time t is

$$\underbrace{B_{t-1}}_{\text{expenditure}} = \underbrace{\tau + \frac{B_t}{1+r}}_{\text{revenue: tax + bond issued}}$$

Notice that $\frac{1}{1+r}$ is the bond price. Since $B_t = B, \forall t$, we have

$$B = \tau + \frac{B}{1+r}$$

which implies

$$\tau = \frac{rB}{1+r}$$

And we assume $\frac{rB}{1+r} \leq w$.

Now the cohort t has preferences

$$U_t(c_t^t, c_{t+1}^t, a_{t+1}^t) = u(c_t^t) + \beta u(c_{t+1}^t) + \alpha V_{t+1}(e_{t+1})$$

subject to two BCs

$$\begin{aligned} c_t^t + \frac{a_t^t}{1+r} &= w - \tau \\ c_{t+1}^t + \frac{a_{t+1}^t}{1+r} &= a_t^t + a_t^{t-1} \\ a_{t+1}^t &\geq 0 \quad \text{(bequest motive)} \end{aligned}$$

where a_t^t denotes the savings of cohort t at time t (when they are young), a_{t+1}^t denotes the savings of cohort t at time $t+1$ (when they are old), we require $a_{t+1}^t \geq 0$ to let it stand for bequest motive. In our previous OLG models obviously $a_{t+1}^t = 0$ was the only optimal choice since individuals were completely selfish, they do not care about their children. Notice that when old the individuals have two sources of funds: their own savings from the previous period and the bequests from the previous generation.

We can consolidate these two-period BCs into a lifetime BC

$$c_t^t + \frac{c_{t+1}^t}{1+r} + \frac{a_{t+1}^t}{(1+r)^2} = w + \frac{a_t^{t-1}}{1+r} - \tau \equiv e_t$$

where e_t is the present value of toal lifetime income for the cohort t. Therefore, $e_{t+1} = w + \frac{a_{t+1}}{1+r} - \tau$ and it is surely a function of cohort t's bequest a_{t+1}^t. $V_{t+1}(e_{t+1})$ is the maximal utility cohort $t+1$ can attain with lifetime resources e_{t+1}. To be consistent with bequest motive, we assume $1 > \alpha > 0$.

The initial old's problem is

$$\begin{aligned} V_0(B) &= \max_{c_1^0, a_1^0 \geq 0} \{\beta u(c_1^0) + \alpha V_1(e_1)\} \\ & \quad s.t. \\ c_1^0 + \frac{a_1^0}{1+r} &= B \\ e_1 &= w + \frac{a_1^0}{1+r} - \tau \end{aligned}$$

Notice that these two constraints can be combined into one as follows

$$c_1^0 + e_1 = w + B - \tau. \tag{9.31}$$

The solution to this problem is a decision rule $c_1^0(w, B, \tau)$ and $a_1^0(w, B, \tau)$. Now assume the bequest motive is operative, i.e., $a_1^0(w, B, \tau) > 0$. Let's answer Barro's question by considering following experiment: increase initial government debt marginally by ΔB (i.e., the initial old BC changes to $c_1^0 + \frac{a_1^0}{1+r} = B + \Delta B$) and repay this additional debt by levying higher taxes on the first young generation–cohort 1.

Clearly, in the OLG model without bequest motives such a change in fiscal policy is not neutral: it increases resources available to the initial old and reduces resources available to the first regular generation. This will change consumption of both generations and interest rate. But now in Barro's economy with altruistic preference, In order to repay the ΔB, from the government budget constraint taxes for the young have to increase by

$$\Delta\tau = \Delta B$$

since by assumption government debt from the second period onwards remains unchanged ($B + \Delta B = \tau + \Delta\tau + \frac{B}{1+r}$ for time $t = 1$, $B = \tau + \frac{B}{1+r}, \forall t \geq 2$). Then from the BC of initial old (9.31), we clearly know that the optimal allocation for

c_1^0 and e_1 are unchanged since the RHS of the BC is unchanged. What happen is the initial old generation receives additional transfers of bonds of magnitude ΔB from the government, but since they care about their descendants, and they know that the government will immediately levy higher taxes $\tau + \Delta\tau$ on their children in order to repay higher government debt, the initial old cohort thus increases its bequests a_1^0 by $(1+r)\Delta B$ so that lifetime resources available to their children (and hence their allocation) is left unchanged. In other words, altruistically motivated bequest motives just undo the change in fiscal policy. Ricardian equivalence is restored.

$$\begin{aligned} e_1 &= w + \frac{a_1^0 + (1+r)\Delta B}{1+r} - (\tau + \Delta\tau) \\ &= w + \frac{a_1^0}{1+r} - \tau + (\Delta B - \Delta\tau) \\ &= w + \frac{a_1^0}{1+r} - \tau \end{aligned}$$

Ricardian equivalence also holds in general with operational altruistic bequests. In doing so we will establish the linkage between Barro's OLG economy and an economy with infinitely lived consumers and borrowing constraints. Again consider the problem of the

initial old generation, we can write it in a FE way, the value function for the initial old is $V_0(B)$

$$\begin{aligned} V_0(B) &= \max_{\substack{c_1^0, a_1^0 \geq 0 \\ c_1^0 + \frac{a_1^0}{1+r} = B}} \{\beta u(c_1^0) + \alpha V_1(e_1)\} \\ &= \max_{\substack{c_1^0, a_1^0 \geq 0 \\ c_1^0 + \frac{a_1^0}{1+r} = B}} \left\{ \beta u(c_1^0) + \alpha \max_{\substack{c_1^1, c_2^1, a_2^1 \geq 0, a_1^1 \\ c_1^1 + \frac{a_1^1}{1+r} = w - \tau \\ c_2^1 + \frac{a_2^1}{1+r} = a_1^1 + a_1^0}} \{u(c_1^1) + \beta u(c_2^1) + \alpha V_2(e_2)\} \right\} \end{aligned}$$

We can rewrite it as

$$\begin{aligned} V_0(B) &= \max_{c_1^0, a_1^0, c_1^1, c_2^1, a_2^1 \geq 0, a_1^1} \{\beta u(c_1^0) + \alpha u(c_1^1) + \alpha\beta u(c_2^1) + \alpha^2 V_2(a_2^1) \\ & \quad s.t. \\ c_1^0 + \frac{a_1^0}{1+r} &= B \\ c_1^1 + \frac{a_1^1}{1+r} &= w - \tau \\ c_2^1 + \frac{a_2^1}{1+r} &= a_1^1 + a_1^0 \end{aligned}$$

Repeating this procedure infinitely many times, and let $\alpha = \beta$, we will end up with

$$\begin{aligned} V_0(B) &= \max_{\{c_t^{t-1}, c_t^t, a_t^{t-1}\}_{t=1}^{\infty}} \{\beta u(c_1^0) + \sum_{t=1}^{\infty} \beta^t (u(c_t^t) + \beta u(c_{t+1}^t))\} \\ & \quad s.t. \\ c_1^0 + \frac{a_1^0}{1+r} &= B \\ c_t^t + \frac{c_{t+1}^t}{1+r} + \frac{a_{t+1}^t}{(1+r)^2} &= w + \frac{a_t^{t-1}}{1+r} - \tau \end{aligned}$$

Essentially, the problem is equivalent to that of an infinitely lived consumer that faces a no-borrowing constraint. This infinitely lived consumer is peculiar in the sense that her periods are subdivided into two subperiods, she eats twice a period, c_t^t in the first subperiod and c_{t+1}^t in the second subperiod, and saves in the second subperiod $a_{t+1}^t \geq 0$. The relative price of the consumption goods in the two subperiods is given by $(1+r)$. Apart from these reinterpretations this is a standard infinitely lived consumer with no-borrowing constraints imposed on her. Consequently one obtains a Ricardian equivalence proposition similar to Proposition 71, where the requirement of "operative bequest motives" $a_{t+1}^t \geq 0$ is the equivalent to condition (9.30). More generally, this argument shows that an OLG economy with two period-lived agents and operative bequest motives is formally equivalent to an infinitely lived agent model.

9.3 Overlapping Generations Models with Production

We want more realistic OLG model with production due to two reasons:

1. We are curious that if the theoretical results we obtained under pure endowment economy still hold in production economy?

2. With more realistic setting with production, we can do policy analysis. Since Auerbach and Kotlikoff (1987), macroeconoimists are able to set up large-scale OLG model with production to analyze a lot of economic policy issues related to social security, taxation, and so on.

9.3.1 Basic Setup of the Model

Demographic Structure

The economy consists of identical agents who live up to two periods: young and old. We use N_t^t to denote number of young people in period t. N_t^{t-1} denotes number of old people in period t. We assume no death until the end of the second period, hence we have $N_t^t = N_{t+1}^t$. We normalize the size of the initial young generation to 1, i.e., $N_0^0 = N_1^0 = 1$. We assume population grows at constant rate n, thus we have

$$N_t^t = (1+n)N_{t-1}^{t-1}.$$

which implies

$$N_t^t = (1+n)^t N_0^0 = (1+n)^t.$$

Therefore, the total population in period t is

$$N_t^{t-1} + N_t^t = (1+n)^t + (1+n)^{t-1}.$$

Preferences

For each cohort t (generation was born in period t), the preference is given by

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

Again, we make standard assumptions about utility function u: u is strictly increasing, strictly concave, twice continuously differentiable and satisfies the Inada conditions. All individuals are assumed to be purely selfish and have no bequest motives whatsoever. The initial old generation has preferences $u(c_1^0)$. Each individual of generation $t \geq 1$ has as endowments one unit of time for working when young and no endowment when old. Each member of the initial old generation is endowed with initial capital stock $(1+n)\bar{k}_1 > 0$.

Production

There is a representative firm that accesses to a CRS technology with production function

$$Y_t = F(K_t, L_t).$$

Given this assumption, profits are zero in equilibrium and we do not have to specify ownership of firms. Since the production function is CRS, we can rewrite it in an intensive form. Define $k_t = \frac{K_t}{L_t}$, we have

$$y_t = \frac{Y_t}{L_t} = F(\frac{K_t}{L_t}, 1) = f(k_t)$$

We also assume that f is twice continuously differentiable, strictly concave and satisfies the Inada conditions.

Timing

The timing of events for a given generation t is as follows:

1. At the beginning of period t production takes place with labor of generation t and capital saved by the current old generation $t-1$ from the previous period. The young generation earns a wage w_t.

2. At the end of period t the young generation decides how much of the wage income to consume, c_t^t, and how much to save for tomorrow, s_t^t. The saving occurs in form of physical capital, which is the only asset in this economy.

3. At the beginning of period $t+1$ production takes place with labor of generation $t+1$ and the saved capital of the now old generation t. The return on savings equals $r_{t+1}+1-\delta$, where again $1+r_{t+1}$ is the rental rate of capital and δ is the rate of depreciation, so that $r_{t+1}+1-\delta$ is the real gross interest rate from period t to $t+1$.

4. At the end of period $t+1$ generation t consumes its savings plus interest rate $(1+r_{t+1}-\delta)s_t^t$ and then dies.

We can now define a SME for this economy.

Definition 72 *Given $\bar{k}_1$, a SME is an allocation for the consumers $\{(c_t^t, c_{t+1}^t, s_t^t)\}_{t=1}^{\infty} \cup c_1^0$, an allocation for the firm $\{K_t, L_t\}_{t=1}^{\infty}$, and prices $\{r_t, w_t\}_{t=1}^{\infty}$ such that:*
(i). $\forall t \geq 1$, given prices (w_t, r_{t+1}), $(c_t^t, c_{t+1}^t, s_t^t)$ solves cohort t's problem

$$\begin{aligned} &\max_{c_t^t, c_{t+1}^t, s_t^t} u(c_t^t) + \beta u(c_{t+1}^t) \\ &s.t. \\ c_t^t + s_t^t &\leq w_t \cdot 1 \\ c_{t+1}^t &\leq (1+r_{t+1}-\delta)s_t^t \\ c_t^t, c_{t+1}^t &\geq 0. \end{aligned}$$

(ii). Given $\bar{k}_1$ and r_1, c_1^0 solves the initial old generation's problem

$$\begin{aligned} &\max_{c_1^0} u(c_1^0) \\ &s.t. \\ c_1^0 &\leq (1+r_1-\delta)\bar{k}_1 \\ c_1^0 &\geq 0. \end{aligned}$$

(iii). $\forall t \geq 1$, given prices (w_t, r_t), (K_t, L_t) solves the firm's problem

$$\max_{K_t, L_t} F(K_t, L_t) - r_t K_t - w_t L_t$$

(iv). Markets clear.
Goods market:

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

Labor market:

$$L_t = N_t^t$$

Capital market:

$$K_{t+1} = N_t^t s_t^t.$$

In many cases, we are particularly interested in a stationary equilibrium.

Definition 73 *A **steady state** (or stationary equilibrium) is an equilibrium with allocations* $\{(c_t^t, c_{t+1}^t, s_t^t)\}_{t=1}^{\infty} \cup c_1^0$, $\{K_t, L_t\}_{t=1}^{\infty}$, *and prices* $\{r_t, w_t\}_{t=1}^{\infty}$ *such that*

$$\begin{aligned} c_t^t &= \bar{c}_1 \\ c_{t+1}^t &= \bar{c}_2 \\ s_t^t &= \bar{s} \\ K_t &= \bar{k} N_t^t \\ L_t &= N_t^t \\ r_t &= \bar{r} \\ w_t &= \bar{w} \end{aligned}$$

for initial capital stock $\bar{k}_1 = \bar{k}$.

In other words, a steady state is an equilibrium for which the allocation (per capita) is constant over time, given that the initial condition for the initial capital stock is exactly right.

9.3.2 Characterization of the Equilibrium

We can derive out an equation to say savings = investment (it is a direct result from goods market clearing condition)

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

Dividing both sides by N_t^t and using labor market clearing condition, we obtain

$$\begin{aligned} \frac{K_{t+1}}{N_{t+1}^{t+1}} \frac{N_{t+1}^{t+1}}{N_t^t} - (1-\delta)\frac{K_t}{N_t^t} &= F(\frac{K_t}{N_t^t}, 1) - (c_t^t + N_t^{t-1} \frac{N_t^{t-1}}{N_t^t} c_t^{t-1}) \\ &\Rightarrow \\ k_{t+1}(1+n) - (1-\delta)k_t &= f(k_t) - c_t^t - \frac{c_t^{t-1}}{1+n} \\ &\Rightarrow \\ f(k_t) &= c_t^t + \frac{c_t^{t-1}}{1+n} - (1-\delta)k_t + k_{t+1}(1+n) \quad (9.32) \end{aligned}$$

The Firm's FOCs

$$\begin{aligned} r_t &= \frac{\partial Y_t}{\partial K_t} = \frac{\partial L_t f(k_t)}{\partial K_t} = L_t f'(k_t) \frac{1}{L_t} = f'(k_t) \\ w_t &= \frac{\partial Y_t}{\partial L_t} = \frac{\partial L_t f(k_t)}{\partial L_t} = f(k_t) - f'(k_t) k_t. \end{aligned}$$

The EE for cohort t's problem

$$\begin{aligned} u'(c_t^t) &= \beta(1 + r_{t+1} - \delta) u'(c_{t+1}^t) \\ &\quad or \\ u'(w_t - s_t^t) &= \beta(1 + r_{t+1} - \delta) u'((1 + r_{t+1} - \delta) s_t^t) \end{aligned}$$

Notice that this EE implicitly define a saving function

$$\begin{aligned} s_t^t &= s(w_t, r_{t+1}) \\ &= s(f(k_t) - f'(k_t)k_t, f'(k_{t+1})) \end{aligned}$$

Therefore, optimal savings are a function of this and next period's capital stock. From the capital clearing condition, we also know

$$s_t^t = \frac{K_{t+1}}{N_t^t} = \frac{N_{t+1}^{t+1}}{N_t^t}\frac{K_{t+1}}{N_{t+1}^{t+1}} = (1+n)k_{t+1}$$

Therefore we have following first-order difference equation

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

Since we know the initial capital-labor ratio $k_1 = \frac{K_1}{L_1} = \frac{K_1}{N_1^1} = \frac{(1+n)\bar{k}_1}{1+n} = \bar{k}_1$, So in principle we could put equation above on a computer and solve for the entire sequence of $\{k_t\}_{t=1}^{\infty}$ and hence for the entire equilibrium. Under some appropriate conditions, we can show that there exists a steady state for this first-order difference condition which satisfies

$$\bar{k} = \frac{s(f(\bar{k}) - f'(\bar{k})\bar{k}, f'(\bar{k}))}{1+n}$$

9.3.3 An Analytical Example

Consider an OLG economy with production. $u(c) = \ln c$, $Y = AK^{\alpha}L^{1-\alpha}$ and no population growth $n = 0$.

The firm's problem is

$$\max_{K_t, L_t} AK_t^{\alpha}L_t^{1-\alpha} - r_tK_t - w_tL_t$$

FOCs are

$$\begin{aligned} w_t &= (1-\alpha)A(\frac{K_t}{L_t})^{\alpha} = A(1-\alpha)k_t^{\alpha} \\ r_t &= \alpha A(\frac{K_t}{L_t})^{\alpha-1} = A\alpha k_t^{\alpha-1}. \end{aligned}$$

Cohort t's problem is

$$\begin{aligned} & \max_{c_t^t, c_{t+1}^t, s_t^t} \ln(c_t^t) + \beta\ln(c_{t+1}^t) \\ & s.t. \\ c_t^t + s_t^t &\leq w_t \\ c_{t+1}^t &\leq (1 + r_{t+1} - \delta)s_t^t \\ c_t^t, c_{t+1}^t &\geq 0 \end{aligned}$$

The lifetime BC thus is

$$c_t^t + \frac{c_{t+1}^t}{(1 + r_{t+1} - \delta)} = w_t$$

EE for cohort t is

$$\frac{c_{t+1}^t}{c_t^t} = \beta(1 + r_{t+1} - \delta)$$

Combining lifetime BC and EE, we have

$$\begin{aligned} c_t^t &= \frac{1}{1+\beta} w_t \\ s_t^t &= w_t - c_t^t = \frac{\beta}{1+\beta} w_t \end{aligned}$$

Since $N_t^t = 1, \forall t$, we have

$$k_{t+1} = s_t^t = \frac{\beta}{1+\beta} w_t = \frac{\beta}{1+\beta} A(1-\alpha) k_t^\alpha$$

This first-order difference equation has two steady states where $k_{t+1} = k_t = \bar{k}$, one is $\bar{k} = 0$. Another one (more meaningful one) is

$$\bar{k} = [\frac{\beta}{1+\beta} A(1-\alpha)]^{1-\alpha} > 0$$

9.3.4 Dynamic Efficiency in the Model

We know in steady state, we have

$$\begin{aligned} f(\bar{k}) &= \bar{c}_1 + \frac{\bar{c}_2}{1+n} - (1-\delta)\bar{k} + \bar{k}(1+n) \\ &\Rightarrow \\ f(\bar{k}) - (n+\delta)\bar{k} &= \bar{c}_1 + \frac{\bar{c}_2}{1+n} \equiv \bar{c} \end{aligned}$$

Here $\bar{c}$ is total (per capita) consumption in the steady state. Using implicit function theorem, we obtain

$$\frac{d\bar{c}}{d\bar{k}} = f'(\bar{k}) - (n+\delta).$$

When we let $\frac{d\bar{c}}{d\bar{k}} = 0$, we have

$$f'(\bar{k}) = n + \delta \tag{9.33}$$

This relation is called the *golden rule of capital accumulation* and characterizes the efficient steady state capital stock which will maximize the steady state consumption. To understand the golden rule is indeed efficient, think about if

$$f'(\bar{k}) < n + \delta, \tag{9.34}$$

The capital stock is *inefficiently high* so that the marginal product of capital is inefficiently low; it is so high that its marginal productivity $f'(\bar{k})$ is outweighed by the cost of replacing depreciated capital, $\delta\bar{k}$ and provide newborns with the steady state level of capital per worker, $n\bar{k}$. Thus a marginal decrease of the capital stock actually leads to higher available overall consumption. We can show in this case the steady state equilibrium is not PO by constructing an alternative Pareto superior allocation.

Suppose the economy is in the steady state at some arbitrary date t and suppose that the steady state satisfies (9.34). Now consider the following alternative allocation: at time t reduce the capital stock per worker to be saved to the next period, k_{t+1}, by a marginal amount $\varepsilon < 0$ to $\tilde{k} = \bar{k} - \varepsilon$ and keep it at $\tilde{k}$ forever since time t. (This alternative allocation since it is a free disposal to the original allocation.) Go back to the resource constraint equation (9.32), let's define $c_t = c_t^t + \frac{c_t^{t-1}}{1+n}$ is the total (per worker) consumption at time t, we have

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

Therefore, staring from time t, the difference between this alternative allocation and the original allocation in terms of consumption is

$$\begin{aligned}\Delta c_t &= -(1+n)\varepsilon > 0 \\ \Delta c_{t+s} &= f'(\bar{k})\varepsilon - (n+\delta)\varepsilon > 0, \forall s \geq 1\end{aligned}$$

In this way we can increase total per capita consumption in every period. Now we can just divide the additional consumption between the two generations alive in a given period in such a way that make both generations better off, which is straightforward to do, given that we have extra consumption goods to distribute in every period. Thus a Pareto superior allocation is found. The original equilibrium is *dynamically inefficient.*

In our analytical example above, now with positive population growth rate $n > 0$, it is straightforward to compute the unique steady state as

$$\bar{k} = [\frac{\beta A(1-\alpha)}{(1+\beta)(1+n)}]^{1-\alpha} > 0$$

so that

$$\bar{r} = f'(\bar{k}) = \alpha A\bar{k}^{\alpha-1} = \frac{\alpha(1+\beta)(1+n)}{\beta(1-\alpha)}.$$

Therefore, the economy is dynamically inefficient iff

$$\frac{\alpha(1+\beta)(1+n)}{\beta(1-\alpha)} < n + \delta.$$

Notice that in the OLG model with peoduction, the real interest rate is $r_t - \delta$, therefore, the dynamic inefficiency condition is equivalent to

$$\text{real interest rate in ss} = r_t - \delta < n = \text{population growth rate}$$

That is exactly the Samuelson case with $n = 0$ in pure exchange economy. Accordingly, we also have the following Balasko-Shell optimality criteria in the production economy.

