Lecture Notes[1]

Advanced Macroeconomic Theory II

Thorsten V. Koeppl

Queen's University

Winter Term 2017

[1]**Please note that I will be revising these notes throughout the term.**

Contents

Part I

Optimal Economies – A Review

Chapter 1

A Road Map

1.1 The Neo-classical Macroeconomic Model

Environment

- discrete time, infinite horizon: $t = 0, 1, \ldots$
- set of states: S at each time t with $s_0 \in S$ being the initial state at $t = 0$
- transition probabilities: $\pi_\tau(s^\tau | s^t)$ where $s^t = (s_0, s_1, \ldots, s_t)$ for all t and τ
  —▷ induce a probability measure $\pi_t(s^t)$ at time t
- alternatively: initial probability measure π_0 on S (where S is the set of all infinite paths over time)

Households

- preferences
$$\sum_{t=0}^{\infty} \sum_{s^t} \beta^t \pi_t(s^t) u(c_t(s^t)) \tag{1.1}$$
- endowment: one unit of labor, i.e. $n_t \in [0, 1]$ for all t

- u is twice continuously differentiable, strictly increasing, strictly concave and satisfies $\lim_{c \to 0} u_c(\cdot) = \infty$

<u>Technology</u>

- initial capital stock is given: k_0
- feasibility

$$c_t(s^t) + x_t(s^t) \leq A_t(s^t) F(k_t(s^{t-1}), n_t(s^t)) \tag{1.2}$$

- capital accumulation

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

- F is homogeneous of degree one: $F(k, n) = n f(\tilde{k})$, where $\tilde{k} \equiv \frac{k}{n}$
- $F_i(\cdot) > 0$, $F_{ii}(\cdot) < 0$, $\lim_{k \to 0} F_k(\cdot) = \lim_{n \to 0} F_n(\cdot) = \infty$, $\lim_{k \to \infty} F_k(\cdot) = \lim_{n \to \infty} F_n(\cdot) = 0$

This is the most stripped down description of an economy where intertemporal decisions (savings and investment) matter.

<u>Goal:</u>

We would like to understand

- what is "best" for this economy,
- what the economy can achieve "in equilibrium",
- and whether there is a role for "macroeconomic policy".

1.2 Social Planner's Problem

Sequential planning problem

$$\sup_{\{c_t,n_t,k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty}\sum_{s^t} \beta^t \pi_t(s^t) u(c_t(s^t)) \tag{1.4}$$

subject to

$$k_{t+1}(s^t) = (1-\delta)k_t(s^{t-1}) + A_t(s^t)F(k_t(s^{t-1}), n_t(s^t)) - c_t(s^t) \text{ for all } s^t$$

$$k_t(s^t) \geq 0, \quad c_t(s^t) \geq 0, \quad n_t(s^t) \in [0,1] \text{ for all } s^t$$

k_0 given

More generally, we look at problems of the sort

$$\sup_{\{u_t\}_{t=0}^{\infty}} E_0 \left[\sum_{t=0}^{\infty} \beta^t r(u_t, x_t)\right] \tag{1.5}$$

subject to

$$x_{t+1} = g(x_t, u_t, \epsilon_{t+1}) \text{ for all } t$$

x_0 given

where

- r is a "return" function
- g a constraint (more generally a constraint set)
- u_t "control variables"
- x_t "state variables"

Bellmann equation

$$V(x) = \sup_u r(u,x) + \beta E\left[V(g(x,u,\epsilon)\right] \tag{1.6}$$

How do we proceed?

- find policy fct. $u = h(x)$ to obtain endogenous law of motion $g(x, h(x), \epsilon)$

  —▷ existence of policy function

  —▷ h is time-invariant

  —▷ given x_0 we can fully describe the optimal dynamic behaviour of the economy

- need to solve for the value function V which is a fixed point of the Bellman equation

1.3 What lies ahead?

<u>Questions:</u>

1. How can we be sure that a solution to the recursive problem is a solution to the sequential problem?

   —▷ Principle of Optimality

2. How can we find the (unique) solution to a Bellman equation?

   —▷ contraction mapping theorem

   —▷ basic computational techniques

3. How can we be sure that the solution to the social planning problem is a competitive equilibrium?

   —▷ Welfare Theorems

<u>Remark:</u> With optimal economies, there is a tight connection between the social planning problem and competitive equilibrium. We can then use the following approach to solve for a competitive equilibrium of such an economy.

1. Define equilibrium and a planning problem.

2. Establish the welfare theorems.

3. Formulate the planning problem recursively.

4. Apply numerical methods to solve the recursive problem.

With "non-optimal economies" things are a bit more complicated as equilibrium and the solution of a planning problem do not correspond to each other. However, one can think of (macroeconomic) policy moving the economy to a (second- or first-best) optimum by changing people's behaviour.