Theorem 74 *(Cass-Balasko-Shell) A feasible allocation is Pareto optimal if and only if*

$$\sum_{t=1}^{\infty}\prod_{s=1}^{t}\frac{(1+r_{s+1}-\delta)}{1+n_{s+1}}=+\infty.$$

In steady state equilibrium, obviously this condition is reduced to $\bar{r} = f'(\bar{k}) \geq n+\delta$. Thus, if $f'(\bar{k}) < n+\delta$, we obtain dynamic inefficiency.

If the competitive equilibrium of the economy features dynamic inefficiency its citizens save more than is socially optimal. Hence government programs that reduce national saving are called for. The possible candidates for the government policy that corrects the dynamic inefficiency are: PAYG social security system, taxation on capital and government debt.

Chapter 10

Bewley-Aiyagari Incomplete Market Model

In this chapter[1] we will look at a class of models that take a first step at explaining the distribution of wealth in actual economies. So far our models abstracted from distributional aspects. As standard in macro up until the early 90's our models had representative agents, that all faced the same preferences, endowments and choices, and hence received the same allocations. Obviously, in such environments one cannot talk meaningfully about the income distribution, the wealth distribution or the consumption distribution. One exception was the OLG model, where, at a given point of time we had agents that differed by age, and hence differed in their consumption and savings decisions. However, with only two (groups of) agents the cross sectional distribution of consumption and wealth looks rather sparse, containing only two points at any time period. A more realistic way to introduce heterogeneity into macroeconomics is needed.

We want to accomplish two things in this chapter. First, we want to summarize the main empirical facts about the current U.S. income and wealth distribution. Second we want to build a class of models which are both tractable and whose equilibria feature a nontrivial distribution of wealth across agents. The basic idea is the following. The is a continuum of agents that are *ex ante* identical and all have a *stochastic* endowment process that follow a Markov chain. Then endowments are realized in each period, and it so happens that some agents are lucky and get good endowment realizations, others are unlucky and get bad endowment realizations. The aggregate endowment is constant across time (no aggregate shock). If there was a complete set of Arrow-Debreu contingent claims, then people would simply fully insure each other against the endowment shocks and we would be back at the standard representative agent model. Here we will assume that people *cannot insure against these shocks* (for reasons exogenous from the model), in that we close down all insurance markets. The only financial instrument that agents, by assumption, can use to

[1] This chapter is heavily based on Kreuger (2002) Lecture Notes Chapter 10.

hedge against endowment uncertainty are one period bonds (or IOU's) that yield a riskless return r. In other words, agents can only *self-insure* by borrowing and lending at a risk free rate r. In addition, we impose tight limits on how much people can borrow (otherwise, it turns out, self-insurance (almost) as good as insuring with Arrow-Debreu claims). As a result, agents will accumulate wealth, in the form of bonds, to hedge against endowment uncertainty. Those agents with a sequence of good endowment shocks will have a lot of wealth, those with a sequence of bad shocks will have low wealth (or even debt). Hence the model will use as input an *exogenously* specified stochastic endowment (income) process, and will deliver as output an *endogenously* derived wealth distribution.

To analyze these models we will need to keep track of the characteristics of each agent at a given point of time, which, in most cases, is at least the current endowment realization and the current wealth position. Since these differ across agents, we need an entire distribution (measure) to keep track of the state of the economy. Hence the richness of the model with respect to distributional aspects comes at a cost: we need to deal with entire distributions as state variables, instead of just numbers as the capital stock. Therefore some knowlege about measure theory is needed.

10.1 Stylized Facts about the Income and Wealth Distribution

In this section[2] we describe the main stylized facts characterizing the U.S. income and wealth distribution. For data on the income and wealth distribution we have to look beyond the national income and product accounts (NIPA) data, since NIPA only contains *aggregated* data. What we need are data on income and wealth of a sample of individual families.

10.1.1 Data Sources

For the US income and wealth distribution data, there are three main data sets:

1. *the Survey of Consumer Finances (SCF).* The SCF is conducted in three year intervals; the four available surveys are for the years 1989, 1992, 1995 and 1998. It is conducted by the National Opinion Research center at the University of Chicago and sponsored by the Federal Reserve system. It contains rich information about U.S. households' income and wealth. In each survey about 4,000 households are asked detailed questions about their labor earnings, income and wealth. One part of the sample is representative of the U.S. population, to give an accurate description of the entire population. The second part oversamples rich households, to get a more precise idea about the precise composition of this groups' income and wealth composition. As we will see, this group accounts for the majority of total household wealth, and hence it is particularly important to

[2] This section is heavily based on Diaz-Gimenez, Quadrini, and Rios-Rull (1997).

have good information about this group. The main advantage of the SCF is the level of detail of information about income and wealth. The main disadvantage is that it is not a panel data set, i.e., households are not followed over time. Hence dynamics of income and wealth accumulation cannot be documented on the household level with this data set.

2. *the Panel Study of Income Dynamics (PSID).* It is conducted by the Survey Research Center of the University of Michigan and mainly sponsored by the National Science Foundation. The PSID is a panel data set that started with a national sample of 5,000 U.S. households in 1968. The same sample individuals are followed over the years, barring attrition due to death or nonresponse. New households are added to the sample on a consistent basis, making the total sample size of the PSID about 8700 households. The income and wealth data are not as detailed as for the SCF, but its panel dimension allows to construct measures of income and wealth dynamics, since the same households are interviewed year after year. Also the PSID contains data on consumption expenditures, albeit only food consumption. In addition, in 1990, a representative national sample of 2,000 Latinos, differentially sampled to provide adequate numbers of Puerto Rican, Mexican-American, and Cuban-Americans, was added to the PSID database. This provides a host of information for studies on discrimination and other relevant labor economics issues.

3. *the Consumer Expenditure Survey (CEX).* The CEX is conducted by the U.S. Bureau of the Census and sponsored by the Bureau of Labor statistics. The first year the survey was carried out was 1980. The CEX is a so-called rotating panel: each household in the sample is interviewed for four consecutive quarters and then rotated out of the survey. Hence in each quarter 20% of all households is rotated out of the sample and replaced by new households. In each quarter about 3000 to 5000 households are in the sample, and the sample is representative of the U.S. population. The main advantage of the CEX is that it contains very detailed information about consumption expenditures. Information about income and wealth is inferior to the SCF and PSID, also the panel dimension is significantly shorter than for the PSID (one household is only followed for 4 quarters). Given our focus on income and wealth we will not use the CEX here, but anyone writing a paper about consumption will find the CEX an extremely useful data set.

10.1.2 Measurements

Before we go to stylized facts, let's study how to measure the inequality and mobility in the economy.

We first define our interested variables: earnings, income, and wealth.

Definition 75 ***Earnings****: Wages, Salaries of all kinds, plus a fraction 0.864 of business income (such as income from professional practices, business and* *farm sources).*

Definition 76 ***Income****: All kinds of household revenues before taxes, including: wages and salaries, a fraction of business income (as above), interest income, dividends, gains or losses from the sale of stocks, bonds, and real estate, rent, trust income and royalties from any other investment or business, unemployment and worker compensation, child support and alimony, aid to dependent children, aid to families with dependent children, food stamps and other forms of welfare and assistance, income form social security and other pensions, annuities, compensation for disabilities and retirement programs, income from all other sources including settlements, prizes, scholarships and grants, inheritances, gifts and so forth.*

Definition 77 ***Wealth****: Net worth of households, defined as the value of all real and financial assets of all kinds net of all the kinds of debts. Assets considered are: residences and other real estate, farms and other businesses, checking accounts, certificates of deposit, and other bank accounts, IRA/Keogh accounts, money market accounts, mutual funds, bonds and stocks, cash and call money at the stock brokerage, all annuities, trusts and managed investment accounts, vehicles, the cash value of term life insurance policies, money owed by friends, relatives and businesses, pension plans accumulated in accounts.*

In summary, earnings is one of the components of income, the one related to *labor* input. Income is the revenue from *all* sources before taxes but after transfers. So this variable includes both labor earnings and income generated by wealth. Finally, wealth is defined as the net worth of the HHs.

Measurement of Inequality

In this section we use data from the SCF. We measure the dispersion of the earnings, income and wealth distribution in a cross section of households by several measures. Let the sample of size n, assumed to be representative of the population, be given by $\{x_1, x_2, ..., x_n\}$, where x is the variable of interest (i.e., earnings, income or wealth). We define the mean

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

The standard deviation is defined by

$$std(x) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}.$$

We thus define a very important measurement of inequality, *coefficient of variation (cv)*, as follows

$$cv(x) = \frac{std(x)}{\bar{x}}.$$

Variable	**Mean**	**cv**	**Gini**	$\frac{\textbf{Top 1\%}}{\textbf{Bottom 40\%}}$	**Location of mean**	$\frac{\textbf{mean}}{\textbf{median}}$
Earnings	\$33074	4.19	0.63	211	65%	1.65
Income	\$45924	3.86	0.57	84	71%	1.72
Wealth	\$184308	6.09	0.78	875	80%	3.61

Table 10.1: Inequality of Earnings, Income, and Wealth in US: 1992

Note this is desirable than Standard Deviation, because it is independent of the unit. Imagine we compare the inequality of income between Italy and US. Even if the real inequality is the same, the Standard Deviation of income across Italian people measures in Lira is much larger than the Standard Deviation of income across US people measured in US dollar, because of the difference of the denomination. By dividing by mean, we can avoid such trouble.

A second commonly used measure is the *Gini coefficient* and the associated *Lorenz curve*. Here is how to construct Lorenz curve of income: First sort the agents by their income. And arrange the agents on the horizontal axis $[0, 1]$. The agent with lowest income comes on the point 0 and the one with highest income comes on the point 1. The vertical axis $[0, 1]$ represents the proportion of total income owned by agents. Specifically, a point on the income Lorenz curve (x, y) represents that fraction of agents in the interval $[0, x]$ owns fraction y of total income of the economy. By definition, the Lorenz curve crosses the point $(0, 0)$ because no agent owns nothing. Similarly, the Lorenz curve crosses the point $(1, 1)$ because all agents in the economy own all of the income in the economy. Note that the Lorenz curve never crosses the 45 degree line by construction, unless everybody has the same amount (in this case Lorenz curve is exactly the 45 degree line). If you are considering income, and there is nobody with negative income in the economy, the Lorenz curve does not have a negative slope in $[0, 1]$. Suppose you are considering the Lorenz curve of the net asset. In this case the poorest agents have usually negative net assets, and the asset Lorenz has a negative slope as long as the corresponding agents have negative net asset. Accordingly, the Gini coefficient is defined by

$$\text{Gini} = \frac{\text{Area enclosed by Lorenz curve and 45 degree line}}{\text{Triangle made by connecting the points (0,0), (1,0), (1,1)}}$$

Gini coefficient falls between zero and 1, with higher Gini coefficients indicating bigger concentration of earnings, income or wealth. As extremes, for complete equality of income the Gini coefficient is zero and for complete concentration (the richest person has all earnings, income, wealth) the Gini coefficient is 1.

Figure 10.1 and Table 10.1 and 10.2 summarize the main stylized facts with respect to the concentration of earnings, income and wealth from the 1992 SCF.

From them, we observe the following stylized facts:

1. There is substantial variability in earnings, income and wealth across U.S. households. The standard deviation of earnings, for example, is about \$140000. The top 1% of earners on average earn 21100% more than the bottom 40%, the corresponding number for income is still 8400%.

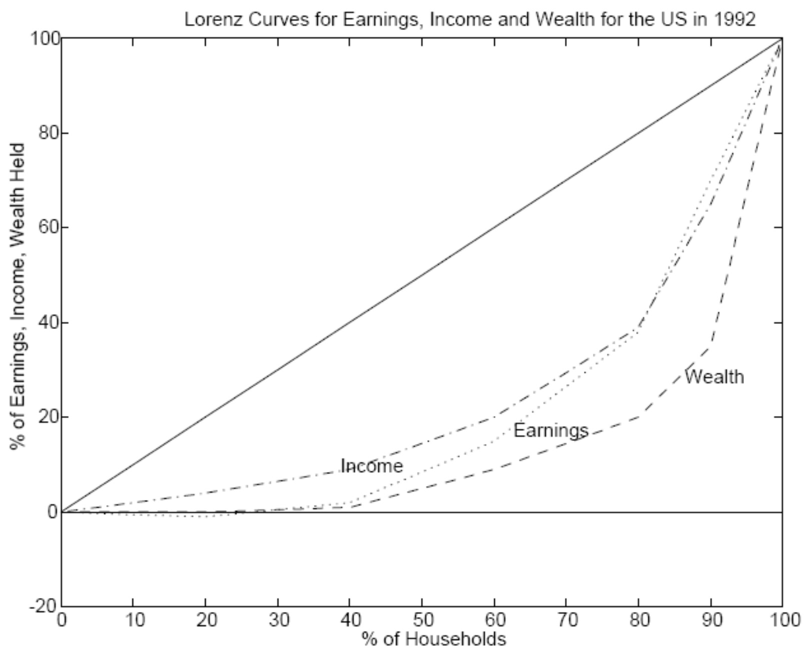


Figure 10.1: Lorenz Curve

	Quintiles				
Measure	1st	2nd	3rd	4th	5th
Earnings	-0.40	3.19	12.49	23.33	61.39
Income	2.18	6.63	11.80	19.47	59.91
Wealth	-0.39	1.74	5.72	13.43	79.49

Table 10.2: Share of Earnings, Income, and Wealth in different quintiles: 1992 SCF

2. Wealth is by far the most concentrated of the three variables, followed by earnings and income. That income is most equally distributed across households makes sense as income includes payments from government insurance programs. The distribution would be even less dispersed if we would look at income after taxes, due to the progressivity of the tax code.

3. Since wealth is accumulated past income minus consumption, it also makes intuitive sense that it is most most concentrated. For example, the top 1% households of the wealth distribution hold about 29.55% of total wealth.

4. The distribution of all three variables is skewed. If the distributions were symmetric, the median would equal the mean and the mean would be located exactly at the 50-percentile of the distribution. For all three variables the mean is substantially higher than the median, which indicates skewness to the right. In accordance with the last stylized fact, the distribution of wealth is most skewed, followed by the distribution of earnings and the distribution of income.

Measurement of Mobility

Not only is there a lot of variability in earnings, income and wealth across households, but also a lot of dynamics within the corresponding distribution. Statistics that are useful in analyzing mobility are the followings:

- Autocorrelation $\rho(x_{t-s}, x_t)$ and $\rho(x_{t+s}, x_t), s = 1, 2, ..., T$.

- Transition matrix. Since SCF is not a panel data, i.e., it does not track people over time. Therefore, we use PSID here. We use data on net worth from the PSID for the years 1984 and 1989 (reported in the 1984 and 1989 PSIDs) and combine them with data on earnings and income for the same HHs for these two years (reported in 1985 and 1990 PSIDs). We then use these data to construct the transition matrices for the 1984 earnings, income, and wealth quintiles as reported in Table 10.3.

For example, the entry in the first row and the first column of Table 10.3 shows that 85.8% of the HHs in the lowest earnings quintile in 1984 stayed in the lowest earnings quintile in 1989. To avoid the role of retirees in shaping the mobility of households with zero earnings, Table 10.4 reports the transition matrices in earnings for HHs with positive earnings in both positive sample periods. To partially control for the role played by age in shaping the properties of the mobility of earnings, Table 10.4 also reports the transition matrices of earnings for those HHs with heads between the so-called prime ages 35-45 in 1984.

Based on these two tables, We find the following stylized facts:

1. There is substantial persistence of labor earnings, in particular at the *lowest* and *highest* quintile. For the lowest quintile this may be due to retirees and long-term unemployed. Stratifying the sample as in Table

Measure	**1984 Quintile**	**1989 Quintile** 1st	2nd	3rd	4th	5th
Earnings	1st	85.8	11.6	1.4	0.6	0.5
	2nd	18.6	40.9	30.0	7.1	3.4
	3rd	7.1	12.0	47.0	26.2	7.6
	4th	7.5	6.8	17.5	46.5	21.7
	5th	5.8	4.1	5.5	18.3	66.3
Income	1st	71.0	17.9	7.0	2.9	1.3
	2nd	19.5	43.8	22.9	10.1	3.7
	3rd	5.1	25.5	37.2	24.9	7.3
	4th	2.5	10.7	23.4	42.5	20.8
	5th	1.9	2.1	9.5	20.3	66.3
Wealth	1st	66.7	23.4	6.6	2.9	0.4
	2nd	25.4	46.6	20.4	5.4	2.3
	3rd	5.8	24.4	44.9	20.5	4.6
	4th	1.8	4.6	22.4	49.6	21.6
	5th	0.7	0.8	5.7	21.6	71.2

Table 10.3: Three Measures of the Economic Mobility of US Households

Type of Household	**1984 Quintile**	**1989 Quintile** 1st	2nd	3rd	4th	5th
With Posotive Earnings in both 1984 and 1989	1st	58.8	25.1	9.0	5.1	2.0
	2nd	20.2	45.6	21.6	8.6	4.0
	3rd	9.7	20.2	40.4	21.9	7.8
	4th	7.7	6.1	20.0	45.9	20.4
	5th	3.6	2.9	9.0	18.4	66.1
With Heads 35-45 years old in 1984	1st	63.3	27.2	4.0	3.3	2.3
	2nd	23.6	44.3	22.3	7.3	2.4
	3rd	4.7	16.7	47.0	25.1	6.6
	4th	6.9	8.1	20.2	44.6	20.1
	5th	1.1	4.0	6.4	19.1	69.3

Table 10.4: A Closer Look at the Economic Mobility of US Households

10.4 indicates that this may be part of the explanation that 85.8% of all the households that were in the lowest earnings quintile in 1984 are still in the lowest earnings quintile in 1989. But even looking at Table 10.4 there seems to be substantial persistence of earnings at the low and high end, with persistence being even more accentuated for prime-age HHs.

2. The persistence properties of income are similar to those of earnings, which is understandable given the high correlation between income and earnings (ρ(income, earnings) = 0.928).

3. Wealth seems to be more persistent than income and earnings.

The inequality facts we just summarized suggest that the following elements are important ingredients for a reliable theory of inequality:

- *Transfers.* Income transfers distort the labor/leisure decision, and they allow households to survive without work. They are an important source of income for earnings- and wealth-poor households; hence, they should play an important role in any attempt to account for the *lower tails* of the distributions.

- *Businesses.* Businesses in financial distress account for the sizable amount of negative income earned by many U.S. households. Moreover, business income is an important source of income for the households in the *upper tails* of the distributions. These facts suggest that both business successes and business failures should be important elements for any theory of inequality.

- *Retirees.* Retirees hold a large share of total wealth. Moreover, their labor earnings are zero. These facts spell trouble for any theory of inequality that abstracts from elements of the life cycle.

- *Education.* Households whose head has a college education have more than twice the earnings, income, and wealth of those households whose head has a high school education. Understanding the determinants of the acquisition of education becomes a crucial part of understanding inequality.

- *Marital Status.* The better financial performance of married households over single households cannot be accounted for only by family size. A successful theory should account for how the patterns of household formation and dissolution shape inequality.

Now it's time to build a model that tries to explain the U.S. wealth distribution, taking as given the earnings distribution, i.e., treating the earnings distribution as an input in the model. Let's start from the classic savings problem (or income fluctuation problem).

10.2 The Classic Income Fluctuation Problem

This section describes a version of what is sometimes called a savings problem. A consumer wants to maximize the expected discounted sum of a concave function of one-period consumption rates. However, the consumer is cut off from all insurance markets and almost all asset markets. The consumer can only purchase nonnegative amount of a single risk-free asset. The absence of insurance opportunities induces the consumer to adjust his asset holdings to acquire "*self-insurance*."

The savings problem is

$$\begin{aligned} & \max_{\{c_t, a_{t+1}\}} E_0 \sum_{t=0}^{\infty} \beta^t u(c_t) \\ & s.t. \\ c_t + a_{t+1} & \leq y_t + (1+r)a_t \\ a_{t+1} & \geq 0, c_t \geq 0 \\ & a_0 \text{ given} \end{aligned}$$

We assume $u(c)$ is a strictly increasing, strictly concave, twice continuously differentiable function of the consumption of a single good c. Here the endowment T time t y_t takes a finite number of values indexed by $s \in S$. So this is an endowment economy (without production). In particular, the set of possible endowments is $\bar{y}_1 < \bar{y}_2 < ... < \bar{y}_S$. Elements of the sequence of endowments are independently and identically distributed with $\Pr(y = y_s) = \Pi_s, \Pi_s \geq 0$, and $\sum_{s\in S} \Pi_s = 1$. There are no insurance markets.

The agent can hold nonnegative amounts of a single risk-free asset that has a net rate of return r. Let $a_t \geq 0$ be the agent's assets at the beginning of period t. We assume that $a_0 = y_0$ is drawn from the time-invariant endowment distribution.

The Bellman equation for this problem is

$$v(a) = \max_c \{u(c) + \sum_{s=1}^{S} \beta \Pi_s v[(1+r)a + y_s - c]\}$$

where y_s is the income realization in state $s \in S$. The value function $v(a)$ inherits the basic properties of $u(c)$, that is, v is increasing, strictly concave, and differentiable.

"Self-insurance" occurs when the agent uses savings to insure himself against income fluctuations. On the one hand, in response to low income realizations, an agent can draw down his savings and avoid temporary large drops in consumption. On the other hand, the agent can partly save high income realizations in anticipation of poor outcomes in the future. We are interested in the longrun properties of an optimal "self-insurance" scheme. Will the agent's future consumption settle down around some level $\bar{c}$? Or will the agent eventually become impoverished? We will show that neither of these outcomes occurs: consumption will diverge to infinity!

Before analyzing it under uncertainty, we'll briefly consider the savings problem under a certain endowment sequence. With a nonrandom endowment that does not grow perpetually, consumption does converge.

10.2.1 Deterministic Income

We break our analysis of the nonstochastic case into two parts, depending on the stringency of the borrowing constraint. We begin with the least stringent possible borrowing constraint, namely, the natural borrowing constraint on one-period Arrow securities, which are risk free in the current context. After that, we'll arbitrarily tighten the borrowing constraint to arrive at the no-borrowing condition $a_{t+1} \geq 0$ imposed in the statement of the problem above.

Natural Debt Limit

Let's specify the borrowing constraint as

$$a_{t+1} \geq -b_t$$

and the borrowing limit b is defined by

$$b_t = \sup \sum_{j=1}^{\infty} \frac{y_{t+j}}{(1+r)^j} < +\infty$$

It is the maximal amount that it is feasible to repay at time t when $c_t \geq 0, \forall t$. The key of specifying the borrowing constraint in this form is that the borrowing constraint will never be binding, i.e., we will never have $a_{t+1} = -b_t$. (Think about why? Hint: Inada condition.) Hence the optimal consumption allocation is completely characterized by the Euler equations

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

and A-D BC is

$$\sum_{t=0}^{\infty} \frac{c_t}{(1+r)^{t+1}} = \sum_{t=0}^{\infty} \frac{y_t}{(1+r)^{t+1}} + a_0.$$

Depending on the value of $\beta(1+r)$, we have following three cases:

1. If $\beta(1+r) > 1$, $u'(c_t) > u'(c_{t+1}), \Rightarrow c_t < c_{t+1}, \forall t$. Consumption is increasing over time. Since borrowing is limited at zero, assets must grow to finance consumption, so also assets diverge. The individual is "too patient" (β is too high) or/and the rate of return on savings is too high (r is too high). Both forces push her to accumulate too much wealth.

2. If $\beta(1+r) < 1$, $u'(c_t) < u'(c_{t+1}), \Rightarrow c_t > c_{t+1}, \forall t$. Consumption is decreasing over time and converges to a finite number.

3. If $\beta(1+r) = 1$, $u'(c_t) = u'(c_{t+1})$, $\Rightarrow c_t = c_{t+1}, \forall t$. Consumption is constant over time. One can prove that the effect of the borrowing constraint lasts until a given time and vanishes thereafter. Therefore, consumption and assets converge to a finite value.

Notice that β measures the patience. Higher β means more patient. $\frac{1}{1+r}$ is the relative price of consumption tomorrow compared to consumption today. If the relative price is lower than the natural discount ($\frac{1}{1+r} < \beta$), people want to consume more tomorrow (by saving more), therefore, we have $c_{t+1} > c_t$. On the other hand, in terms of saving, the desire to save is increasing in patience β and the interest rate r . When $1 + r$ is too large, HHs' desire to accumulate is "too strong" hence assets also diverge to infinity. Therefore, in deterministic case, all we need to prevent divergence is $\beta(1 + r) \leq 1$.

No-Borrowing Constraint

Let us make the same assumptions as before, but now assume that the borrowing constraint is tighter than the natural borrowing limit. For simplicity assume that the consumer is prevented from borrowing completely, i.e., assume $a_{t+1} \geq 0$. We also assume that the endowment sequence $\{y_t\}_{t=0}^{\infty}$ is constant at $y_t = y$. Now in the optimization problem we have to take the borrowing constraints into account explicitly. Forming the Lagrangian and denoting by λ_t the Lagrange multiplier for the BC at time t and by μ_t the Lagrange multiplier for the non-negativity constraint for a_{t+1}, we have the following FOCs:

$$\begin{aligned} c_t &: \beta^t u'(c_t) = \lambda_t \\ c_{t+1} &: \beta^{t+1} u'(c_{t+1}) = \lambda_{t+1} \\ a_{t+1} &: -\lambda_t + \mu_t + (1+r)\lambda_{t+1} = 0 \end{aligned}$$

The complementary slackness condition is

$$a_{t+1}\mu_t = 0, a_{t+1} \geq 0, \mu_t \geq 0$$

Or

$$a_{t+1} > 0 \Rightarrow \mu_t = 0$$

Combining FOCs, we have

$$\begin{aligned} u'(c_t) &\geq \beta(1+r)u'(c_{t+1}) \\ &= \beta(1+r)u'(c_{t+1}), \text{ if } a_{t+1} > 0 \end{aligned}$$

Notice that for this economy, a_t and y only enter as sum in the dynamic programming problem. Hence we can define a variable $x_t = (1 + r)a_t + y$, which we call "cash at hand", i.e., the total resources of the agent available for consumption or saving at time t. Now we can rewrite the FE as

$$\begin{aligned} v(x_t) &= \max_{c_t, a_{t+1} \geq 0} \{u(c_t) + \beta v(x_{t+1})\} \\ & \quad s.t. \\ c_t + a_{t+1} &= x_t \\ x_{t+1} &= (1+r)a_{t+1} + y \end{aligned}$$

or more concisely

$$v(x_t) = \max_{x_t \geq a_{t+1} \geq 0} \{u(x_t - a_{t+1}) + \beta v((1+r)a_{t+1} + y)\}$$

EE for this FE is

$$\begin{aligned} u'(c_t(x_t)) &\geq \beta(1+r)v'((1+r)a_{t+1} + y) \\ &= \beta(1+r)v'((1+r)a_{t+1} + y) \text{ if } a_{t+1}(x_t) > 0 \end{aligned}$$

The envelope condition is

$$v'(x_t) = u'(c_t(x_t))$$

Therefore

$$v'(x_{t+1}) = u'(c_{t+1}(x_{t+1}))$$

Substituting into EE, we have

$$\begin{aligned} u'(c_t(x_t)) &\geq \beta(1+r)u'(c_{t+1}(x_{t+1})) \\ &= \beta(1+r)u'(c_{t+1}(x_{t+1})), \text{ if } a_{t+1}(x_t) > 0. \end{aligned}$$

We end up with the same EE as derived from SP. Our first result about characterizing the equilibrium path is following:

Proposition 78 *In this economy, first, consumption is strictly increasing in cash at hand, i.e., $c_t'(x_t) > 0, \forall t$. Second, there exists a $\bar{x}_t$ such that $a_{t+1}(x_t) = 0$ for all $x_t \leq \bar{x}_t$ and $a_{t+1}'(x_t) > 0$ for all $x_t > \bar{x}_t$. Finally, we have $c_t'(x_t) < 1$ and $a_{t+1}'(x_t) < 1$.*

Proof. For the first part, take first-order derivative w.r.t. x_t on the envelope condition, we obtain

$$v''(x_t) = u''(c_t(x_t))c_t'(x_t)$$

$\Rightarrow$

$$c_t'(x_t) = \frac{v''(x_t)}{u''(c_t(x_t))} > 0$$

since both utility and value function are strictly concave.

Suppose the borrowing constraint is not binding, i.e., $a_{t+1}(x_t) > 0$, then from differentiating the EE w.r.t. x_t we get

$$u''(c_t(x_t))c_t'(x_t) = \beta(1+r)^2 v''((1+r)a_{t+1} + y)a_{t+1}'(x_t)$$

$\Rightarrow$

$$a'_{t+1}(x_t) = \frac{u''(c_t(x_t))c'_t(x_t)}{\beta(1+r)^2 v''((1+r)a_{t+1}+y)} > 0$$

Now we want to show that $a_{t+1}(x_t) = 0$ for all $x_t \leq \bar{x}_t$ (this implies $a_{t+1}(\bar{x}_t) = 0$). Suppose not, for example, suppose that $a_{t+1}(x_t) > a_{t+1}(\bar{x}_t) = 0$, then by EE, we know

$$\begin{aligned} u'(c_t(x_t)) &= \beta(1+r)v'((1+r)a_{t+1}(x_t)+y) \\ u'(c_t(\bar{x}_t)) &\geq \beta(1+r)v'((1+r)a_{t+1}(\bar{x}_t)+y) \end{aligned}$$

But since v' is strictly decreasing (v is strictly concave), we have

$$\beta(1+r)v'((1+r)a_{t+1}(x_t)+y) < \beta(1+r)v'((1+r)a_{t+1}(\bar{x}_t)+y)$$

$\Rightarrow$

$$u'(c_t(\bar{x}_t)) > u'(c_t(x_t))$$

$\Rightarrow$

$$c_t(\bar{x}_t) < c_t(x_t)$$

Since we already know that $c'_t(x_t) > 0$, we have

$$x_t > \bar{x}_t$$

Contradiction! Same reasoning applies to exclude the case that $a_{t+1}(x_t) < a_{t+1}(\bar{x}_t) = 0$.

Finally to show $c'_t(x_t) < 1$ and $a'_{t+1}(x_t) < 1$, we differentiate the BC at time t w.r.t. x_t

$$c'_t(x_t) + a'_{t+1}(x_t) = 1$$

Also notice that $c'_t(x_t), a'_{t+1}(x_t) > 0$. ■

The proposition basically states that the more cash at hand an agent has, coming into the period, the more he consumes and the higher his asset accumulation, provided that the borrowing constraint is not binding. It also states that there is a cut-off level for cash at hand below which the borrowing constraint is always binding. Obviously when the borrowing constraint is binding, i.e., when $x_t \leq \bar{x}_t$, we have $a_{t+1}(x_t) = 0$ which implies $c_t = x_t$, the agent consumes all his income.

Let's especially consider the case when $\beta(1+r) < 1$ hence consumption is declining over time. We have following proposition.

Proposition 79 *When $\beta(1+r) < 1$, if $a'(x) > 0$, then we have $x' < x$. We also have $a'(y) = 0$[3]. And there exists a cut-off level $\bar{x} > y$ such that $a'(x) = 0$ for all $x \leq \bar{x}$.*

[3] Here a little bit notation abuse, $a'(x)$ is not a first-order derivative of a w.r.t. x, in stead, it is the function of next period asset holding.

Proof. When borrowing constraint is not binding, $a'(x) > 0$, the EE is

$$v'(x) = \beta(1+r)v'(x') < v'(x')$$

since v is strictly concave, we have $x' < x$.

For the second part, suppose that $a'(y) > 0$. Then from the EE and the strict concavity of the value function, we have

$$\begin{aligned} v'(y) &= \beta(1+r)v'((1+r)a'(y)+y) \\ &< v'((1+r)a'(y)+y) \\ &< v'(y) \end{aligned}$$

Contradiction! Hence we have $a'(y) = 0$ and $c(y) = y$.

For the last part, we also prove by contradiction. Suppose $a'(x) > 0, \forall x > \bar{x} > y$. Pick arbitrary such x and define the sequence $\{x_t\}_{t=0}^{\infty}$ recursively by

$$\begin{aligned} x_0 &= x \\ x_t &= (1+r)a'(x_{t-1}) + y > y \end{aligned}$$

If there exists a smallest T such that $x_T = y$ then we found a contradiction, since then $a'(x_{T-1}) = 0$ and $x_{T-1} > y$. So suppose that $x_t > y$ for all t, therefore we have $a'(x) > 0, \forall t$. Hence

$$\begin{aligned} v'(x_0) &= \beta(1+r)v'(x_1) \\ &= [\beta(1+r)]^2 v'(x_2) \\ &= ... \\ &= [\beta(1+r)]^t v'(x_t) \\ &< [\beta(1+r)]^t v'(y) \\ &= [\beta(1+r)]^t u'(y) \end{aligned}$$

where the inequality follows from the fact that $x_t > y$ and the strict concavity of v. the last equality follows from the envelope theorem and the fact that $a'(y) = 0$ and $c(y) = y$.

But since v and u both are strictly increasing, $v'(x_0) > 0$, $u'(y) > 0$. By assumption $\beta(1+r) < 1$, therefore, there must exist a finite t such that $v'(x_0) > [\beta(1+r)]^t u(y)$. A contradiction! ■

This last result bounds the optimal asset holdings (and hence cash at hand) from above since the asset holding is decreasing given the borrowing constraint is not binding (or is constant at y, in the case the borrowing constraint binds). Since computational techniques usually rely on the finiteness of the state space, we want to make sure that for our theory the state space can be bounded from above (see section 4.2.1). This theorem thus is a desirable result. The theorem also implies that the agent eventually becomes credit-constrained: there exists a finite time τ such that the agent consumes his endowment in all periods following τ. We summarize it in the following proposition.

Proposition 80 *When $\beta(1+r)<1$, there exists a finite time τ such that the agent consumes his endowment in all periods following τ.*

Proof. We prove it be contradiction. Suppose we have $a'(x)=0$ but $a'(x')>0$, here x' is the next period cash in hand. By definition

$$x' = (1+r)a'(x) + y = y$$

But from the previous proposition, we know $a'(x') = a'(y) = 0$. Contradiction! ■

In summary, for the infinite horizon, if $\beta(1+r)<1$, under deterministic and constant income we have a full qualitative characterization of the allocation: If $a_0 = 0$ then the consumer consumes his income forever from time 0. If $a_0 > 0$, then cash at hand and hence consumption is declining over time, and there exists a time $\tau(a_0)$ such that for all $t > \tau(a_0)$ the consumer consumes his income forever from thereon, and consequently does not save anything.

10.2.2 Stochastic Income and Borrowing Limits

Now let's consider the stochastic income case. We first look at i.i.d. case.

Stochastic Income: i.i.d. case

We first assume that endowments are independently and identically distributed with $\Pr(y = y_s) = \Pi_s, \Pi_s \geq 0$, and $\sum_{s\in S}\Pi_s = 1$. And we maintain the assumption $\beta(1+r)<1$. Thus, the FE is

$$v(x) = \max_{0\leq a'\leq x}\{u(x-a') + \sum_{s=1}^{S}\beta\Pi_s v((1+r)a' + y_s)\}$$

FOC is

$$\begin{aligned} u'(x-a') &\geq \beta(1+r)\sum_{s=1}^{S}\Pi_s v'((1+r)a'(x) + y_s) \\ &= \beta(1+r)\sum_{s=1}^{S}\Pi_s v'((1+r)a'(x) + y_s), \text{ if } a'(x) > 0 \end{aligned}$$

Envelope condition is

$$v'(x) = u'(x - a'(x)) = u'(c)$$

For simplicity, denote

$$Ev'(x') \equiv \sum_{s=1}^{S}\Pi_s v'((1+r)a'(x) + y_s).$$

The proof of the following proposition is identical to the deterministic case.

Proposition 81 *When $\beta(1+r) < 1$, consumption is strictly increasing in cash at hand, i.e. $c'(x) \in (0, 1]$. Optimal asset holdings are either constant at the borrowing limit or strictly increasing in cash at hand, i.e., $a'(x) = 0$ or $\frac{da(x)}{dx} \in (0, 1)$. Moreover, there exists $\bar{x} \geq y_1$ such that for all $x \leq \bar{x}$, we have $c(x) = x$ and $a'(x) = 0$.*

The proposition states that there is a cutoff level for cash at hand below which the consumer consumes all cash at hand and above which he consumes less than cash at hand and saves $a'(x) > 0$ and the saving is also strictly increasing in cash at hand.

What if $\beta(1 + r) = 1$? Previously in the deterministic case, we know that we still get convergence results for consumption and assets. Goes back to the EE

$$\begin{aligned} u'(x - a') &\geq \beta(1 + r) \sum_{s=1}^{S} \Pi_s v((1 + r)a'(x) + y_s) \\ &= \beta(1 + r) \sum_{s=1}^{S} \Pi_s v((1 + r)a'(x) + y_s), \text{ if } a'(x) > 0 \end{aligned}$$

Since now $\beta(1 + r) = 1$, the EE reduces to

$$\begin{aligned} v'(x) &\geq Ev'(x') \\ &= Ev'(x'), \text{ if } a'(x) > 0 \end{aligned}$$

Since cash at hand x now is a random variable, so does $v'(x)$. Therefore, according to probability theory (see Richard Durrett (1996): "*Probability: Theory and Examples*" page 231), $v'(x)$ is a nonnegative *supermartingale*, and it is a *martingale* if the borrowing constraint is not binding. By the martingale convergence theorem (see Durrett page 235-236), a nonnegative supermartingale will converge almost surely to a finite limit.

But we can show that the limiting value of $v'(x)$ must be zero because of the following argument: Suppose to the contrary that $v'(x)$ converges for some realization to a strictly positive limit. That supposition implies that x converges to a finite positive value. But this implication is immediately contradicted by the nature of the optimal policy function, which makes c a function of x (hence a), together with the budget constraint $c + a' - (1 + r)a \leq y$, the LHS of BC is a deterministic finite number, but the RHS is random due to y. Therefore, the limit of $v'(x)$ cannot be positive, it must converge to zero, implying that assets diverge to infinity since value function v is strictly concave. Since $x \to \infty$, so does the consumption. Unfortunately, the convergence result of the deterministic case does not hold in the stochastic case if $\beta(1 + r) = 1$.

We can give a simple proof (and intuition) of this result if we assume that $u''' > 0$, i.e., the marginal utility is strictly convex. From the EE

$$u'(c_t) \geq E_t[u'(c_{t+1})]$$

From strict convexity of the marginal utility, by Jensen's inequality

$$u'(c_t) \geq E_t[u'(c_{t+1})] > u'[E_t(c_{t+1})]$$

which by strict concavity of utility function implies

$$E_t(c_{t+1}) > c_t$$

Therefore, consumption will always tend to ratchet upward over time and never converge.

Stochastic Income: general case

The result that consumption diverges to infinity with an i.i.d. endowment process is extended by Chamberlain and Wilson (2000) to an arbitrary stationary stochastic endowment process that is sufficiently stochastic. Let I_t denote the information set at time t. When $\beta(1+r)=1$, the general version of EE becomes

$$u'(c_t) \geq E[u'(c_{t+1}) \mid I_t]$$

where $E[\cdot \mid I_t]$ is the expectation operator conditioned upon information set I_t. Assuming a bounded utility function, Chamberlain and Wilson prove the following result:

Proposition 82 *If there is an $\varepsilon > 0$ such that for any $\alpha \in R$*

$$\Pr(\alpha \leq w_t \leq \alpha + \varepsilon \mid I_t) < 1 - \varepsilon$$

for all I_t and $t \geq 0$, then $\Pr(\lim_{t\to\infty} c_t = \infty) = 1$. Here x_t is the annuity value of the tail of the income process starting from period t

$$w_t \equiv \frac{r}{1+r} \sum_{j=t}^{\infty} (1+r)^{t-j} y_j .$$

Under certainty, the limiting value of the consumption path is given by the highest annuity of the endowment process across all starting dates t: $\bar{c} = \sup_t w_t$. Under uncertainty, Proposition above says that the consumption path will never converge to any finite limit if the annuity value of the endowment process is sufficiently stochastic. Instead, the optimal consumption path will then converge to infinity. This stark difference between the case of certainty and uncertainty is quite remarkable.

In summary, In presence of borrowing constraints and uncertain income, the condition $\beta(1+r) < 1$ is necessary for the optimal consumption sequence and for the asset space to be bounded.

10.3 Economies with Idiosyncratic Risk and Incomplete Markets

Let's go back to our target in this chapter: build a class of models whose equilibria feature a nontrivial endogenous distribution of income and wealth across agents. We are going to describe a particular type of incomplete markets model in this section. The models have a large number of *ex ante* identical but *ex post* heterogeneous agents who trade a single security. For most of this section, we study models with no aggregate uncertainty and no variation of an aggregate state variable over time (so macroeconomic time series variation is absent). But there is much uncertainty at the individual level. Households' only option is to "self-insure" by managing a stock of a single asset to buffer their consumption against adverse shocks. We study several models that differ mainly with respect to the particular asset that is the vehicle for self-insurance, for example, fiat currency or capital.

The tools for constructing these models are discrete-state discounted dynamic programming, used to formulate and solve problems of the individuals, and Markov chains, used to compute a stationary wealth distribution. The models produce a stationary wealth distribution that is determined simultaneously with various aggregates that are defined as means across corresponding individual-level variables.

We begin by recalling our discrete-state formulation of a single-agent infinite horizon savings problem. We then describe several economies in which households face some version of this infinite horizon savings problem, and where some of the prices taken parametrically in each household's problem are determined by the *average* behavior of all households.

This class of models was invented by Bewley (1977, 1980, 1983, 1986), partly to study a set of classic issues in monetary theory. Then Huggett (1993) and Aiyagari (1994) are the two pathbreaking papers that establish its quantitative significance in the field. Researchers have used calibrated versions of Bewley-Aiyagari models to give quantitative answers to questions including the welfare costs of business cycles (Imrohoroglu, 1989), the risk-sharing benefits of unfunded social security systems (Imrohoroglu, Imrohoroglu, and Joines, 1995), the benefits of insuring unemployed people (Hansen and Imrohoroglu, 1992), and the welfare costs of taxing capital (Aiyagari, 1995).

10.3.1 A Savings Problem

Recall the discrete-state savings problem as in section 4.2.1 and the previous section. The household's labor income at time t, s_t evolves according to an m-state Markov chain with transition matrix P. If the realization of the process at t is $\bar{s}_i$, then at time t the HH receives labor income $w\bar{s}_i$. Thus, employment opportunities determine the labor income process. We sometimes assume that m is 2, and that s_t takes the value 0 in an unemployed state and 1 in an employed state. This is the idiosyncratic risk (shock) in the model.

We constrain holdings of a *single* asset to a grid $\mathcal{A} = [a_0 = 0 < a_1 < a_2 < ... < a_n = a_{\max}]$. For given values of prices (w, r) and given initial values (a_0, s_0), the HH chooses a policy for $\{a_{t+1}\}_{t=0}^{\infty}$ to maximize

$$\begin{aligned} & E_0 \sum_{t=0}^{\infty} \beta^t u(c_t) \\ & s.t. \\ c_t + a_{t+1} &= (1+r)a_t + ws_t \\ c_t &\geq 0, a_{t+1} \in \mathcal{A}. \end{aligned}$$

$u(c)$ is a strictly increasing, strictly concave, twice continuously differentiable one-period utility function satisfying the Inada condition $\lim_{c \to 0} u'(c) = +\infty$ and $\beta(1+r) < 1$. (Recall that we learned from the previous section that this is a necessary condition for the optimal consumption sequence and for the asset space to be bounded.)

The Bellman eauation is $\forall i \in [1, 2, ..., m]$ and $\forall h \in [1, 2, ..., n]$

$$v(a_h, \bar{s}_i) = \max_{a' \in \mathcal{A}} \{U[(1+r)a_h + w\bar{s}_i - a'] + \beta \sum_{j=1}^{m} P_{ij} v(a', \bar{s}_j)\}$$

where a' is next period's value of asset holdings. Here $v(a, s)$ is the optimal value of the objective function, starting from asset-employment state (a, s). Note that the grid $\mathcal{A}$ incorporates upper and lower limits on the quantity that can be borrowed (i.e., the amount of the asset that can be issued). The upper bound on $\mathcal{A}$ is restrictive. In some of our theoretical discussion to follow, it will be important to dispense with that upper bound. We assume that there is no government, no physical capital or no supply or demand of bonds from abroad. Hence the net supply of assets in this economy is zero.

As we showed in the previous section, we can solve this BE for a value function $v(a, s)$ and an associated policy function $a' = g(a, s)$ mapping this period's (a, s) pair into an optimal choice of assets to carry into next period.