1.4 Literature

Sargent & Ljunqvist – Ch. 3/4

Adda & Cooper – Ch. 2 and Ch. 5

Chapter 2

The Principle of Optimality

Question: How can we be sure that a solution of the Bellman equation solves the sequential social planning problem?

We only look at the case of no uncertainty.

Principle of Optimality:

"An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."

2.1 Main Result

Notation (follows SLwP):

- X state space
- $\Gamma : X \longrightarrow X$ constraint correspondence
- $F : A \longrightarrow I\!R$ return function, where $A = \{(x, y) \in X \times X | y \in \Gamma(x)\}$

Set of Feasible Plans:

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

Assumption 2.1.1. *1. $\Gamma(x)$ is nonempty for all $x \in X$.*

2. For all $x_0 \in X$ and all $x \in \Pi(x_0)$,

$$u(x) = \lim_{n\to\infty} \sum_{t=0}^{n} \beta^t F(x_t, x_{t+1}) \tag{2.1}$$

exists.

Define the sequential problem (SP), $V^* : X \longrightarrow I\!R \cup \{-\infty, +\infty\}$, as

$$V^*(x_0) = \sup_{x\in\Pi(x_0)} u(x) = \sup_{x\in\Pi(x_0)} \lim_{n\to\infty} \sum_{t=0}^{n} \beta^t F(x_t, x_{t+1}). \tag{2.2}$$

Define the corresponding functional equation (FE) as

$$V(x) = \sup_{y\in\Gamma(x)} F(x,y) + \beta V(y). \tag{2.3}$$

The key step is to realize that for all $x_0 \in X$ and any feasible plan $x \in \Pi(x_0)$,

$$\begin{aligned} u(x) &= \lim_{n\to\infty} \sum_{t=0}^{n} \beta^t F(x_t, x_{t+1}) && (2.4)\\ &= F(x_0, x_1) + \beta \lim_{n\to\infty} \sum_{t=0}^{n} \beta^t F(x_{t+1}, x_{t+2}) && (2.5)\\ &= F(x_0, x_1) + \beta u(x') && (2.6) \end{aligned}$$

where x' is the continuation of the plan after x_0, i.e. starting with x_1.

Theorem 2.1.2. *The function V^* satisfies the FE. Furthermore, if $\lim_{n\to\infty} \beta^n V(x_n) = 0$ for all $x \in \Pi(x_0)$ for all $x_0 \in X$, then $V = V^*$.*

2.1.1 Proof

See tutorial by TA and SLwP pp. 71-74.

2.1.2 Some intuition

Suppose the sup is attained in SP. Then, we have

$$V^*(x_0) \geq u(\hat{x}) \qquad \text{for all } \hat{x} \in \Pi(x_0).$$

In particular, fix any $x_1 \in \Gamma(x_0)$. Then, using the key step we discussed above,

$$V^*(x_0) \geq F(x_0, x_1) + \beta u(\hat{x}') \qquad \text{for all } \hat{x}' \in \Pi(x_1)$$

and in particular

$$V^*(x_0) \geq F(x_0, x_1) + \beta V^*(x_1).$$

Since x_1 was arbitrary, the result follows.

Conversely, suppose the sup is attained in the FE. Expanding the LHS of the functional equation

$$V(x_0) \geq F(x_0, x_1) + \beta V(x_1)$$

for any $x \in \Pi(x_0)$, we obtain

$$V(x_0) \geq F(x_0, x_1) + \beta F(x_1, x_2) + \cdots + \beta^n V(x_n).$$

Using the assumption on the limit in the statement of the Theorem, we get

$$V(x_0) \geq u(x_0)$$

which is just the definition of $V^*(x_0)$, since x was chosen arbitrarily.

<u>Remark:</u> We need the boundedness condition, since there could be solutions to FE which are *not* solutions to the SP (for example $V(x_0) = +\infty$ for all x_0 certainly is a solution to FE, but might not be a solution of the SP).

Remark: It can be very useful to establish uniqueness of the solution to a functional equation. If this is the case, we have by our result on the equivalence of value functions that $V = V^*$ immediately.

2.2 Equivalence of Policy Functions

Question: Given the value function corresponds to the value of the sequential problem, what about the policy function and the solution to the SP? Are they equivalent?

We call a plan $x^* \in \Pi(x_0)$ optimal if it attains the supremum V^*.

Theorem 2.2.1. *Let x^* be an optimal plan. Then*

$$V^*(x_t^*) = F(x_t^*, x_{t+1}^*) + V^*(x_{t+1}^*) \tag{2.7}$$

for all t.

Furthermore, any feasible plan $x^* \in \Pi(x_0)$ with $\limsup_{t\to\infty} \beta^t V^*(x_t^*) \leq 0$ that satisfies this equation is an optimal plan.

2.2.1 Proof

See discussion by TA.