10.3.2 Wealth-employment Distributions

Let's define the unconditional distribution of *individual* state pairs at time t (a_t, s_t) to be

$$\lambda_t(a, s) = \Pr(a_t = a, s_t = s)$$

Notice that we have

$$\sum_a \sum_s \lambda_t(a, s) = 1, \forall t$$

This cross-sectional distribution over individual characteristics describes the *aggregate* state of the economy. The exogenous Markov chain P on s and the optimal policy function $a' = g(a, s)$ induce a law of motion for the distribution

λ_t as following

$$\begin{aligned} \Pr(a_{t+1} &= a', s_{t+1} = s') = \sum_{a_t} \sum_{s_t} \Pr(a_{t+1} = a' \mid a_t = a, s_t = s) \\ \cdot \Pr(s_{t+1} &= s' \mid s_t = s) \cdot \Pr(a_t = a, s_t = s). \end{aligned}$$

Or

$$\lambda_{t+1}(a', s') = \sum_{a} \sum_{s} \lambda_t(a, s) \cdot P_{ss'} \cdot \mathcal{I}(a', a, s)$$

where the indicator function $\mathcal{I}(a', a, s)$ is defined as

$$\mathcal{I}(a', a, s) = \begin{cases} 1 & \text{if } a' = g(a, s) \\ 0 & \text{otherwise} \end{cases}$$

The preceding equation can also be expressed concisely as

$$\lambda_{t+1}(a', s') = \sum_{\{a : a' = g(a,s)\}} \sum_{s} \lambda_t(a, s) \cdot P_{ss'} \tag{10.1}$$

Definition 83 *A time-invariant distribution λ that solves equation (10.1), i.e., one for which $\lambda_{t+1} = \lambda_t, \forall t$, is called a* ***stationary distribution****.*

A Digression on Markov Chain and Stationary Distribution

What is a Markov chain? Formally we have the following definition:

Definition 84 *A stochastic process $\{x_t\}$ is said to have the* ***Markov property*** *if for all $k \geq 1$ and all t,*

$$\Pr(x_{t+1} \mid x_t, x_{t-1}, ..., x_{t-k}) = \Pr(x_{t+1} \mid x_t).$$

We assume the Markov property and characterize the process by a *Markov chain.* A time-invariant Markov chain is defined by a triple of objects $(\bar{x}, P, \pi_0)$ where $\bar{x}$ is an n-demensional vector of possible state values, P is an $n \times n$ transition matrix which records the probabilities of moving from one value of the state to another in one period; and an $n \times 1$ vector π_0 whose ith element is the probability of being in state i at time 0: $\pi_0 = \Pr(x_0 = e_i)$.

The elements of transition matrix P are

$$P_{ij} = \Pr(x_{t+1} = e_j \mid x_t = e_i) \geq 0.$$

For these interpretations to be valid, the matrix P and the vector π_0 must satisfy the following assumption:

Assumption M: (1). $\forall i = 1, 2, ..., n$, the matrix P satisfies

$$\sum_{j=1}^{n} P_{ij} = 1.$$

(2). The vector π_0 satisfies

$$\sum_{i=1}^{n} \pi_{0i} = 1.$$

A matrix P that satisfies Assumption M is called a *stochastic matrix*. A stochastic matrix defines the probabilities of moving from one value of the state to another in one period. The probability of moving from one value of the state to another in two periods is determined by P^2 because

$$\begin{aligned}\Pr(x_{t+2} &= e_j \mid x_t = e_i) \\ &= \sum_{h=1}^{n} \Pr(x_{t+2} = e_j \mid x_{t+1} = e_h) \cdot \Pr(x_{t+1} = e_h \mid x_t = e_i) \\ &= \sum_{h=1}^{n} P_{ih} \cdot P_{hj} = P_{ij}^{(2)},\end{aligned}$$

where $P_{ij}^{(2)}$ is the (i, j) element of $P^2 = P \cdot P$. Continuing this process, we have

$$\Pr(x_{t+k} = e_j \mid x_t = e_i) = P_{ij}^{(k)}.$$

Therefore, the unconditional probability distributions of x_t are determined by

$$\begin{aligned}\underbrace{\pi_1'}_{1\times n} &= \Pr(x_1) = \underbrace{\pi_0'}_{1\times n} \cdot \underbrace{P}_{n\times n} \\ \pi_2' &= \Pr(x_2) = \pi_0' \cdot P^2 \\ &\ldots\ldots \\ \pi_k' &= \Pr(x_k) = \pi_0' \cdot P^k\end{aligned}$$

where $\pi_t' = \Pr(x_t)$ is the $1 \times n$ vector whose ith element is $\Pr(x_t = e_i)$. More concisely

$$\pi_{t+1}' = \pi_t' \cdot P. \tag{10.2}$$

An unconditional distribution is called *stationary* or *invariant* if it satisfies

$$\pi_{t+1} = \pi_t,$$

that is, if the unconditional distribution remains unaltered with the passage of time. From the law of motion (10.2) for unconditional distributions, a stationary distribution must satisfy

$$\pi' = \pi' P$$

$\Rightarrow$

$$\pi'(I - P) = 0$$

where I is an $n \times n$ identity matrix. Transposing both sides of this equation gives

$$(I - P')\pi = 0 \tag{10.3}$$

which implies π as an eigenvector associated with matrix P'.

The fact that P is a stochastic matrix (i.e., it has nonnegative elements and satisfies $\sum_{j=1}^{n} P_{ij} = 1, \forall i, j$) guarantees that P has at least one unit eigenvalue, and that there is at least one eigenvector π that satisfies equation (10.3). This stationary distribution may not be unique because P can have a repeated unit eigenvalue.

Example 85 *A Markov chain*

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0.2 & 0.5 & 0.3 \\ 0 & 0 & 1 \end{bmatrix}$$

has two unit eigenvalues with associated two eigenvectors $\pi' = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$ *and* $\pi' = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$. *You can easily confirm that by computing* $\pi' \cdot P = \pi'$. *Or you can use MATLAB function [V,D]=eig(P') to confirm it. This function returns matrices of eigenvalues (D) and eigenvectors (V) of matrix P', so that* $P'V = VD$. *In this case,*

$$\begin{aligned} V &= \begin{bmatrix} 1 & 0 & -0.3244 \\ 0 & 0 & 0.8111 \\ 0 & 1 & -0.4867 \end{bmatrix} \\ D &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.5 \end{bmatrix} \end{aligned}$$

The diagonal elements of matrix D are three eigenvalues. Here state 1 and 3 are both absorbing states. Furthermore, any initial distribution that puts zero probability on state 2 is a stationary distribution. Fo example, given $\pi_0' = \begin{bmatrix} 0.5 & 0 & 0.5 \end{bmatrix}$, *we have* $\pi_0' \cdot P = \pi_0'$.

Example 86 *A Markov chain*

$$P = \begin{bmatrix} 0.7 & 0.3 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 0.9 & 0.1 \end{bmatrix}$$

has one unit eigenvalue with associated eigenvector $\pi' = \begin{bmatrix} 0 & 0.6429 & 0.3571 \end{bmatrix}$. *Here states 2 and 3 form an absorbing subset of the state space.*

We often ask the following question about a Markov process: for an arbitrary initial distribution π_0, do the unconditional distributions π_t approach a stationary distribution

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

where π_∞ solves equation (10.3)? If the answer is yes, then does the limit distribution π_∞ depend on the initial distribution π_0? If the limit π_∞ is independent of the initial distribution π_0, we say that the process is *asymptotically stationary with a unique invariant distribution.* We call a solution π_∞ a *stationary distribution* or an *invariant distribution* of P.

We state these concepts formally in the following definition:

Definition 87 *Let π_∞ be a unique vector that satisfies $(I - P')\pi_\infty = 0$. If for all initial distributions π_0 it is true that $P^{t\prime}\pi_0$ converges to the same π_∞, we say that the Markov chain is asymptotically stationary with a unique invariant distribution.*

When does a Markov chain is asymptotically stationary with a unique invariant distribution? We have following theorems.

Theorem 88 *Let P be a stochastic matrix with $P_{ij} > 0, \forall i, j$. Then P has a unique stationary distribution, and the process is asymptotically stationary.*

Theorem 89 *Let P be a stochastic matrix with $P_{ij}^{(k)} > 0, \forall i, j$ for some $k \geq 1$. Then P has a unique stationary distribution, and the process is asymptotically stationary.*

The conditions of Theorems above state that from any state there is a positive probability of moving to any other state in one (or k) steps.

Computing Invariant Distribution

Now let's go back to the law of motion of distribution in the incomplete market model equation (10.1). From the digression above, we know there are two ways to compute a stationary distribution in this equation. One way is to directly iterate to convergence on equation (10.1). An alternative way is to create a Markov chain that describes the solution of the optimum problem, then to compute an invariant distribution from a left eigenvector associated with a unit eigenvalue of the stochastic matrix.

Here we go for the second method. To deduce this Markov chain, we map (or"stack") the pair (a, s) of vectors into a single state vector x as follows. For $i = 1, 2, ..., n,\ h = 1, 2, ..., m$, let the jth element of the $n \times m$ vector x be the pair (a_i, s_h) where $j = (i-1)m + h$. Thus we denote $x' = [\ (a_1, s_1),\ (a_1, s_2),\ ...,\ (a_1, s_m),\ (a_2, s_1),\ (a_2, s_2),\ ...,\ (a_2, s_m),\ ...,\ (a_n, s_1),\ ..,\ (a_n, s_m)\ ]$. As we said above, the optimal policy function $a' = g(a, s)$ and the Markov chain P on s induce a Markov chain on x_t via the formula

$$\begin{aligned}
\Pr[(a_{t+1} &= a', s_{t+1} = s') \mid (a_t = a, s_t = s)] \\
&= \Pr(a_{t+1} = a' \mid a_t = a, s_t = s) \cdot \Pr(s_{t+1} = s' \mid s_t = s) \\
&= \mathcal{I}(a', a, s) \cdot P_{ss'}
\end{aligned}$$

This formula defines an $N \times N$ matrix $\mathcal{P}$, where $N = n \times m$. This is the Markov chain on the household's state vector x.

Now finding an invariant distribution implies

$$x' = x'\mathcal{P}$$

where again $\mathcal{P}$ is the Markov chain on the HH's state vector x. Suppose that the Markov chain associated with $\mathcal{P}$ is asymptotically stationary and has a unique invariant distribution π_∞, which is an $n \times 1$ vector. Typically, all states in the Markov chain will be recurrent, and the individual will occasionally revisit each state. The distribution π_∞ tells the fraction of time that the HH spends in each state. Once we get it, we can "unstack" the state vector x and use π_∞ to deduce the stationary probability measure $\lambda(a_i, s_h)$ over (a, s) pairs

$$\lambda(a_i, s_h) = \Pr(a_t = a_i, s_t = s_h) = \pi_\infty(j)$$

where $j = (i-1)m + h$ is the jth element of the vector π_∞. Notice that at a particular point of time t, $\lambda(a_i, s_h)$ also measures the fraction of people who has state vector (a_i, s_h), i.e., it is also a probability distribution of all HHs over state (a, s).

10.3.3 Stationary Equilibrium

Here we basically summarize the setup as in Aiyagari (1994). He used a version of the savings problem in an economy with many agents and interpreted the single asset as homogeneous physical capital, denoted k. The law of motion for capital holdings is

$$k_{t+1} = (1-\delta)k_t + i_t$$

The HH's BC is

$$c_t + i_t = \tilde{r}k_t + ws_t,$$

where $\tilde{r}$ is the rental rate on capital and w is a competitive wage, both will be determined later. Combining these two equations, we have

$$c_t + k_{t+1} = (1 + \tilde{r} - \delta)k_t + ws_t$$

This BC coincides with the one in the benchmark savings problem if we define $a_t \equiv k_t$ and $r \equiv \tilde{r} - \delta$.

In Aiyagari model, there is a large number of HHs with *identical* preferences whose distribution across (k, s) pairs is given by $\lambda(k, s)$, and whose *average* behavior determines prices (w, r) as follows: HHs are identical in their preferences, the Markov processes governing their employment opportunities, and the prices that they face. However, they differ in their histories $s_0^t = \{s_h\}_{h=0}^t$ of employment opportunities, and therefore in the capital that they have accumulated. Each household has its own history s_0^t as well as its own initial capital k_0. The productivity processes are assumed to be independent across households. The behavior of the collection of these HHs determines the wage and interest rate (w, r).

Assume an initial distribution across HHs of $\lambda(k, s)$. The *average* level of capital per HH K satisfies

$$K = \sum_k \sum_s \lambda(k, s) g(k, s),$$

where the optimal policy function is

$$k' = g(k, s).$$

The average level of employment is

$$N = \sum_k \sum_s \lambda(k, s) \bar{s} \tag{10.4}$$

where s is the exogenously specified vector of individual employment rates (we can also view it as efficiency units of labor, i.e., idiosyncratic labor productivity). It is constant over time (remember there is no aggregate uncertainty). In other words, aggregate labor supply in this economy is constant.

There is an CRS aggregate production function

$$F(K, N) = AK^{\alpha} N^{1-\alpha}$$

The production function determines the rental rates on capital and labor from the marginal conditions

$$\tilde{r} = \frac{\partial F(K, N)}{\partial K} \tag{10.5}$$

$$w = \frac{\partial F(K, N)}{\partial N} \tag{10.6}$$

Now we are ready to define a stationary equilibrium in this economy.

Definition 90 *A* ***stationary recursive equilibrium*** *is a policy function $g(k, s)$ and $c(k, s)$, a value function $v(k, s)$, a probability distribution $\lambda(k, s)$, and positive real numbers $(K, N, \tilde{r}, w)$ such that*

(i). Given $(\tilde{r}, w)$, the policy function $g(k, s)$, $c(k, s)$and the value function $v(k, s)$ solve the HH's problem.

$$\max E_0 \sum_{t=0}^{\infty} \beta^t u(c_t)$$

$$s.t.$$

$$\begin{aligned} c_t + k_{t+1} &= (1 + \tilde{r} - \delta)k_t + ws_t \\ c_t, k_{t+1} &\geq 0. \end{aligned}$$

(ii). The prices $(\tilde{r}, w)$ satisfies the firm's FOCs (10.5)-(10.6).

(iii). The probability distribution $\lambda(k, s)$ is a stationary distribution associated with policy function $g(k, s)$ and Markov chain P on s; that is, it satisfies

$$\lambda(k', s') = \sum_{\{k: k' = g(k,s)\}} \sum_s \lambda(k, s) \cdot P_{ss'}$$

(iv). Individual and aggregate behavior are consistent:

$$
\begin{aligned}
K &= \sum_k \sum_s \lambda(k,s) g(k,s) \\
N &= \sum_k \sum_s \lambda(k,s) \bar{s}.
\end{aligned}
$$

(v). Goods market clears.

$$
\begin{aligned}
\sum_k \sum_s \lambda(k,s) c(k,s) + \sum_k \sum_s \lambda(k,s) i(k,s) &= \sum_k \sum_s \lambda(k,s) \tilde{r} g(k,s) + \sum_k \sum_s \lambda(k,s) w \bar{s} \\
&\quad or \\
C + I &= \tilde{r}K + wN = Y.
\end{aligned}
$$

In this equilibrium, the individual-wide state variables are (k_t, s_t), the economy-wide state variable is K_t. Therefore, for individuals the state vector is (K_t, k_t, s_t). Notice that since the distribution λ is time-invariant, it does not enter into the state space. This dramatically reduces the "curse of dimensionality" problem. Given the stationary property, in the equilibrium, *aggregate* capital and labor are constant, so are interest rate and wage rate. But the indivudual variables are different depending on their status. In other words, we have individual heterogeneity and aggregate stability.

10.3.4 Computing Stationary Equilibrium

Aiyagari (1994) computed an equilibrium of the model above by defining a mapping from $K \in R$ into R, with the property that a fixed point of the mapping is an equilibrium capital stock K. Here is an algorithm for finding a fixed point:

1. Guess the initial capital stock $K = K_j$ with $j = 0$. Make some initial guess of distribution $\lambda(k,s)$, using equation (10.4) to obtain the total labor input N_j. Then using (10.5)-(10.6) to obtain w and $\tilde{r}$.

2. Given the prices, now we can slove the HH's optimization problem and obtain optimal policy function $g_j(k,s)$. Now use this decision rule to duduce an associated stationary distribution $\lambda_j(k,s)$ by following the law of motion of distribution (10.1).

3. Compute the new aggregate capital stock associated with $\lambda_j(k,s)$, namely

$$K_j^* = \sum_k \sum_s \lambda_j(k,s) g_j(k,s)$$

If

$$\left\| K_j^* - K_j \right\| < \varepsilon$$

Stop. Otherwise, for a fixed "relaxation parameter" $\xi \in (0,1)$, compute a new estimate of K

$$K_{j+1} = \xi K_j + (1-\xi) K_j^*.$$

and go back to step 1.

4. Iterate on this scheme to the convergence.

10.3.5 Borrowing Constraint: Natural vs. Ad Hoc

Now let's go back to our savings problem and change our borrowing constraint to a broader one

$$\begin{aligned} & E_0 \sum_{t=0}^{\infty} \beta^t u(c_t) \\ & s.t. \\ c_t + a_{t+1} &= (1+r)a_t + ws_t \quad \text{(BC)} \\ a_{t+1} &\geq -\phi \quad \text{(borrowing constraint)} \\ c_t &\geq 0. \end{aligned}$$

We consider two types of borrowing constraints here.

First of all, imposing $c_t \geq 0, \forall t$ implies in BC

$$a_{t+1} \leq (1+r)a_t + ws_t$$

Or

$$a_t \geq \frac{1}{1+r}(a_{t+1} - ws_t)$$

which by iteration implies

$$\begin{aligned} a_t &\geq \frac{1}{1+r}([\frac{1}{1+r}(a_{t+2} - ws_{t+1})] - ws_t) \\ &= -\frac{1}{1+r}ws_t - (\frac{1}{1+r})^2 ws_{t+1} + (\frac{1}{1+r})^2 a_{t+2} \\ &= \ldots \\ &= -\frac{1}{1+r}\sum_{j=0}^{T} ws_{t+j}(1+r)^{-j} + (\frac{1}{1+r})^{T+1} a_{t+T+1} \end{aligned}$$

Take limit $T \to \infty$, impose the transversality condition (or non-Ponzi condition), we have

$$a_t \geq -\frac{1}{1+r}\sum_{j=0}^{\infty} ws_{t+j}(1+r)^{-j}. \tag{10.7}$$

Since the right hand side of inequality above is a random variable, not known at time t, we have to supplement equation (10.7) to obtain the borrowing constraint. One possible approach is to replace the RHS of equation (10.7) with its conditional expectation, and to require equation (10.7) to hold in expected value. But this expected value formulation is incompatible with the notion that the loan is risk free, and that the household can repay it for sure. If we insist that equation (10.7) hold *almost surely*, for all $t \geq 0$, then we obtain the constraint that emerges by replacing s_t with $\min s \equiv s_1$, which yields

$$\begin{aligned} a_t &\geq -\frac{1}{1+r}\sum_{j=0}^{\infty} w s_1 (1+r)^{-j} \\ &= -\frac{s_1 w}{r}. \end{aligned}$$

Aiyagari (1994) call it the *natural borrowing limit*. Note that $\frac{s_1 w}{r}$ is the maximum amount that the HH can repay for sure without violating nonnegativity of sonsumption. If this borrowing limit holds, the inequality in (10.7) definitely holds.

If the borrowing limit b as defined

$$a_t \geq -b$$

is tighter than the natural borrowing limit (e.g., $b = 0$ so that no borrowing is permitted), then the borrowing limit b may be regarded as *ad hoc* in the sense that it is not a consequence of nonnegativity of consumption. Combining these two borrowing limits, we have

$$a_t \geq -\phi$$

where

$$\phi = \min[b, \frac{s_1 w}{r}].$$

10.3.6 Equilibrium Interest Rate and Savings under Idiosyncratic Risk

For the special case of Aiyagari model in which s is i.i.d., Aiyagari showed how to cast the model in terms of a single state variable to appear in the HH's value function. To synthesize a single state variable, note that the maximum "disposable resources" available to be allocated at t are

$$z_t = (1+r)a_t + w s_t + \phi$$

i.e., borrow to the limit $a_{t+1} = -\phi$. Define $\hat{a}_t \equiv a_t + \phi$, we have

$$z_t = w s_t + (1+r)\hat{a}_t - r\phi.$$

In terms of the single state variable z_t, the HH's BC can be represented recursively as

$$\begin{aligned} c_t + \hat{a}_{t+1} &\leq z_t \\ z_{t+1} &= w s_{t+1} + (1+r)\hat{a}_{t+1} - r\phi \\ c_t &\geq 0, \hat{a}_{t+1} \geq 0 \end{aligned}$$

The FE is

$$
\begin{aligned}
v(z_t) &= \max_{\hat{a}_{t+1} \geq 0} \{u(z_t - \hat{a}_{t+1}) + \beta E v(z_{t+1})\} \\
& \quad s.t. \\
z_{t+1} &= w s_{t+1} + (1+r)\hat{a}_{t+1} - r\phi.
\end{aligned} \tag{10.8}
$$

The optimal asset demand rule (or policy function) is obtained by solving the FE above. This yields the following single-valued asset demand function

$$\hat{a}_{t+1} = A(z_t).$$

The EE to this FE is

$$
\begin{aligned}
u'(c_t) &\geq \beta(1+r)E_t u'(c_{t+1}) \\
&= \beta(1+r)E_t u'(c_{t+1}), \text{ if } \hat{a}_{t+1} > 0
\end{aligned} \tag{10.9}
$$

We can use equation (10.9) to deduce significant aspects of the limiting behavior of mean assets $E\hat{a}$ as a funtion of interest rate r from the HH's side. Define

$$M_t = [\beta(1+r)]^t u'(c_t) \geq 0$$

Thus the EE is equivalent to the following

$$M_t \geq E_t M_{t+1}$$

which asserts that M_t is a supermartingale. As we showed in section 10.2.2, the supermartingale converegnce theorem applies here to guarantee that M_t converges almost surely to a nonnegative random variable $\bar{M}$: $M_t \rightarrow_{a.s.} \bar{M}$.

As in the section 10.2.2, let's consider the following three cases.

1. $\beta(1+r) > 1$. Since $M_t = [\beta(1+r)]^t u'(c_t)$, the fact that M_t converges implies that $u'(c_t)$ has to converge to zero almost surely, which in turn implies that $c_t \rightarrow_{a.s.} +\infty$ and that the consumer's asset holdings must be diverging to $+\infty$ too.

2. $\beta(1+r) = 1$. As we showed in section 10.2.2, we have the same divergence results here as in case 1.

3. $\beta(1+r) < 1$. convergence of Mt leaves open the possibility that $u'(c)$ does not converge almost surely, that it remains finite and continues to vary randomly. Indeed, when $\beta(1+r) < 1$, the average level of assets remains finite, and so does the level of consumption.

It is easier to analyze the borderline case $\beta(1+r) = 1$ in the special case that the employment process s_t is i.i.d.. Aiyagari (1994)[4] uses the following argument by contradiction to show that if $\beta(1+r) = 1$, then total available

[4]See Aiyagari (1994) footnote 21.

resource z_t diverges to $+\infty$. Assume that there is some upper limit $z_{\max}$ such that

$$z_{t+1} \leq z_{\max} = ws_{\max} + (1+r)A(z_{\max}) - r\phi.$$

Then when $\beta(1+r) = 1$, the strict concavity of the value function, the Benveniste-Scheinkman formula, and Euler equation (10.9) imply

$$\begin{aligned} v'(z_{\max}) &\geq \beta(1+r)E_t v'(z' = ws' + (1+r)A(z_{\max}) - r\phi) \\ &> E_t v'[ws_{\max} + (1+r)A(z_{\max}) - r\phi] \\ &= v'[(ws_{\max} + (1+r)A(z_{\max}) - r\phi)] \\ &= v'(z_{\max}) \end{aligned}$$

The first inequality is the FOC w.r.t. $\hat{a}_{t+1}$ for the FE. The second inequality comes from the fact that the value function is strictly concave, i.e., $v'(\cdot)$ is strictly decreasing, and $z' < z_{\max}$ (since $Es' < s_{\max}$). Contradiction!

Define $\beta = \frac{1}{1+\rho}$, i.e., ρ is the time discount rate. We can use a similar graph as in Aiyagari (1994) to characterize the asset demand function $Ea(r)$. The average asset demand curve, as a function of the interest rate, is upward sloping, i.e., $Ea'(r) > 0$ (this is because $A'(z) > 0$, asset holding is a positive function of total resource, and $\frac{\partial z}{\partial r} > 0$). It is equal to $-b$ for sufficiently low r (if the interest rate is low enough, everybody wants to borrow hence hit the borrow limit), asymptotes towards $+\infty$ as r approaches ρ from below.

Now let's talk about equilibrium. Define the capital stock as a function of interest rate r $K(r)$. From the firm's side, we know $K(r)$ is implicitly determined by the firm's FOC

$$\begin{aligned} r &= \frac{\partial F(K,N)}{\partial K} - \delta \\ &= F_K - \delta \end{aligned}$$

Notice that since production function F is strictly concave, marginal product of capital is strictly decreasing, therefore, $K(r)$ is downward-slopping.

The equilibrium is achieved when the $Ea(r)$ curve intersects the $F_K - \delta$ curve. Thus (r_1, K_1) is our equilibrium interest rate and capital stock in this economy with idiosyncratic shock and borrowing limit. Notice that when there is no uncertainty, consumption and assets converge to finite number if $\beta(1+r) \leq 1$, thus the asset demand function $a(r)$ does not depend on r. Therefore, the horizontal line crosses the r-axis at ρ is $a(r)$ curve under certainty[5]. The equilibrium interest rate is $r = \rho$ and the equilibrium capital stock under certainty is K_0.

From Figure 10.3, clearly under uncertainty, the equilibrium interest rate r_1 is lower than its certainty counterpart, and its equilibrium capital stock K_1 is higher than the equilibrium capital stock under full insurance (certainty)

[5] More accurately, when $r < \rho$, i.e., $\beta(1+r) < 1$, interest rate is too low, hence the HHs always want to borrow up to the limit, $a(r) = -\phi$. When $r = \rho$, i.e., $\beta(1+r) = 1$, the HH's asset holdings equal to his initial assets a_0 whatever they might be. Therefore, the asset demand curve is actually a two-part curve. When $r < \rho$, it is the vertical line through $-\phi$ on x-axis. When $r = \rho$, it is the horizontal curve through ρ on y-axis.

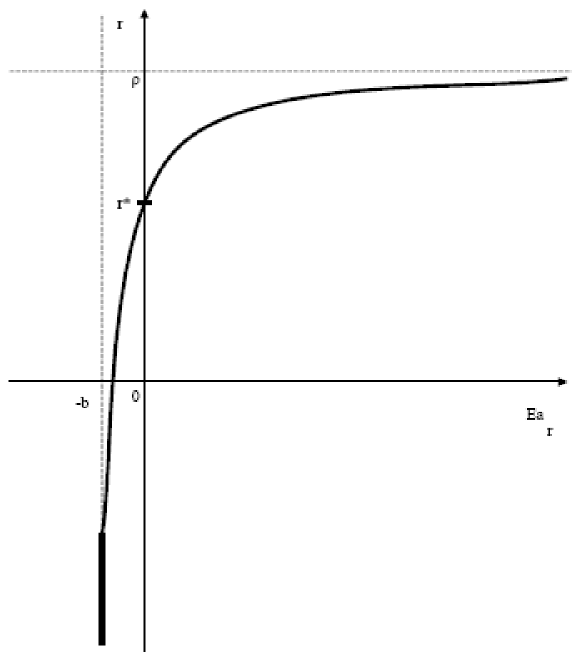


Figure 10.2: Average assets as a function of r in Aiyagari economy

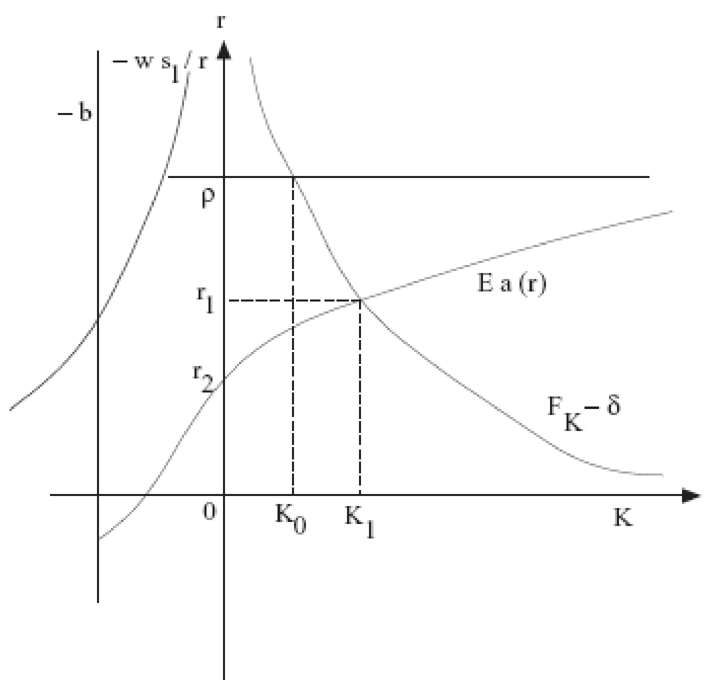


Figure 10.3: Demand for capital and determination of equilibrium interest rate

K_0. Uninsured idiosyncratic risk brings the *overaccumulation* of capital stock. Obviously, if $b = 0$, i.e., if we do not allow borrow at all ($a_t \geq 0, \forall t$), the difference between K_1 and K_0 will be even bigger.

This completes our description of the theoretical features of the Bewley-Aiyagari models. Now we will turn to the quantitative results that applications of these models have delivered.

10.3.7 Numerical Results

In this subsection, we are going to show some quantitative results of Aiyagari model.

The model period is taken to be one year. The discount factor β is chosen to be 0.96. The production function $F(K, N)$ is assumed to be Cobb-Douglas $F(K, N) = K^{\alpha}N^{1-\alpha}$ with the capital share $\alpha = 0.36$. The depreciation rate of capital δ is set at 8%. The period utility function is taken as CRRA

$$u(c) = \frac{c^{1-\mu} - 1}{1 - \mu}$$

where μ is the relative risk aversion coefficient. Results are reported for three different values of $\mu \in \{1, 3, 5\}$. The above technology and preference specifications and parameter values are chosen to be consistent with aggregate features of the postwar U.S. economy and are also commonly used in aggregative models of growth and business cycles.

Next, we use Tauchen's (1986) method to get a discrete-state Markov chain to approximate a first-order autoregressive process

$$\log s_t = \rho \log s_{t-1} + \sigma\sqrt{1 - \rho^2}\varepsilon_t$$

where ε_t is a sequence of i.i.d. random variables that belongs to a standard normal distribution $N(0, 1)$. We take $\sigma \in \{0.2, 0.4\}$ and $\rho \in \{0, 0.3, 0.6, 0.9\}$. This process is approximated by a seven-state Markov chain.

Finally, the borrowing limit b is set to be zero, i.e., borrowing is prohibited.

The algorithm for approximating the steady state uses simulated series and the bisection method by following steps below.

1. Start from some value of interest rate r_1 that is close but less than the rate of time preference $\rho = \frac{1}{\beta} - 1 = 0.04167$. Given r_1, we compute the asset demand function.

2. Simulate the Markov chain for the shock s_t using a random number generator and obtain a series of 10000 draws. These are used with the asset demand function to obtain a simulated series of assets. Ea is calculated as the sample mean of these series. We then calculate r_2 such that $K(r_2) = Ea$.

3. If $r_2 = r_1$, stop. Otherwise, define

$$r_3 = \frac{r_1 + r_2}{2}$$

A. Net return to capital (%) / Aggregate saving rate (%): $\sigma = 0.2$			
$\rho \backslash \mu$	1	3	5
0	4.1666/23.67	4.1456/23.71	4.0858/23.83
0.3	4.1365/23.73	4.0432/23.91	3.9054/24.19
0.6	4.0912/23.82	3.8767/24.25	3.5857/24.86
0.9	3.9305/24.14	3.2903/25.51	2.5260/27.36
B. Net return to capital (%) / Aggregate saving rate (%): $\sigma = 0.4$			
$\rho \backslash \mu$	1	3	5
0	4.0649/23.87	3.7816/24.44	3.4177/25.22
0.3	3.9554/24.09	3.4188/25.22	2.8032/26.66
0.6	3.7567/24.50	2.7835/26.71	1.8070/29.37
0.9	3.3054/25.47	1.2894/31.00	-0.3456/37.63

Table 10.5: Equilibrium interest rate and savings rate in Aiyagari model

Go back to step 1. Iterate until convergence.

Once the steady state is approximated, we use the solution to calculate the following objects of interest. We calculate the mean, median, standard deviation, coefficient of variation (c.v.), skewness, and serial correlation coefficient for labor income, asset (capital) holdings, net income, gross income, gross saving, and consumption. We also calculate measures of inequality for each of these variables. We use the simulated series for each variable to construct its distribution, and then we compute the Lorenz curves and calculate the associated Gini coefficients.

Here in Table 10.5[6], we report the net return to capital (r) and the savings rate which is the investment/output ratio in stationary equilibrium of Aiyagari model given all the parameter values as we described above. For comparison, it is easy to calculate that the net return to capital and the savings rate under the certainty case are 4.17% and 23.67% respectively.

The main message from this table is under idiosyncratic risk, r does go down and savings rate does increase for all the combination of serial correlation coefficient ρ and CRRA coefficient μ.

Since long-run distributions for an individual coincide with cross-section distributions for the population, results for variabilities of individual consumption, income, and assets have immediate implications for cross-section distributions. These results are qualitatively consistent with casual empiricism and more careful empirical observation. There is much less dispersion across households in consumption compared with income and much greater dispersion in wealth compared with income For example, when $\sigma = 0.4$, $\rho = 0.9$, and $\mu = 1$, the model generates the coefficients of variation (cv) of consumption, gross income, and gross savings are 0.37, 0.59, and 1.33. But they are quite lower than the data, which are 3.86 for income and 6.09 for wealth. In all cases, the fraction of liquidity-constrained households is close to zero. Skewness coefficients reveal

[6] It replicates Table 2 in Aiyagari (1994).

another aspect of inequality. All of the cross-section distributions are positively skewed (median<mean). However, the degree of skewness is somewhat less than in the data. For example, median net income, gross income, and capital are all over 90 percent of their respective mean values. In contrast, U. S. median household income is about 80 percent of U. S. mean household income. Lorenz curves and Gini coefficients show that the model does generate empirically plausible relative degrees of inequality. Consumption exhibits the least inequality, followed by net income, gross income, and then capital, and saving exhibits the greatest inequality. However, the model cannot generate the observed degrees of inequality. For example, when μ is 5, ρ is 0.6, and σ is 0.2, the Gini coefficients for net income and wealth are 0.12 and 0.32, respectively. In U. S. data, however, the Gini coefficient for income is about 0.57 and that for net wealth is about 0.78.

Overall, the model underachieves in terms of income and wealth dispersion. This is largely because it fails to generate the fat upper tail of the wealth distribution. There have been several suggestions to cure this failure; for example to introduce potential entrepreneurs that have to accumulate a lot of wealth before financing investment projects. But this is still an open question for future research.

10.4 Extension of Aiyagari Model with Aggregate Shocks

Aiyagari model is absent of aggregate shock. This helps makes the model tractable since the distribution does not enter into the state space. This section describes a way to extend such models to situations with time-varying stochastic aggregate state variables.

Krusell and Smith (1998) modified Aiyagari's (1994) model by adding an aggregate state variable z, a technology shock that follows a Markov process. Each household continues to receive an idiosyncratic labor-endowment shock s that averages to the same constant value for each value of the aggregate shock z. The aggregate shock causes the size of the state of the economy to expand dramatically, because every HH's wealth will depend on the *history* of the aggregate shock z, call it z^t, as well as the history of the household-specific shock s^t. That makes the joint histories of z^t, s^t correlated across households, which in turn makes the cross-section distribution of (k, s) vary randomly over time. Therefore, the interest rate and wage will also vary randomly over time.

One way to specify the state is to include the cross-section distribution $\lambda(k, s)$ each period among the state variables. Thus, the state includes a cross-section probability distribution of (capital, employment) pairs. In addition, a description of a recursive competitive equilibrium must include a law of motion mapping today's distribution $\lambda(k, s)$ into tomorrow's distribution.

Let's recall the original Aiyagari model. The FE of the model is

$$\begin{aligned} v(k,s) &= \max_{c,k'}\{u(c)+\beta E[v(k',s')]\} \\ & \quad s.t. \\ c+k' &= (1+\tilde{r}-\delta)k+ws \end{aligned}$$

where the wage w and interest rate $\tilde{r}$ are determined by the firm's FOCs

$$\begin{aligned} \tilde{r} &= \frac{\partial F(K,N)}{\partial K} = \alpha(\frac{K}{N})^{\alpha-1} \\ w &= \frac{\partial F(K,N)}{\partial N} = (1-\alpha)(\frac{K}{N})^{\alpha}. \end{aligned}$$

And the total capital stock and labor input K and N are the average values of k and s over the stationary distribution $\lambda(k,s)$

$$\begin{aligned} K &= \int k\lambda(k,s)dkds \\ N &= \int s\lambda(k,s)dkds. \end{aligned}$$

10.4.1 Krusell-Smith's Model Economy

Krusell and Smith (1998) consider a version of stochastic growth model with a large (measure one) population of infinitely lived consumers. The preference is given by

$$\max E_0 \sum_{t=0}^{\infty} \beta^t u(c_t)$$

where period utility function takes the CRRA form

$$u(c) = \frac{c^{1-\mu}-1}{1-\mu}.$$

The BC is standard as the one in Aiyagari (1994).

$$c+k' = (1+\tilde{r}-\delta)k+ws.$$

However, the aggregate production is subject to a stochastic shock z:

$$Y = zF(K,N) = zK^{\alpha}N^{1-\alpha}.$$

Thus the wage and interest rate are the function of K, N and z:

$$\tilde{r} = \tilde{r}(K,N,z) = z\alpha(\frac{K}{N})^{\alpha-1} \tag{10.10}$$

$$w = w(K,N,z) = z(1-\alpha)(\frac{K}{N})^{\alpha}. \tag{10.11}$$

Let λ denote the current measure (distribution) of HHs over holdings of capital k and employment status s. Then the law of motion of the distribution now is

$$\lambda' = H(\lambda, z, z'). \tag{10.12}$$

The aggregate shock today and tomorrow will affect the transition of distribution. Therefore, we do not have time-invariant distribution any more. The distribution itself is a time-variant random variable distributed by the aggregate shock z. Thus at time t, the aggregate capital and labor input are

$$\begin{aligned} K_t &= \int k\lambda_t(k,s)dkds \qquad (10.13) \\ N_t &= \int s\lambda_t(k,s)dkds. \qquad (10.14) \end{aligned}$$

In this economy, the aggregate state variables are the distribution and aggregate shock (λ, z). Notice that in the complete market representative agent model, the aggregate state are total capital K and shock z (see section 1.4.5). But now in this incomplete market economy, since different HHs have different amount of wealth and their propensities to save are not equal, the distribution of wealth will affect the total savings. It is clearly seen from the equation (10.13) above. Thus the state variables are indeed (K, λ, z). But since K is determined by λ (of course also the individual state k and s), we can ignore K here to save one state variable.

For the individual HH, the relevant state variables are his asset holdings k, his employment status s, and the aggregate state. The role of the aggregate state is to allow the HHs to predict future prices. We thus can write the HH's optimization problem as a FE below

$$\begin{aligned} v(k,s;\lambda,z) &= \max_{c,k'}\{u(c) + \beta E[v(k',s';\lambda',z') \mid (s,z,\lambda)]\} \\ & \quad s.t. \\ c + k' &= (1 + \tilde{r}(K,N,z) - \delta)k + w(K,N,z)s \qquad (10.15) \\ K &= \int k\lambda(k,s)dkds, N = \int s\lambda(k,s)dkds \\ \lambda' &= H(\lambda,z,z') \\ c,k' &\geq 0 \end{aligned}$$

The solution to this FE is a decision rule for consumption $c = c(k,s;\lambda,z)$ and a decision rule for savings $k' = g(k,s;\lambda,z)$.

10.4.2 Recursive Competitive Equilibrium

Now we are ready to define a RCE for this economy.

Definition 91 *A RCE is a value function $v(k,s;\lambda,z)$, a decision rule $c(k,s;\lambda,z)$ and $g(k,s;\lambda,z)$, a pair of price functions $(\tilde{r}, w)$, and a law of motion H for distribution $\lambda(k,s)$ such that:*

(i). Given prices $(\tilde{r}, w)$*,* v, c, g *solve the HH's DP problem (10.15).*

(ii). $\tilde{r}$ *and* w *are the solutions to the firm's profit-maximization problem, i.e.,* $\tilde{r}$ *and* w *satisfy (10.10)-(10.11).*

(iii). The decision rule g *and the Markov processes for* s *and* z *imply that today's distribution* $\lambda(k,s)$ *is mapped into tomorrow's* $\lambda'(k,s)$ *by* H *as in equation (10.12)*

$$\begin{aligned} \lambda' &= H(\lambda, z, z') \\ &= \int Q_{z,z'}(k,s)\lambda(k,s)dkds \end{aligned}$$

where $Q_{z,z'}$ *is the transition function between two periods where the aggregate shock goes from* z *to* z' *and is defined by*

$$Q_{z,z'}(k,s) = \sum_{s'} \mathcal{I}(k', k, s, z, \lambda)\pi(z', s' \mid z, s)$$

where $\mathcal{I}(k', k, s, z, \lambda)$ *is an idcator function defined below*

$$\mathcal{I}(k', k, s, z, \lambda) = \begin{cases} 1 & \text{if } k' = g(k, s; \lambda, z) \\ 0 & \text{otherwise} \end{cases}$$

(iv). Consistency or market clearing conditions

$$\begin{aligned} K &= \int k\lambda(k,s)dkds \\ N &= \int s\lambda(k,s)dkds \\ \int c(k,s;\lambda,z)\lambda(k,s)dkds + \int g(k,s;\lambda,z)\lambda(k,s)dkds &= zF(K,N) + (1-\delta)K. \end{aligned}$$

10.4.3 Computing the Equilibrium

The main difficulty of computing the RCE for this economy is due to the "curse of dimensionality". The state variable contains the distribution λ, which is a variable of, in principle, infinite dimension (it is a function!). Krusell and Smith propose a smart way to *approximate* an equilibrium using simulations. Their underlying assumption is that agents are boundedly rational in their perceptions of how λ evolves over time and to increase the sophistication of these perceptions until the errors that agents make because they are not fully rational become negligible.

Therefore, let's assume that agents do not perceive current or future prices as depending on anything more than the first I moments of λ. Denote those moments as $m \equiv (m_1, m_2, ..., m_I)$. Since current prices depend only on the total amount of capital and not on its distribution, limiting agents to a finite set of moments is restrictive only as far as future prices are concerned. In particular, to know future prices, it is necessary to know how the aggregate capital stock evolves. Since savings decisions do not aggregate, the total capital stock in the future is a nontrivial function of all the moments of the current distribution. Agents thus perceive the law of motion for m to be given by a function H_I that belongs to a class φ. Each of these functions expresses m', that is, the vector of I moments in the next period, as a function of the I current moments

$$m' = H_I(m, z, z')$$

Given the law of motion H_I, each agent's optimal savings decision can then be represented by a decision rule f_I. Given such a decision rule f_I for individuals and an initial wealth and labor shock distribution, it is possible to derive the implied aggregate behavior–a time series path of the distribution of income and wealth–by simulating the behavior of a large number of consumers. The resulting distributions, therefore, are restricted only by initial conditions, shocks, and the decision rules of agents. Moreover, they can be used to compare the simulated evolution of the specific vector of moments m to the perceived law of motion for m on which agents base their behavior. Our approximate equilibrium is a function H_I that, when taken as given by the agents, (i). yields the best fit within the class φ to the behavior of m in the simulated data and (ii). yields a fit that is close to perfect in the sense that H_I tracks the behavior of m in the simulated data almost exactly, that is, with very small errors. In a computed, approximate equilibrium, thus, agents do not take into account all the moments of the distribution, but the errors in forecasting prices that result from this omission are very small. In other words, we think thr first I moments of distribution are enough for summarizing relevant information for the economic decision. We thus reduce the infinite-dimensional problem of finding a law of motion of λ to a finite-dimensional one.

According to thsi idea, we will have following computation algorithm:

1. Select I.

2. Guess on a parameterized functional form for H_I, and guess on the parameters of this function.

3. Solve the consumer's DP problem given H_I

$$\begin{aligned} v(k, s; m, z) &= \max_{c,k'}\{u(c) + \beta E[v(k', s'; m', z') \mid (s, z, m)]\} \\ & \quad s.t. \\ c + k' &= (1 + \tilde{r} - \delta)k + ws \\ m' &= H_I(m, z, z') \end{aligned}$$

4. Draw a single long realization from the Markov process for $\{z_t\}$ for a length of a large number of time say T. For that particular realization of z, using consumers' decision rules derived from step 3 to simulate the time paths of $\{k_t, s_t\}$ for N agents (with N a large number) over T time periods.

5. Assemble these N simulations into a history of T empirical cross-section distributions $\lambda_t(k,s)$. Then use the cross section at t to compute the cross-section moments $m(t)$, thereby assembling a time series of length T of the cross-section moments $m(t)$. Next use this sample and nonlinear least squares to estimate the parameter values of transition function H_I mapping $m(t)$ into $m(t+1)$. They return to the beginning of the

6. If the estimation gives parameter values that are very close to those guessed initially and the goodness of fit is satisfactory, stop. If the parameter values have converged but the goodness of fit is not satisfactory, go back to step 1 to increase I or, as an intermediate step, try a different functional form for H_I. Follow the procedure again until the convergence.

An example serves to illustrate the workings of the algorithm. Suppose that $I = 1$, so we only look at the first moment which is the aggregate capital stock. The aggregate shock only takes two values $z \in \{z_g, z_b\}$, i.e., it is either in good state z_g or in bad state z_b. And we further assume that H_1 is log-linear

$$\begin{aligned} z &= z_g : \log K' = a_0 + a_1 \log K \\ z &= z_b : \log K' = b_0 + b_1 \log K. \end{aligned}$$

The FE for this example is

$$v(k,s;K,z) = \max_{c,k'} \{u(c) + \beta E[v(k',s';K',z') \mid (s,z,K)]\}$$

subject to

$$\begin{aligned} c + k' &= (1 + \tilde{r}(K,N,z) - \delta)k + w(K,N,z)s \qquad (10.16) \\ \log K' &= a_0 + a_1 \log K, \text{ if } z = z_g \\ \log K' &= b_0 + b_1 \log K, \text{ if } z = z_b \\ c, k' &\geq 0 \end{aligned}$$

By solving it, we obtain a (nonlinear) decision rule $k' = f_1(k,s;K,z)$. This rule is then used to simulate the economy with a sample of N agents (N is a large number). We then sum up among these N agents to compute a sequence of aggregate capital stock $\{K_t\}_{t=0}^T$ as follows

$$K_t = \frac{1}{N} \sum_{i=1}^{N} k_i'(k,s;K,z).$$

Next, use least-squares regression to estimate parameters (a_0, a_1, b_0, b_1) of the linear law of motion for K. Using these estimates, the law of motion is updated, and the procedure is repeated until a fixed point in these parameters is found. If this H_1 is satisfactorily reproduced in simulations (i.e., is the goodness-of-fit, which represents a measure of how close approximate aggregation is to exact[7]), stop. Otherwise, consider a more flexible functional form for H_1 or add another moment.

[7] Krusell and Smith (1998) discuss different goodness-of-fit measures, but the simplest one is R^2. Thus, if the R^2 is very close to 1, we say that we have approximate aggregation.

10.4.4 Numerical Results

Krusell and Smith calibrate the two-state example above and show the approximate aggregation results. They use $\beta = 0.99$ and $\delta = 0.025$, reflecting a period in the model is one quarter in the real world. CRRA coeffcient μ is set to be 1, i.e., we have log utility. Capital share $\alpha = 0.36$. They set the shock values to $z_g = 1.01$ and $z_b = 0.99$. To keep things simple, they also assume only two values for the individual productivity shock, i.e., employment status $s \in \{s_g, s_b\}$, where $s_g = 1$ means the agent is employed at good time, $s_b = 0$ means unemployed at bad time. Denote

$$\pi(z', s' \mid z, s) = \Pr(z_{t+1} = z', s_{t+1} = s' \mid z_t = s, s_t = s)$$

More concisely, we write it as $\pi_{zz'ss'}$. In our (2×2) case, the joint Markov transition matrix has 16 entries. It is easy to show that the transition parobabilities have to satisfy the following restrictions

$$\pi_{zz'00} + \pi_{zz'01} = \pi_{zz'10} + \pi_{zz'11} = \pi_{zz'}$$

The number of people who are unemployed always equal to u_g in good time and u_b in bad time. And we have

$$u_z \frac{\pi_{zz'00}}{\pi_{zz'}} + (1 - u_z) \frac{\pi_{zz'10}}{\pi_{zz'}} = u_{z'}, z = g, b$$

Krusell and Smith set the unemployment rates to $u_g = 0.04$ and $u_b = 0.1$, implying that the fluctuations in the macroeconomic aggregates have roughly the same magnitude as the fluctuations in observed postwar U.S. time series. The process for (z, s) is chosen so that the average duration of both good and bad times is eight quarters and so that the average duration of an unemployment spell is 1.5 quarters in good times and 2.5 quarters in bad times.

Given these parameters, Krusell and Smith solve the consumer's problem by computing an approximation to the value function on a grid of points in the state space. They use cubic spline and polynomial interpolation to compute the value function at points not on the grid. In their simulations they include $N = 5000$ agents and $T = 11000$ periods; they discard the first 1,000 time periods. Typically, the initial wealth distribution in the simulations is one in which all agents hold the same level of assets. They find that their results are not sensitive to changes in the initial wealth distribution.

With a log-linear functional form and only the mean of the capital stock as a state variable, They obtain the following approximate equilibrium

$$\begin{aligned} \log K' &= 0.095 + 0.962 \log K, \text{ if } z = z_g \\ \log K' &= 0.085 + 0.965 \log K, \text{ if } z = z_b \end{aligned}$$

The $R^2 = 0.999998$ in both cases, which means that the agents with this simple forecasting rule by ignoring higher moments make very small errors, for example the maximal error in forecasting the prices 25 years into the future is around

0:1%. In other words, although the inclusion of more moments in these predictions must significantly improve forecasts in a statistical sense (since strict aggregation does not obtain), these improvements are minuscule in quantitative terms.

Why is only mean, the first moment, is important? What is the intuition behind this? Krusell and Smith interpret this result in terms of an "*approximate aggregation theorem*" that follows from two properties of their parameterized model. First, consumption as a function of wealth is concave but close to linear for moderate to high wealth levels, i.e., as wealth increases, the slope of savings decision rule increases towards one, the marginal propensities to consume (hence to save) are almost identical for agents with different empolyment states and level of capital. Second, most of the saving is done by the high-wealth people. The capital held by the very poor people is negligible. These two properties mean that fluctuations in the distribution of wealth have only a small effect on the aggregate amount saved and invested. Thus, distribution effects are small.

So, although the agents in the Krusell-Smith economy are prudent (they have CRRA utility), face liquidity constraints which bind with positive probability, and lack full insurance against productivity shocks, the economy in the aggregate behaves almost like a complete-markets economy where aggregation holds perfectly. The reason is that most of the consumers in this economy can smooth consumption very effectively through self-insurance, hence their saving behavior is guided mostly by their intertemporal motive rather than their insurance motive.

Krusell and Smith also compare the distributions of wealth from their model to the U.S. data. Relative to the data, the model with a constant discount factor generates too little skewness in wealth. More precisely, too few agents hold low levels of wealth, and the concentration of wealth among the richest agents is far too small in the model. In the data, the richest 1% of the US population holds 30% of all household wealth, the top 5% hold 51% of all wealth. For the model described above the corresponding numbers are 3% and 11%, correspondingly. In the model people save to buffer their consumption against unemployment shocks, but since these shocks are infrequent and of short duration, they don't save all that much. Also the model misses on the low wealth housholds. In the US data 11% of households have negative wealth, while in the model this fraction is 0%.

As an attempt to solve this problem, Krusell and Smith find that a modest difference in discount factors among agents can bring the model's wealth distribution much closer to the data. Patient people become wealthier, impatient people eventually become poorer, skewness in wealth thus increases. They assume that agents' preferences are ex ante identical but that discount factors β are random and follow a Markov process with transition probabilities $\gamma(\beta' \mid \beta)$.

Model	**% of Wealth Held by Top** **1%**	**5%**	**10%**	**20%**	**30%**	**% Holding Zero Wealth**	**Gini**
Benchmark	3	11	19	35	46	0	0.25
Stochastic-β	24	55	73	88	92	11	0.82
Data	30	51	64	79	88	11	0.79

Table 10.6: Distribution of wealth: model vs. data

Thus the FE for the HH's problem changes to

$$\begin{aligned} v(k,s,\beta;K,z) &= \max_{c,k'}\{u(c)+\beta E[v(k',s'\beta';K',z') \mid (s,z,\beta,K)]\} \\ &= \max_{c,k'}\{u(c)+\beta\sum_{k'}\sum_{s'}\sum_{\beta'}\pi(z',s' \mid z,s)\gamma(\beta' \mid \beta)v(k',s'\beta';K)\} \end{aligned}$$

subject to the BC as in equation (10.16) above.

More precisely, Krusell and Smith assume that β can take one three possible values, $\beta \in \{0.9858, 0.9894, 0.9930\}$. Hence annual discount rates differ between 2.8% for the most patient agents and 5.6% for the most impatient agents. As transition matrix Krusell and Smith propose (remember that the model period is one quarter)

$$\gamma = \begin{pmatrix} 0.995 & 0.005 & 0 \\ 0.000625 & 0.99875 & 0.000625 \\ 0 & 0.005 & 0.995 \end{pmatrix}$$

to match (i). the invariant distribution for β's has 80% of the population at the middle β and 10% at each of the other β's, (ii). immediate transitions between the two extreme values of β occur with probability zero, and (iii). the average duration of the highest and lowest β's is 50 years.

This stochastic-β model has done a much better job in replicating the wealth distribution in the data. See Table which replicates Table 1 in Krusell and Smith (1998) for the results.

In the modified stochastic-β economy the richest 1% of the population holds 24% of all wealth, the richest 5% hold 55% percent of all wealth, pretty closer to the data. In particular, the large fraction of agents with negative wealth in the stochastic-β model matches the data, and the Gini coefficient is also quite close to that in the data.

In the stochastic-β model, poor agents are poor because they have chosen to be poor. This choice, in turn, is based purely on (exogenous) genetics: the degree of patience turns out to be crucial for the accumulation of wealth of an agent, and the equilibrium interaction between agents with different degrees of patience forces very large differences in wealth from seemingly small differences in patience.

10.5 Some Applications of Bewley-Aiyagari Model

10.5.1 Cost of Business Cycles with Indivisibilities and Liquidity Constraints

In an interesting and influential study, Lucas (1987) estimates the magnitude of the costs of business cycles to be remarkably small, only around 0.1% of total U.S. consumption. But he used the complete-market model that assumes perfect insurance of the idiosyncratic risk. Ayse Imrohoroglu (1989) examines whether the magnitude of the costs of business cycles in economies with incomplete insurance markets differs significantly frorn the cost estimates found in an environment with perfect insurance.

Model Setup

It is very similar to Aiyagari model in a partial equilibrium setting. The economy consists of many infinitely lived individuals who are different at a point in tirne only in their asset holdings and ernployment opportunities. They maximize

$$\max E_0 \sum_{t=0}^{\infty} \beta^t u(c_t)$$

where period utility function takes the CRRA form

$$u(c) = \frac{c^{1-\sigma} - 1}{1-\sigma}, \sigma > 0.$$

Agents are endowed with one *indivisible* unit of time in each period and each face an individual-specific stochastic employment opportunity that has two states, $i = e$ or $i = u$. If the employed state occurs ($i = e$), an agent produces y units of the consurnption good using the time allocation. In the unemployed state ($i = u$), the agent produces θy units of consumption good through household production, where $0 < \theta < 1$. Thus an individual's asset holdings evolve over time according to

$$a_{t+1} = \begin{cases} (1+r)(a_t - c_t + y) & \text{if } i = e \\ (1+r)(a_t - c_t + \theta y) & \text{if } i = u \end{cases}.$$

In order to assess the cost of business cycles, two economies are considered. The first economy is exposed to aggregate shocks, hence experiences business cycles, whereas in the second there is no aggregate uncertainty. The average rates of unemployment for these two economies are the same.

Economy 1: The state of the national economy n is assumed to follow a first-order Markov chain. $n = g$ for a good time and $n = b$ for a bad time. The transition matrix of n is

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

where 1 for good time and 2 for bad time.

As we describe above, variable $i = e, u$ denotes the individual-specific employment state and is assumed to follow a first-order Markov chain. The transition matrix for i is P^g in good times and P^b in bad times.

$$\begin{aligned} P^g &= \begin{bmatrix} p^g_{u,u} & p^g_{u,e} \\ p^g_{e,u} & p^g_{e,e} \end{bmatrix} \\ P^b &= \begin{bmatrix} p^b_{u,u} & p^b_{u,e} \\ p^b_{e,u} & p^b_{e,e} \end{bmatrix}, \end{aligned}$$

where for example $p^g_{e,u} = \Pr(i_{t+1} = u^g \mid i_t = e)$ is the probability that an agent will be unemployed in good times at period $t+1$ given that the agent was employed at period t.

We will have following structure on the transition probabilities of P^g and P^b to let the story make economic sense

$$\begin{aligned} p^g_{e,e} &> p^b_{e,e}, p^g_{u,u} < p^b_{u,u} \\ p^g_{u,e} &> p^b_{u,e}, p^g_{e,u} < p^b_{e,u} \end{aligned}$$

The overall employment prospects state, s, faced by each individual is a combination of the aggregate and individual states, that is, $s = \{i, n\}$. It has four possible values, s_1, s_2, s_3, s_4, which stand for employed in good times $s_1 = \{i = e, n = g)$, unemployed in good times $s_2 = \{i = u, n = g)$, employed in bad times $s_3 = \{i = e, n = b\}$, and unernployed in bad times $s_4 = \{i = u, n = b\}$, respectively. The process governing s is a first-order Markov chain with the transition matrix Π

$$\Pi = \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} & \pi_{14} \\ \pi_{21} & \pi_{22} & \pi_{23} & \pi_{24} \\ \pi_{31} & \pi_{32} & \pi_{33} & \pi_{34} \\ \pi_{41} & \pi_{42} & \pi_{43} & \pi_{44} \end{bmatrix}$$

where $\pi_{ij} = \Pr(s_{t+1} = s_j \mid s_t = s_i)$. The transition probabilities of this rnatrix are determined from the P, P^g, P^b matrices. For example, if $s_t = s_1$, then the probability of s_{t+1} being equal to s_2 is given by $\pi_{12} = p_{11} p^g_{e,u}$.

Economy 2: this economy, there is no aggregate uncertainty. The state of employment, i, is assumed to follow a Markov chain with two possible states: u and e. The transition matrix for i is

$$\chi = \begin{bmatrix} \chi_{ee} & \chi_{eu} \\ \chi_{ue} & \chi_{uu} \end{bmatrix}$$

Ayse considers two types of borrowing constraint in the economy. The first one is no-borrowing constraint, i.e., borrowing is not allowed

$$a_{t+1} \geq 0$$

Since event-contingent insurance is not permitted, individuals can insure only through holdings of liquid assets. In equilibrium they will accumulate assets during the periods when they work to provide for consumption during the periods when they are unemployed.

Another one is an environment that there is a perfectly competitive intermediation sector. Agents can now borrow as well as save. However, agents are only allowed to borrow from the intermediary at a borrowing rate that exceeds the lending rate. The difference between these rates reflects the costs of intermediation.

Finally for comparison, we describe the environment with perfect insurance of the idiosyncratic shock. An event-contingent insurance scheme is assumed to exist that eliminates all but aggregate uncertainty. At a given point in time a certain fraction of the population is employed, producing y units of the consumption good. On the other hand, those who are unemployed produce θy units of the consumption good. However, regardless of the individual-specific employment state, each agent receives the same amount of per capita income. Of course, at each period, the amount of income produced in the economy depends on the aggregate shock and the fraction of the people employed.

Let κ be the fraction of people employed in the current period and $\bar{y}^n$ be the per capita income in the current priod, where $n = g, b$. We have

$$\bar{y}^n = \kappa y + (1 - \kappa)\theta y.$$

The fraction of the people employed next period κ' is

$$\kappa_n' = \kappa p_{e,e}^n + (1 - \kappa) p_{u,e}^n, n = g, b.$$

Calibration

For the economies to be fully specified, it is necessary to choose the invariant transition probabilities for P, P^g, P^b and specific parameter values for β, σ, r and θ. Ayse follows Kydland and Prescott (1982) and chooses these so that certain key statistics for the model economies match those for the U.S. economy. Since the assets in the model economy are liquid assets, the net return on assets r is set to be zero. For the model with intermediation technology, we assume rate on borrowing is 8%. The income fraction of an unemlployed individual, θ, is assumed to be $\frac{1}{4}$. The subjective time discount factor, β, is assumed to be 0.995, which implies an annual subjective time discount rate of 4 percent. In order to compare the results here with the findings in Lucas (1987), the costs of business cycles are computed for CRRA coefficient $\sigma = 6.2$. The value of the risk aversion parameter utilized for the economies with an intermediation technology is 1.5. The time period in the model corresponds to 6 weeks in the real world.

The time-invariant transition probabilities for P, P^g, P^b, for all the environments studied, are selected so that the variation in per capita employment between good and bad times is 8 percent. This implies a variation in unernployrnent of 8 percent, namely from 4 percent in good times ($\kappa^g = 0.96$) to 12 percent in bad times ($\kappa^b = 0.88$). We end up with

$$P = \begin{bmatrix} 0.9375 & 0.0625 \\ 0.0625 & 0.9375 \end{bmatrix}$$
$$P^g = \begin{bmatrix} 0.3976 & 0.6024 \\ 0.0251 & 0.9749 \end{bmatrix}$$
$$P^b = \begin{bmatrix} 0.5708 & 0.4292 \\ 0.0585 & 0.9415 \end{bmatrix}$$

This implies

$$\Pi = \begin{bmatrix} 0.9140 & 0.0235 & 0.0588 & 0.0037 \\ 0.5625 & 0.3750 & 0.0269 & 0.0356 \\ 0.0608 & 0.0016 & 0.8813 & 0.0563 \\ 0.0375 & 0.0250 & 0.4031 & 0.5344 \end{bmatrix}$$

Next we will select the transition probabilities for the state of employment in an economy without any business cycles. Requiring the average rate of unemployment and the average duration of umemployment to be the same across the economies with and without business cycles determines these probabilities. The average duration of unemployment the umployment rate in good time are ($D^g = 10$ weeks ($=\frac{10}{6}$ model period), $\kappa^g = 0.96$). The average duration of unemployment the umployment rate in bad time are ($D^b = 14$ weeks ($=\frac{14}{6}$ model period), $\kappa^g = 0.88$). Thus the simple average of duration and rate of umemployment is ($\bar{D} = 12$ weeks (=2 model period), $\bar{\kappa} = 0.92$). Thus χ_{uu} is determined by

$$\bar{D} = \frac{1}{1 - \chi_{uu}} = 2 \Rightarrow \chi_{uu} = 0.5$$

And χ_{ee} is determined by

$$\bar{\kappa} = \bar{\kappa}\chi_{ee} + (1 - \bar{\kappa})\chi_{ue}$$

But we have

$$\chi_{ue} = 1 - \chi_{uu} = 0.5$$

which implies $\chi_{ee} = 0.9565$. Therefore we obtain

$$\chi = \begin{bmatrix} 0.9565 & 0.0435 \\ 0.5000 & 0.5000 \end{bmatrix}.$$

This concludes our calibration process.

Computation of the Equilibrium

The Bellman equation that individuals face is

$$\begin{aligned} v(a, s) &= \max_{a'}\{u(c) + \beta \sum_{s'} \pi_{ss'} v(a', s')\} \\ & s.t. \\ a_{t+1} &= \begin{cases} (1+r)(a_t - c_t + y) & \text{if } i = e \\ (1+r)(a_t - c_t + \theta y) & \text{if } i = u \end{cases} \end{aligned}$$

Or more concisely by substituting out BC

$$v(a, s) = \max_{a'}\{u(a, a', s) + \beta \sum_{s'} \pi_{ss'} v(a', s')\}.$$

Notice that the DP problem for the economy with no business cycles is the same as that of the economy with business cycles except for the fact that the state $s = (i, n)$ in the latter economy and $s = i$ for the former.

Ayse uses the standard discrete-state DP method to numerically solve the BE above (see section 4.2.1). For the non-borrowing constraint case, $a \in [0, 8]$ with a grid of 301 points (increment=0.027). For the intermediation technplogy case, $a \in [-8, 8]$ with a grid of 601 points (increment=0.027). The overall state of employment prospects s takes one of four possible values, s_1, s_2, s_3, s_4. The total number of possible states for the individual is then $4 \times 301 = 1204$, the number of employment states times the number of asset states. At each point in time the number of possible outcomes is finite, never exceeding 301. Consequently the problem is a finite state, discounted dynamic program.

Ayse then uses value function iteration method to obtain the value function and the optimal decision rule for asset holdings. By using the Markov properties of s and the obtained decision rule, she compute the invariant distribution $\lambda^*(a, s)$ by solving the following law of motion of distribution

$$\lambda_{t+1}(a', s') = \sum_{a':a'=g(a,s)} \sum_{s'} \pi_{ss'} \lambda_t(a, s).$$

The equilibrium process for the economies with perfect insurance is given by the equality between per capita consumption and per capita income each period because storage of the aggregate output is not allowed in this study.

In order to compute the steady-state average utility in the economy with business cycles, it is necessary to generate time series by using Monte Carlo methods. Average income fluctuates each period and is a function of the fraction of the people employed. For this economy, individual time series that consist of 500,000 periods are generated and average utility is found. For the economy with no business cycles, average income each period is constant. Thus average utility is computed simply by setting average consumption equal to average income.

Findings

Table 10.7 reports the increase in average consumption that is necessary to compensate the individual for the loss of utility due to business cycles in the environment with only a storage technology (i.e., non-borrowing constraint).

Eliminating perfect insurance significantly increases the cost of business cycles. Increasing the risk aversion coefficient from 1.5 to 6.2 increased the average asset holdings by a factor of 2.6. This confirms the fact that as individuals get more risk averse, their precautionary asset holdings increase. The time average of asset holdings in the economy with business cycles is 2.20 for $\sigma = 1.5$ and

CRRA coefficient	perfect insurance	Storage tech.	Intermediation tech.
$\sigma = 1.5$	0.08%	0.30%	0.05%
$\sigma = 6.2$	0.30%	1.50%	n.a.

Table 10.7: Cost of Business Cycles' As a Percentage of Consumption

5.67 for $\sigma = 6.2$. The corresponding averages for the economy without business cycles are 2.26 and 5.76, respectively.

In the economies in which borrowing is allowed and $\sigma = 1.5$, only a 0.05 percent increase in average consumption is needed to compensate the individual for the loss of utility due to business cycles. That is, the cost estimate is reduced by a factor of six when borrowing is permitted, even though the borrowing rate exceeds the lending rate by 8 percent. This finding seems to suggest that the ability to store along with an intermediation technology significantly reduces the magnitude of the cost of fluctuations, i.e., allowing self-insurance through borrowing to smooth consumption buffers the effect from the business cycles successfully.

10.5.2 Wealth Distribution in Life-Cycle Economies

Huggett (1996) compares the age wealth distribution produced in Diamond-type life-cycle economies to the corresponding distribution in the US economy. By adding in several new features into the basic life-cycle framework, such as (1) earnings, health, and longevity uncertainty; (2) household structure; (3) institutional features such as social security, income taxation, and social insurance, and (4) market features such as borrowing constraints and the absence of some insurance markets, Huggett shows that the calibrated model economies can replicate measures of both aggregate wealth and transfer wealth in the US. Furthermore, the model economies produce the US wealth Gini and a significant fraction of the wealth inequality within age groups.

Model Setup

Huggett (1996) considers an overlapping generations economy. Each period a continuum of agents are born. Agents live a maximum of N periods and face a probability s_t of surviving up to age t conditional on surviving up to age $t-1$. The population grows at rate n. These demographic patterns are stable so that age t agents make up a constant fraction μ_t, of the population at any point in time. The law of motion for age structure μ_t defined by

$$\mu_{t+1} N_{t+1} = s_{t+1} \mu_t N_t$$

which implies

$$\mu_{t+1} = s_{t+1} \mu_t \frac{N_t}{N_{t+1}} = \frac{s_t}{1+n} \mu_t.$$

All age-1 agents have identical preferences for consumption:

$$E\left[\sum_{t=1}^{N}\beta^t\left(\prod_{j=1}^{t}s_j\right)u(c_t)\right]$$

Again, the period utility function is CRRA

$$u(c)=\frac{c^{1-\sigma}-1}{1-\sigma},\sigma>0.$$

An agent's labor endowment is given by a function $e(z,t)$ that depends on the agent's age t and on an idiosyncratic labor productivity shock z. The shock z takes on a finite number of possible values in the set Z and follows a Markov process. Labor productivity shocks are independent across agents. Law of large number then implies that there is no uncertainty over the aggregate labor endowment even though there is uncertainty at the individual level.

The production function is Cobb-Douglas

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

The individual's BC is

$$\begin{aligned} c+a' &\leq a(1+r(1-\tau))+(1-\theta-\tau)e(z,t)w+T+b_t \\ a' &\geq -\phi, c\geq 0 \end{aligned} \tag{10.17}$$

where τ is capital and labor income tax rate, θ is social security tax rate. T is a lump-sum transfer. b_t is social security benefit defined by

$$b_t=\begin{cases} 0 & \text{if } t<t_R \\ b & \text{otherwise}\end{cases}$$

t_R is the retirement age. Since the benefit level is the same for all agents, there is no linkage between an individual agent's earnings and future social security benefits. This assumption can be viewed as a rough first approximation to the highly redistributive nature of social security system. It also eases the computational burden significantly as a variable capturing an agent's earnings history need not be included as part of an individual agent's state.

Now for an age-t agent, the state variables are (a,z,t). Let $x=(a,z)$, the FE for an age-t agent is

$$v(x,t)=\max_{c,a'}\{u(c)+\beta s_{t+1}E[v(a',z',t+1)\mid x]\} \tag{10.18}$$

subject to BC (10.17). The solution to this FE is optimal decision rule for consumption $c(x,t)$ and asset holdings $a(x,t)$.

Stationary Equilibrium

At a point in time agents are heterogeneous in their age t and in their individual state x. Denote $X = [-\phi, +\infty)$ the state space and $B(X)$ the Borel σ-algebra on X. For each $B \in B(X)$, a probability measure $\lambda_t(B)$ defined on subsets of the state space describes the distribution of individual states across age t agents. It is the fraction of age-t agents whose individual states lie in B as a proportion of all age-t agents. Since the share of all age-t agents in the whole population is μ_t, These agents then make up a fraction $\mu_t \lambda_t(B)$ of all agents in the economy.

The distribution of individual states across age-1 agents is determined by the exogenous initial distribution of labor productivity since all agents start out with no assets. The distribution of individual states across agents age $t = 2, 3.....N$ is then given recursively as follows:

$$\lambda_t(B) = \int_X P(x, t-1, B) d\lambda_{t-1}, \ \forall B \in B(X) \tag{10.19}$$

The function $P(x, t, B)$ is a transition function which gives the probability that an age-t agent transits to the set B next period given that the agent's current state is x. The transition function is determined by the optimal decision rule on asset holdings $a(x, t)$ and by the exogenous transition probabilities on the labor productivity shock z, i.e., $P(x, t, B) = \Pr(\{z' \in Z : (a(x, t), z') \in B\} \mid z)$.

Now we are ready to define the stationary equilibrium.

Definition 92 *A Stationary equilibrium is* $(v(x,t), c(x,t), a(x,t), r, w, K, L, T, G, \tau, \theta, b)$ *and distributions* $\{\lambda_t\}_{t=1}^N$ *such that:*

(i). Given prices and government policies $(r, w, T, G, \tau, \theta, b)$*,* $v(x,t), c(x,t), a(x,t)$ *are the solutions to age-t agents' problem (10.18),* $\forall t$*.*

(ii). Competitive input markets:

$$w = \frac{\partial F(K, L)}{\partial L} \tag{10.20}$$

$$r = \frac{\partial F(K, L)}{\partial K} - \delta. \tag{10.21}$$

(iii). Markets clear:

$$\sum_{t=1}^N \mu_t \int_X (c(x,t) + a(x,t)) d\lambda_t + G = F(K, L) + (1-\delta)K \tag{10.22}$$

$$\sum_{t=1}^N \mu_t \int_X a(x,t)) d\lambda_t = (1+n)K \tag{10.23}$$

$$\sum_{t=1}^N \mu_t \int_X e(z,t)) d\lambda_t = L. \tag{10.24}$$

(iv). Distributions are consistent with individual behavior:

$$\lambda_t(B) = \int_X P(x, t-1, B) d\lambda_{t-1}, \ \forall B \in B(X), \forall t$$

(v). Government budget constraint is balanced per period:

$$G = \tau(rK + wL).$$

(vi). Social security is self-financed:

$$\theta wL = b(\sum_{t=t_R}^{N} \mu_t).$$

(vii). Transfers equal accidental bequests:

$$T = \left[\sum_{t=1}^{N} \mu_t(1 - s_{t+1}) \int_X a(x,t)(1 + r(1-\tau))d\lambda_t\right] / (1+n). \tag{10.25}$$

Calibration

The model period is one year. The preference parameter $\beta = 1.011$ is taken from Hurd's (1989) estimate in economies where mortality risk is accounted for separately. CRRA coefficient $\sigma = 1.5$ follows the estimates of the microeconomic studies reviewed by Auerbach and Kotlikoff (1987) and Prescott (1986).

The technology parameters (A, α, δ) are set as follows. The TFP level A is normalized so that the wage equals 1.0 when the capital output ratio equals 3.0 and the labor input per capita is normalized at 1.0. Capital's share of output $\alpha = 0.36$ is set following the discussion in Prescott (1986). The depreciation rate δ is set to match the US depreciation-output ratio following the estimate of Stokey and Rebelo (1995).

The demographic parameters (N, t_R, s_t, n) are set using a model period of one year. Thus, agents are born at a real-life age of 20 (model period 1) and live up to a maximum real-life age of 98 (model period $N = 79$). Agents retire at a real-life age of 65 (model period $t_R = 46$). The survival probabilities $\{s_t\}$ are set according to the actuarial estimates in Jordan (1975). The population growth rate n is set to equal the average growth rate in the US from 1950-92 1.2%.

Tax rates (τ, θ) are set as follows. The income tax τ is set to match the average share of government consumption in output $(\frac{G}{Y})$ 0.195 for 1959-93 period. Notice that

$$G = \tau(rK + wL) = \tau(F(K,L) - \delta K)$$

$\Rightarrow$

$$\begin{aligned} \frac{G}{Y} &= \tau(\frac{F(K,L) - \delta K}{Y}) \\ &= \tau(1 - \delta\frac{K}{Y}) \end{aligned}$$

$\frac{G}{Y} = 0.195, \frac{K}{Y} = 3, \delta = 0.06 \Rightarrow \tau = 23.78\%$. The social security tax rate θ equals the average for the 1980's of the contribution to social security programs as a fraction of labor income.

β	σ	A	α	δ	N	t_R	s_t	n	τ	θ	ϕ
1.011	1.5	0.896	0.36	0.06	79	46		0.012	0.2378	0.10	0,$-w$

Table 10.8: Model Parameters of Huggett (1996)

The credit limit (ϕ) is set at 0 and for comparison purposes at $-w$. A credit limit of 0 means that agents are not allowed to hold net debt. A credit limit of $-w$ means that agents are allowed to borrow up to one year's average earnings in the economy.

Table 10.8 summarizes the model parameters.

Huggett also investigates a labor endowment process with regression towards the mean in log labor endowment. This labor endowment or earnings process has been estimated in a number of studies. Denote y_t the log labor endowment of an age-t agent and $\bar{y}_t$ the mean log endowment of age-t agents. Also let ε_t be a labor endowment shock that is distributed $N(0, \sigma_\varepsilon^2)$ and independently over time.

$$y_t - \bar{y}_t = \gamma(y_{t-1} - \bar{y}_{t-1}) + \varepsilon_t$$

Based on a series of micro studies, Huggett sets $\sigma_\varepsilon^2 = 0.045$ and $\gamma = 0.96$. He also chooses $\bar{y}_t$ to match the US age-earnings profile. Define a Markov process z, where $z_t \equiv y_t - \bar{y}_t$

$$z_t = \gamma z_{t-1} + \varepsilon_t$$

Tauchen's method then is used to approximate the AR(1) process above with 18 possible states of z. The transition probabilities between states are calculated by integrating the area under the normal distribution conditional on the current value of the state. In summary, the labor endowment process is given by

$$e(z,t) = e^{(z_t + \bar{y}_t)}$$

where z_t is a finite Markov chain.

Computing the Equilibrium

The algorithm for computing equilibria is as follows:

1. Choose initial capital stock K_1 and accidental bequest transfer T_1. Obtain L_1 by using (10.24).

2. Set w and r according to equilibrium conditions (10.20)-(10.21).

3. Given w and r, find decision rule $a(x,t)$ by solving the individual's DP problem (10.18) backwardly.

4. Calculate the wealth distribution $\lambda_t(a,t)$ according to (10.19), then calculate new capital K' and transfer T' according to (10.23) and (10.25) respectively.

5. If K' and T' are approximately equal to K_1 and T_1, stop. Otherwise use relaxation method to adjust K and T and go back to step 2.

				% Wealth in the top			
Credit limit (ϕ)	**Earnings shock (σ_ε^2)**	K/Y	**Wealth Gini**	1%	5%	20%	wealth≤ 0 (%)
US Data		3.0	0.72	28	49	75	5.8-15.0
Model with uncertain lifetimes and $\sigma = 1.5$							
0.0	0.00	3.1	0.46	2.5	11.7	42.8	11.0
$-w$	0.00	3.0	0.49	2.6	12.1	44.3	12.0
0.0	0.045	3.4	0.69	10.9	32.9	70.0	12.0
$-w$	0.045	3.2	0.76	11.8	35.6	75.5	24.0
Model with uncertain lifetimes and $\sigma = 3.0$							
0.0	0.00	2.5	0.50	2.6	12.6	45.8	21.0
$-w$	0.00	2.3	0.61	3.1	14.3	51.1	29.0
0.0	0.045	3.0	0.72	12.1	35.7	71.7	19.0
$-w$	0.045	2.8	0.84	13.8	40.4	80.2	40.0

Table 10.9: Wealth distribution: model vs. data

Results

Huggett (1996) reports the model simulated welath distribution measures and compare with the data. The notion of wealth used in the model economies is net asset holdings, a.

Table 10.9 compares the US economy and the model economies along a number of dimensions. Focus first on the capital-output ratio. There is considerable variation in this ratio. Part of the variation occurs as a result of changes in the credit limit. When the credit limit is relaxed by one year's average earnings, the capital-output ratio declines as agents can borrow to buffer shock and smooth consumption. The effect of lowering the credit limit is to decrease the capital-output ratio by 3 to 13 percent.

A large part of the variation in the capital-output ratio is also due to earnings uncertainty. Earnings uncertainty adds a precautionary savings motive to the model. The partial equilibrium literature on precautionary savings shows that savings increase with added uncertainty when marginal utility is a convex function. In steady state the gross savings rate ($\frac{S}{Y}$) in the model economy is related to the capital-output ratio ($\frac{K}{Y}$) as follows

$$\frac{S}{Y} = \frac{I}{Y} = (n + \delta)\frac{K}{Y}$$

Thus, the general equilibrium effect of adding earnings uncertainty is to increase the gross savings rate by 1-2 percent of output when risk aversion is $\sigma = 1.5$. Savings increase by 4-5 percent of output when risk aversion is $\sigma = 3.0$. Clearly, the savings effect is much stronger when agents are more risk-averse. The magnitude of the general equilibrium effects of earnings uncertainty calculated here are within the range calculated by Aiyagari (1994) using the infinitely-lived agent abstraction.

The table also shows that the model economies are capable of generating the US wealth Gini coefficient. However, it is clear that the model economies generate the US Gini by generating a high fraction of zero and negative wealth-holders and not by concentrating enough wealth in the extreme upper tail of the wealth distribution (especially top 1%). So the problem raised in Aiyagari (1994) and Krusell-Smith (1998) remains.

The fraction of agents with zero or negative wealth is at or above US levels even when agents are not allowed to go into debt. This does not seem to be a property of the infinitely-lived agent model (for example, see Aiyagari (1994)). In the infinitely-lived agent model relatively few agents tend to be exactly at the corner of the borrowing constraint. Thus, life-cycle considerations seem to be important for generating low wealth levels.

10.5.3 A Life-Cycle Analysis of Social Security

As an application of Aiyagari type incomplete market model, Imrohoroglu, Imrohoroglu, and Joines (1995) develop an general equilibrium OLG model to examine the optimal social security replacement rate and welfare benefits associated with it.

Model Setup

The model is similar to Huggett (1996). The economy is populated by a large number of *ex-ante* identical individuals who maximize the following expected life-time utility:

$$E\left[\sum_{j=1}^{J}\beta^{j-1}\left(\prod_{k=1}^{j}\psi_k\right)u(c_j)\right]$$

The period utility function is CRRA

$$u(c_j)=\frac{c^{1-\gamma}-1}{1-\gamma},\gamma>0.$$

The share of age-j individuals in the population is given by the fraction $\mu_j > 0, j = 1, 2, ..., J, \sum_{j=1}^{J}\mu_j = 1$, where J is the maximum possible lifetime.

Each period individuals are subject to a stochastic employment opportunity $s \in S = \{e, u\}$. Employment state s follows a first-order Markov process with a 2×2 transition matrix $\Pi(s, s') = [\pi_{ij}], i, j = e, u$ and $\pi_{ij} = \Pr(s_{t+1} = j \mid s_t = i)$.

An age-j individual's BC is

$$\begin{aligned} c_j + a_{j+1} &= (1+r)a_j + q_j + T \\ a_{j+1} &\geq 0, c_j \geq 0 \end{aligned} \tag{10.26}$$

where q_j is the disposable income of this age-j individual, a_{j+1} is the end-of-period asset holdings of an age-j individual, and T is a lump-sum transfer from the government. Agents in this economy is not allowed to borrow and have no access to private insurance market.

Before the mandatory retirement age of j_R, an age-j individual who is given the opportunity to work receives

$$w_j^e = w\varepsilon_j n_j$$

where ε_j is the efficient unit (labor productivity) at age j, n_j is the number of hours worked. The labor indivisibility is introduced here. So if $s = e$, then $n_j = \bar{h}$; if $s = u$, then $n_j = 0$. If an individual is unemployed, he receives umemployment insurance benefits in the amount $w_j^u = \zeta w\bar{h}$[8], where ζ is the replacement ratio.

After the mandatory retirement age of j_R, the disposable income of a retiree is equal to his social security bebefits b. These benefits are calculated to be a fraction, θ, of some base income which we take as the average lifetime employed income. That is

$$b = \begin{cases} 0 & \text{if } j < j_R \\ \theta \frac{\sum_{i=1}^{j_R-1} w_i^e}{j_R-1} & \text{otherwise} \end{cases}$$

Hence we have

$$q_j = \begin{cases} (1-\tau_s-\tau_u)w\varepsilon_j\bar{h} & \text{if } j < j_R \text{ and } s = e \\ \zeta w\bar{h} & \text{if } j < j_R \text{ and } s = u \\ b & \text{otherwise} \end{cases} .$$

where τ_s is the social security tax rate and τ_u is the tax rate for unemployment insurance.

The government in this economy administrate the unemployment insurance and the social security systems. The policy instruments that government has is $(\theta, \tau_s; \zeta, \tau_u)$.

The production function is Cobb-Douglas

$$Y = F(K, N) = AK^{1-\alpha}N^{\alpha}.$$

Firm's FOCs

$$r = (1-\alpha)A(\frac{K}{N})^{-\alpha} - \delta \quad (10.27)$$

$$w = \alpha A(\frac{K}{N})^{1-\alpha}. \quad (10.28)$$

Now for an age-j individual, the state variables are (a, s). The FE for an age-j individual is

$$v_j(a, s) = \max_{c,a'}\{u(c) + \beta\psi_{j+1}E[v_{j+1}(a', s') \mid s]\}, \forall j \quad (10.29)$$

subject to BC (10.26). The solution to this FE is optimal decision rule for consumption $c_j(a, s)$ and asset holdings $a_j(a, s)$. This problem is a standard finite-state, finite-horizon discounted DP problem.

[8] More precisely, $w_j^u = \xi w\bar{\varepsilon}\bar{h}$, where $\bar{\varepsilon}$ is the average efficient unit over lifetime, i.e., the unemployment insurance is a fractrion of average lifetime labor income. Since we normalize the average to be one, we end up with $w_j^u = \xi w\bar{h}$.

Stationary Equilibrium

Now we are ready to define the stationary equilibrium for this economy. Let $D = \{d_1, d_2, ..., d_m\}$ denote the discrete drid of points on which asset holdings will be required to fall. Recall that employment status $s \in S = \{e, u\}$.

Definition 93 *A Stationary equilibrium for a given set of policy arrangements $(\theta, \tau_s; \zeta, \tau_u)$ is a collection of value functions $v_j(a, s)$, individual decision rules $c_j : D \times S \to R_+, a_j : D \times S \to D$, age-dependent stationary measures of agent types $\lambda_j(a, s)$ for each age $j = 1, 2, ..., J$, prices $\{w, r\}$, and a lump-sum transfer T such that:*

(i). Given prices, government policies, and the lump-sum transfer T, $v_j(a, s)$, $c_j(a, s)$, $a_j(a, s)$ are the solutions to age-j agent's problem (10.29), $\forall j$.

(ii). Prices $\{w, r\}$ are the solutions to the firm's FOCs (10.27)-(10.28).

(iii). Markets clear:

$$\begin{aligned}
\sum_j \sum_a \sum_s \mu_j \lambda_j(a, s)[c_j(a, s) + a_j(a, s)] &= F(K, N) + (1 - \delta) \sum_j \sum_a \sum_s \mu_j \lambda_j(a, s) a_{j-1}(a, s) \\
K &= \sum_j \sum_a \sum_s \mu_j \lambda_j(a, s) a_{j-1}(a, s) \\
N &= \sum_{j=1}^{j_R - 1} \sum_a \mu_j \lambda_j(a, s = e) \varepsilon_j \bar{h}
\end{aligned}$$

(iv). Distributions $\lambda_j(a, s)$ satisfies $\forall j$

$$\lambda_{j+1}(a', s') = \sum_s \sum_{a: a' = a_j(a,s)} \Pi(s, s') \lambda_j(a, s)$$

where the initial measure of agents at birth, λ_0 is given.

(v). Social security is self-financing, i.e., social security is a pay-as-you-go system:

$$\tau_s = \frac{\sum_{j=j_R}^{J} \sum_a \sum_s \mu_j \lambda_j(a, s) b}{\sum_{j=1}^{j_R - 1} \sum_a \mu_j \lambda_j(a, s = e) w \varepsilon_j \bar{h}} = \frac{b \sum_{j=j_R}^{J} \mu_j}{wN}.$$

(vi). Unemployment insurance benefits program is self-financing:

$$\tau_u = \frac{\sum_{j=1}^{j_R - 1} \sum_a \mu_j \lambda_j(a, s = u) \zeta w \bar{h}}{\sum_{j=1}^{j_R - 1} \sum_a \mu_j \lambda_j(a, s = e) w \varepsilon_j \bar{h}}.$$

(vii). Transfers equal accidental bequests:

$$T = \sum_j \sum_a \sum_s \mu_j \lambda_j(a, s)(1 - \psi_{j+1}) a_j(a, s).$$

Calibration

The model period is one year. Individuals are assumed to be born at the real-time age of 21 (model age $j = 1$) and they can live a maximum of $J = 65$ years, to the real-time age of 85. The sequence of conditional survival probabilities $\{\psi_j\}_{j=1}^J$ is taken from Faber (1982). The age share μ_j is calculated from the following equation

$$\mu_{j+1} = \frac{\psi_{j+1}}{1+\rho}\mu_j$$

where ρ is the growth rate of population which is average 1.2% per year over the last fifty years. Since we have

$$\sum_{j=1}^{J} \mu_j = 1$$

This implies

$$\sum_{j=1}^{J} \frac{\prod_{s=1}^{j} \psi_j}{(1+\rho)^{j-1}} \mu_1 = 1$$

In turn it implies[9]

$$\mu_1 = 1/\sum_{j=1}^{J} \frac{\prod_{s=1}^{j} \psi_j}{(1+\rho)^{j-1}}.$$

The mandatory retirement age is taken to be $j_R = 45$, which corresponds to the real-time age 65. The efficiency index $\{\varepsilon_j\}$ is intended to provide a realistic cross-sectional age distribution of earnings at a point of time. The index is taken from Hansen (1993), interpolated to in-between ages, and normalized to average one between $j = 1$ and $j = j_R - 1$; and we assume $\varepsilon_j = 0, \forall j \geq j_R$. Figure 10.4 shows the age-efficiency profile $\{\varepsilon_j\}$.

Raw hours of work, $\bar{h}$, is taken to be 0.45, which assumes that individuals devote 45 hours a week (out of a possible 98 hours). Given an employment rate of 94%, the aggregate labor input is computed as

$$N = 0.94\bar{h} \sum_{j=1}^{j_R-1} \mu_j \varepsilon_j.$$

Following Prescott (1986), the labor share α is taken to be 0.64. The TFP parameter is fixed at 1.3193 so that output is normalized at one for a capital-output ratio of 3 given an aggregate labor input of 0.3496. The depreciation rate $\delta = 0.08$.

[9] Notice that $\psi_1 = 1$.

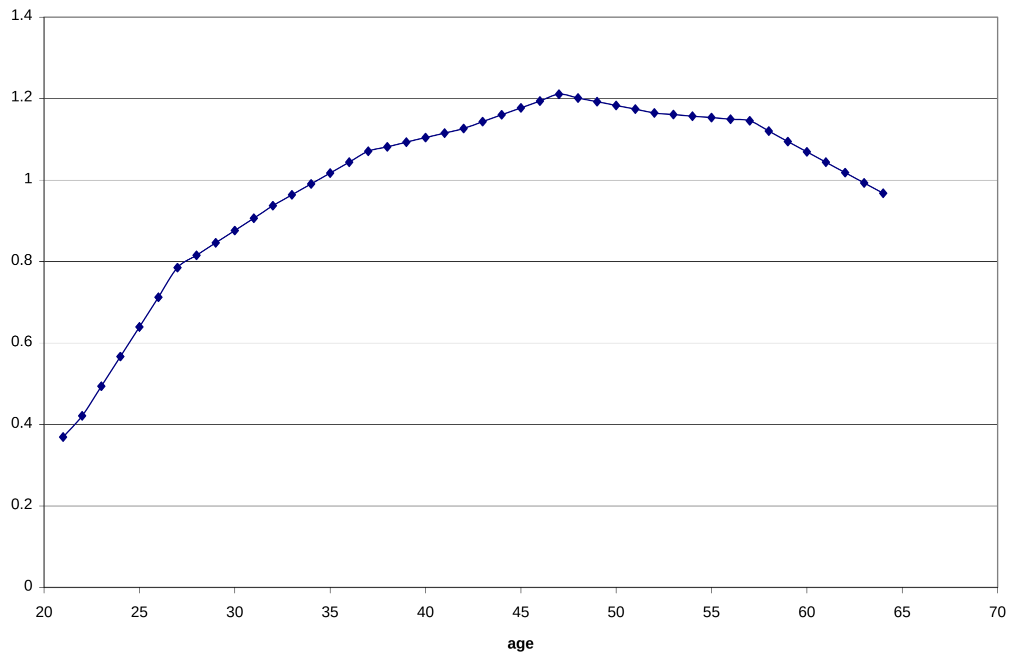


Figure 10.4: Age-efficiency profile

The CRRA coefficient γ is taken to be 2. We take $\beta = 1.011$ as our benchmark value for discount factor.[10]

The transition probabilities are chosen to make the probability of employment equal to 0.94, independent of the availability of the opportunity in the previous period. The transition matrix is given by

$$\Pi = \begin{bmatrix} 0.94 & 0.06 \\ 0.94 & 0.06 \end{bmatrix}$$

The average duration of umemployment is therefore $1/(1-0.94) = 1.0638$ model periods (i.e., years).

In most of the simulations, the discrete set of asset holdings $D = \{d_1, d_2, ..., d_m\}$ is chosen so that $d_1 = 0$ (non-borrowing limit), $d_m = 15$, and number of grids $m = 601$. The upper bound $d_m = 15$ is about 15 times the annual income. When necessary, m and d_m are increased so that d_m is never binding in the simulations.

The umemployment insurance replacement ratio ζ is taken to be 0.40. The social security replacement rate θ, is our target here. We search over values of $\theta \in (0, 1)$ in order to find the optimal benefit level.

Computing the Equilibrium

Computing the Decision Rules The Bellman equation for the individual's optimization problem is

$$v_j(a_j, s_j) = \max_{c_j, a_{j+1}} \{u(c_j) + \beta\psi_{j+1} E[v_{j+1}(a_{j+1}, s_{j+1}) \mid s_j]\}, \forall j$$

subject to

$$\begin{aligned} c_j + a_{j+1} &= (1+r)a_j + q_j + T \\ a_{j+1} &\geq 0, c_j \geq 0 \end{aligned}$$

where

$$q_j = \begin{cases} (1-\tau_s-\tau_u)w\varepsilon_j\bar{h} & \text{if } j < j_R \text{ and } s = e \\ \zeta w\bar{h} & \text{if } j < j_R \text{ and } s = u \\ b & \text{otherwise} \end{cases}.$$

The optimal decision rules and value function for each cohort can be found by a single recursion working backwards from the last period of life.

Since death is certain beyond age J, we know that $v_{J+1}(\cdot) = 0$. Hence the solution the the BE at age J is trivial

$$\begin{aligned} v_J(a_J, s_J) &= \max_{c_J}\{u(c_J)\} \\ & \quad s.t. \\ c_J &= (1+r)a_J + b + T \\ c_J &\geq 0 \end{aligned}$$

[10] See Imrohoroglu, Imrohoroglu, and Joines (1995) for detailed explanation for why β can be larger than one here.

The resulting value function $v_J(\cdot)$ is an $m \times 1$ vector whose entries correspond to the value of the utility function at $(1+r)a_J + b + T$ with a_J taking on the m possible values $d_1, d_2, ..., d_m$. And the optimal decision rule for asset holdings is also simple, it is an $m \times 1$ vector of zeros because corresponding to the m possible values of a_J, since there is no bequest motive and the death is certain, $a_{J+1} = 0$.

The value function v_J is then passed on to the next step where age-$(j-1)$ decision rule and value function are calculated.

$$\begin{aligned} v_{J-1}(a_{J-1}, s_{J-1}) &= \max_{c_{J-1}, a_{J-1}} \{u(c_{J-1}) + \beta \psi_J V_J(a_J, s_J)\} \\ & \quad s.t. \\ c_{J-1} + a_J &= (1+r)a_{J-1} + b + T \\ c_{J-1} &\geq 0, a_J \geq 0 \end{aligned}$$

Now the decision rule is found as follows. For $a_{J-1} = d_1$, we look at all possible values of a_J and evaluate the RHS of the FE above at each point on the grid D. We will pick the a_{J-1} that maximizes the value of RHS. Then this value is reported as the first element of the $m \times 1$ vector of decision rule $a_J(a_{J-1}, s_{J-1})$. By repeating this procedure for all possible initial asset levels $a_{J-1} \in D$ the entire vector a_J is filled. Simultaneously, the age-$(j-1)$ value function v_{J-1} is founded as an $m \times 1$ vector with entries corresponding to the RHS of above objective function evaluated at the decision rule a_J.

Working backwards, now we come to age $j_R - 1$. The FE is

$$\begin{aligned} v_{j_R-1}(a_{j_R-1}, s_{j_R-1}) &= \max_{c_{j_R-1}, a_{j_R}} \{u(c_{j_R-1}) + \beta \psi_{j_R} v_{j_R}(a_{j_R}, s_{j_R})\} \\ & \quad s.t. \\ c_{j_R-1} + a_{j_R} &= (1+r)a_{j_R-1} + q_{j_R-1} + T \end{aligned}$$

Now the decision rule and value function is no longer independent of the idiosyncratic employment shock s_j. Both them are an $m \times 2$ matrix.

For age $j = 1, 2, ..., j_R - 1$, the FE becomes

$$\begin{aligned} v_j(a_j, s_j) &= \max_{c_j, a_{j+1}} \{u(c_j) + \beta \psi_{j+1} \sum_{s'} \Pi(s, s') v_{j+1}(a_{j+1}, s_{j+1})\} \\ & \quad s.t. \\ c_j + a_{j+1} &= (1+r)a_j + q_j + T \\ a_{j+1} &\geq 0, c_j \geq 0 \end{aligned}$$

Repeating this procedure until we go back to the period 1. We will have the decision rules and value functions for all ages, i.e., we will have $(j_R - 1)$ $m \times 2$ matrix and $(J - j_R + 1)$ $m \times 1$ vector decision rules and $(j_R - 1)$ $m \times 2$ matrix and $(J - j_R + 1)$ $m \times 1$ vector value functions.

Computing the Age-dependent Distributions To obtain the distribution of agents, $\lambda_j(a, s)$, we start from a given initial wealth distribution λ_1. We assume that newborns have zero assets, so λ_1 is taken to be an $m \times 2$ matrix with zeroes everywhere except the first row, which is equal to $(0.94, 0.06)$, the expected (or average) employment and unemployment rates, respectively. This implies that initially 94% of newborns are employed and 6% are unemployed, but they all hold zero assets. Then the distribution of agents at the beginning of age 2 is

$$\lambda_2(a, s) = \sum_s \sum_{a:a'=a_2(a,s)} \Pi(s, s')\lambda_1(a, s)$$

And the distributions of following ages are

$$\lambda_{j+1}(a', s') = \sum_s \sum_{a:a'=a_j(a,s)} \Pi(s, s')\lambda_j(a, s), \forall j = 3, ..., J$$

Notice that for $j \geq j_R$, λ_j is an $m \times 1$ vector since the retirees are not subject to idiosyncratic employment shock.

Computing the Sationary Equilibrium We have following algorithm to compute a stationary equilibrium for the economy.

1. Make a guess at the aggregate capital stock K_0 and lump-sum transfer of accidental bequests T_0. Compute the aggregate labor input $N = 0.94\bar{h}\sum_{j=1}^{j_R-1} \mu_j \varepsilon_j$. Given K_0 and N, use the firm's FOCs to calculate factor prices w and r.

2. Given prices, government policies, and the lump-sum transfer T_0, solve the Bellman equation for each cohort. Obtain the decision rules and the age-dependent distributions following the procedure described in the previous section.

3. Computing the new aggregate capital stock $K_1 = \sum_j \sum_a \sum_s \mu_j \lambda_j(a, s)a_j(a, s)$ and the new lump-sum transfer $\sum_j \sum_a \sum_s \mu_j \lambda_j(a, s)(1 - \psi_{j+1})a_j(a, s)$ by using distribution and decision rules obtained in step 2.

4. If $K_1 \approx K_0$ and $T_1 \approx T_0$ up to a tolerance criterion of 0.001, stop; otherwise, go back to step 1 and replace K_0 with $\frac{K_0+K_1}{2}$ and T_0 with $\frac{T_0+T_1}{2}$, and iterate until convergence is achieved.

Measures of Utility and Welfare Benefits In order to compare alternative social security arrangements, we need a measure of "average utility." Given a policy arrangement $\Omega = (\theta, \tau_s; \zeta, \tau_u)$, we calculate

$$W(\Omega) = \sum_{j=1}^{J} \sum_a \sum_s \beta^{j-1} \left(\prod_{k=1}^{j} \psi_k \right) \lambda_j(a, s)u(c_j(a, s))$$

as our measure of utility. $W(\Omega)$ is the expected discounted utility a newly born individual derives from the lifetime consumption policy function $c_j(a, s)$ under a given social security arrangement.

$\boldsymbol{\theta}$	$\boldsymbol{\tau_s}$	**w**	**r**	**average** c	**K**	**average income**	**average utility**
0.00	0.000	2.236	0.004	0.740	5.224	1.220	-97.859
0.10	0.020	2.161	0.009	0.742	4.751	1.179	-96.293
0.20	0.041	2.096	0.014	0.742	4.365	1.143	-95.476
0.30	0.061	2.038	0.019	0.741	4.060	1.114	-95.175
0.40	0.081	1.989	0.024	0.738	3.772	1.085	-95.339
0.50	0.102	1.947	0.028	0.735	3.553	1.062	-95.801
0.60	0.122	1.907	0.032	0.732	3.358	1.040	-96.358
1.00	0.203	1.781	0.046	0.716	2.773	0.971	-101.570

Table 10.10: Model variables under different socisl security replacement rates

Second, we need a measure to quantify the welfare benefits (or costs) of alternative social security arrangements. As our reference economy, we take the benchmark equilibrium under a zero social security replacement rate $\theta = 0$. Our measure of welfare benefits (or costs) is the compensation (relative to output in the reference economy) required to make an individual indifferent between the reference economy and an economy under an alternative social security arrangement. More formally, let $W_0 = W(\Omega_0)$ and $W_1 = W(\Omega_1)$ denote the utility under policy arrangement $\Omega_0 = (\theta = 0, \tau_{s0} = 0; \zeta, \tau_u)$ and $\Omega_1 = (\theta_1 > 0, \tau_{s1} > 0; \zeta, \tau_u)$, respectively. Then our measure of welfare benefits is $\kappa = \frac{L}{Q_0}$ where L is a lump-sum compensation required to make a newborn indifferent between policy arrangement Ω_0 with compensation L in each period of life, and an alternative policy arrangement Ω_1 without compensation. If $L > 0$, it is a welfare benefits of the alternative social security arrangements, otherwise if $L < 0$, it is a welfare loss. Q_0 is real GDP under arrangement Ω_0.

Findings

In Table 10.10, we present our benchmark results for different social security replacement rate θ. We find that an increase in replacement ratio θ monotonically reduces the capital stock and consequently raises the net return to capital. Social security system "crowds out" perivate savings. According to the results that we do not show here, when θ increases from 0 to 0.30, then to 0.60, the capital-output ratio $\frac{K}{Y}$ decreases from 4.28 to 3.65, then to 3.23.

In this economy, the optimal social security replacement rate is 30% since it maximizes the average utility $W(\Omega)$. Notice that when we do not have social security ($\theta = 0$), the net return to capital $r - \delta = 0.4\%$ which is less than the growth rate of population $\rho = 1.2\%$. Hence, according to section 9.3.4, the benchmark economy without social security system is *dynamically inefficient*. PAYG social security system corrects this dynamic inefficiency. It also reduces private savings. From the table, it is clearly that a replacement rate somewhere between 10% and 20% makes net return to capital $r - \delta$ exactly equals the population growth rate $\rho = 1.2\%$, hence we correct the overaccumulation of capital due to dynamic inefficiency and achieve the golden rule of capital accumulation.

θ	0.10	0.20	0.30	0.40	0.50	0.60	1.00
κ	0.0120	0.0184	0.0208	0.0195	0.0158	0.0100	-0.0268

Table 10.11: The welfare benefits of social security

However, a newly born individual would prefer a social security replacement rate 30% higher than this golden rule level. The reason that optimal replacement rate is higher than the golden rule level in this economy is exactly due to the market incompleteness. Market incompleteness (missing annuity markets) deviates the individual's optimal from the social optimal, which is the golden rule. Higher optimal replacement rate arises because social security substitutes for missing annuity markets in providing insurance against uncertain life expectations in smoothing old-age consumption.

The associated welfare benefits (or loss) of alternative social security replacement rate is summarized in Table 10.11. It shows that the optimal social security replacement rate of 30% produces a welfare benefits of 2.08% of GNP, which is th ehighest among the alternative replacement rates we consider here. This number is huge! Recall that Lucas (1987) estimates the magnitude of the welfare costs of business cycles to be remarkably small, less than 0.1% of total U.S. consumption. With incomplete market setting and reasonable parameters, Ayse Imrohoroglu finds that this number could be 0.3%. Hansen and Imrohoroglu (1992) study the role of umemployment benefits and calculate the welfare cost of eliminating unemployment benefits to be 0.67% of GNP. Even with a suboptimal 50% replacement rate which is close to the current replacement rate in the US economy, the welfare benefits are still 1.58% of GNP. Social security system, due to eliminating dynamic inefficiency and substituting missing private annuity markets, is welfare-enhancing.

Chapter 11

References

1. Abel, A. B. (1990): "Asset Prices under Habit Formation and Catching Up with Joneses," *American Economic Review*, vol. 80, 38-42.

2. Aiyagari, R. (1994): "Uninsured Risk and Aggregate Saving," *Quarterly Journal of Economics*, vol. 109, 659-684.

3. Arrow, K. J. (1962): "The Economic Implications of Learning by Doing," *Review of Economic Studies*, vol. 29, 155–173.

4. Auerbach, A. J. and L. J. Kotlikoff (1987): *Dynamic Fiscal Policy*, Cambridge University Press.

5. Balasko, Y. and K. Shell (1980): "The Overlapping-Generations Model I: The Case of Pure Exchange without Money," *Journal of Economic Theory*, vol. 23, 281–306.

6. Barro, R. (1974): "Are Government Bonds Net Wealth?," *Journal of Political Economy*, vol. 82, 1095-1117.

7. Barro, R. and X. Sala-i-Martin (1995): *Economic Growth*, McGraw-Hill.

8. Bartle, R. G. (1975): *The Elements of Real Analysis*, John Wiley & Sons Inc.

9. Bewley, T. F. (1986): "Stationary Monetary Equilibrium with a Continuum of Independently Fluctuating Consumers," In Werner Hildenbrand and Andreu Mas-Colell (eds.), *Contributions to Mathematical Economics in Honor of Gerard Debreu*. Amsterdam: North-Holland, pp. 79–102.

10. Campbell, J. Y. and J. H. Cochrane (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, vol. 107, 205–251.

11. Cochrane, J. (2005): *Asset Pricing*, revised edition, Priceton University Press.

12. Cooley, T. and E. Prescott (1995): "Economic Growth and Business Cycles," in T. Cooley (ed.) *Frontiers of Business Cycle Research*, Princeton University Press.

13. Diamond, P. (1965): "National Debt in a Neo-Classical Growth Model," *American Economic Review*, vol. 55, 1126-1150.

14. Diaz-Jimenez, J., V. Quadrini and J.V. Rios-Rull (1997), "Dimensions of Inequality: Facts on the U.S. Distributions of Earnings, Income, and Wealth," *Federal Reserve Bank of Minneapolis Quarterly Review*, Spring.

15. Durrett, R. (1996): *Probability: Theory and Examples*, second edition, Duxbury Press.

16. Epstein, L. G. and S. E. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, vol. 57, 937–969.

17. Hansen, G. D. and A. Imrohoroglu (1992): "The Role of Unemployment Insurance in an Economy with Liquidity Constraints and Moral Hazard." *Journal of Political Economy*, vol. 100, 118–142.

18. Gale, D. (1973): "Pure Exchange Equilibrium of Dynamic Economic Models," *Journal of Economic Theory*, vol. 6, 12-36.

19. Huggett, M. (1993): "The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies," *Journal of Economic Dynamics and Control*, vol. 17, 953-969.

20. Huggett, M. (1996): "Wealth Distribution in Life-Cycle Economies," *Journal of Monetary Economics*, vol. 38, 469-494.

21. Hurd, M. D. (1989): "Mortaility Risk and Bequests," *Econometrica*, vol. 57, 779-813.

22. Imrohoroglu, A., S. Imrohoroglu and D. Joines (1995): "A Life Cycle Analysis of Social Security," *Economic Theory*, vol. 6, 83–114.

23. Judd, K. L. (1998): *Numerical Methods in Economics*, MIT Press.

24. King, R. G., C. I. Plosser and S. T. Rebelo. 1988. "Production, Growth and Business Cycles: I. The Basic Neoclassical Model," *Journal of Monetary Economics*, vol. 21, 195–232.

25. Kocherlakota, N. R. (1996): "The Equity Premium: It's Still a Puzzle," *Journal of Economic Literature*, vol. 34, 42–71.

26. Krueger D. (2002): *"Macroeconomic Theory" Lecture Notes*, Stanford University.

27. Krusell P. (2004): *Lecture Notes for Macroeconomic I*, Princeton University.

28. Krusell, P. and Smith, A. (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, vol. 106, 867-896.

29. Kydland, F. E. and E. C. Prescott (1982): "Time to Build and Aggregate Fluctuations," *Econometrica*, vol. 50, 1345–1371.

30. Ljungqvist, L. and T. Sargent (2000): *Recursive Macroeconomic Theory*, MIT Press.

31. Long, J. and C. Plosser (1983): "Real Business Cycles," *Journal of Political Economy*, vol. 91, 1345-1370.

32. Lucas, R. E., Jr. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, vol. 46, 1429–1445.

33. Lucas, R. E., Jr. (1987): *Models of Business Cycles*, Basil Blackwell.

34. Lucas, R. E. Jr. (1988): "On the Mechanics of Economic Development," *Journal of Monetary Economics*, vol. 22, 3-42.

35. Mehra, R., and E. C. Prescott (1985): "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, vol. 15, 145–162.

36. Prescott, E. C. (1986): "Theory Ahead of Business Cycle Measurement," *Federal Reserve Bank of Minneapolis Quarterly Review*, 9-22.

37. Rebelo, S. (1991): "Long-Run Policy Analysis and Long-Run Growth," *Journal of Political Economy*, vol. 99, 500-521.

38. Romer (1986): "Increasing Returns and Long Run Growth," *Journal of Political Economy*, vol. 94, 1002-1037.

39. Rosenlicht, M. (1968): *Introduction to Analysis*, Dover Publications Inc.

40. Solow, R. M. (1956): "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, vol. 70, 65–94.

41. Stokey, N. and R. Lucas, with E. C. Prescott (1989): *Recursive Methods in Economic Dynamics*, Harvard University Press.

42. Stokey, N. and S. Rebelo (1995): "Growth Effects of Flat-Rate Taxes," *Journal of Political Economy*, vol. 103, 519-550.

43. Tauchen, G. (1986): "Finite State Markov Chain Approximations to Univariate and Vector Autoregressions," *Economic Letters*, vol. 20, 177–181.