2.3 Digression: Variational Approach

Suppose $\{x_{t+1}^*\}_{t=0}^{\infty}$ solves the sequential problem. Then for all t, x_{t+1}^* must solve

$$\max_y F(x_t^*, y) + \beta F(y, x_{t+2}^*) \tag{2.8}$$

subject to

$$y \in \Gamma(x_t^*) \tag{2.9}$$

$$x_{t+2}^* \in \Gamma(y) \tag{2.10}$$

There is no variation of x_{t+1}^* for any t that can improve upon the optimal policy. This yields a system of second-order difference equations described by the (necessary) Euler equations:

$$0 = F_y(x_t^*, x_{t+1}^*) + \beta F_x(x_{t+1}^*, x_{t+2}^*) \tag{2.11}$$

We have an initial condition x_0. Hence, we need the transverality condition as a boundary condition to characterize the optimal plan,

$$\lim_{t\to\infty} \beta^t F_x(x_{t+1}^*, x_{t+2}^*) \cdot x_t^* = 0. \tag{2.12}$$

Note that this is related to the functional equation. The first-order necessary condition for a solution to the functional equation is given by

$$0 = F_y(x, g(x)) + \beta v'(g(x)) \tag{2.13}$$

where $g(x)$ is the policy function. Using the envelope condition $v'(x) = F_x(x, g(x))$ yields the above expression. Note, however, that given v we have now a first-order difference equation.

2.4 Literature

Stockey, Lucas with Prescott – Ch. 4

Bertsekas (1976) – Main reference on Stochastic Dynamic Programming

Chapter 3

Basic Numerical Methods

3.1 Basic Algorithms

Recall that the social planning problem for the stochastic growth model can be reduced to a Bellmann equation of the form

$$V(x) = \sup_u r(u, x) + \beta E\left[V(g(x, u, \epsilon)|x\right]. \tag{3.1}$$

3.1.1 Guess and Verify

Good luck!

3.1.2 Value Function Iteration

Step 1: Define an appropriate state space $\mathcal{D}$ for x. Define an initial guess V_0. Define a convergence criterion.

Step 2: Solve $\max_u r(u, x) + \beta E\left[V_j(g(x, u, \epsilon)|x\right]$ for every $x \in \mathcal{D}$.

Step 3: Calculate $V_{j+1} = r(h(x), x) + \beta E\left[V_j(g(x, h(x), \epsilon)|x\right]$ as a new guess, where $h(x)$ is the solution of Step 2.

Step 4: Iterate until the convergence criterion is met.

Question: Why does this algorithm work?

Algorithm is built upon the operator

$$TV = \max_u r(u, x) + \beta E \left[V(g(x, u, \epsilon)|x\right]. \tag{3.2}$$

- operator T that is a contraction
- hence: convergence to a unique fixed point

FE corresponds to solution of the sequential problem due to the Principle of Optimality.

3.1.3 Policy Function Iteration

Step 1: Define an appropriate state space $\mathcal{D}$ for x. Define a convergence criterion. Define an initial guess for the policy function $h_0(x) \in \Gamma(x)$ for all x.

Step 2: Calculate $V_j(x) = r(h_j(x), x) + \beta E \left[V_j(g(x, h_j(x), \epsilon)|x\right]$ for every $x \in \mathcal{D}$.

Step 3: Find $h_{j+1} = \arg\max_u r(u, x) + \beta E \left[V_j(g(x, u, \epsilon)|x\right]$

Step 4: Iterate until the convergence criterion is met.

Idea:

- The value function takes into account the policy function *forever*, not only for one period.
- Each iteration (i.e. finding V_j) is costly, but fewer iterations on the value function.
- Finding V_j boils usually down to solving a system of linear equation – once we make the state space discrete (see below).

Two Questions:

1. How do we implement continuous state variables and value functions on a computer?
2. How can we speed up our algorithms?

3.2 Discretizing the State Space

- assume that a continuous state variable takes on only a finite number of values
- renders the problem finite-dimensional
- there are also a finite number of possible control policies

—▷ put state variable x on a grid: $x \in \{x_1, x_2, x_3, \ldots, x_N\}$

—▷ N-dimensional vector x

Question:

How can we implement the two algorithms?

3.2.1 Value function iteration

Our static maximization problem becomes

$$V_{l+1}(i) = \max_u r(x(i), u) + \beta \sum_{j=1}^{N} q_{ij}(u) V_l(j) \tag{3.3}$$

for all i.

—▷ V is now an N-dimensional vector

—▷ q_{ij} describes the transition probability of going from state x_i today to state x_j tomorrow

—▷ $\sum_{j=1}^{N} q_{ij} = 1$

—▷ in a deterministic problem: $q_{ij} \in \{0, 1\}$

3.2.2 Policy function iteration

Our problem of finding a new policy function becomes

$$h(i) = \arg\max_u r(x_i, u) + \beta \sum_{j=1}^{N} q_{ij}(u) V_l(j) \tag{3.4}$$

—▷ solution is an N-dimensional vector

—▷ to calculate the update for V, use the solution h and stack the vectors for q to obtain

$$V_{l+1} = r(x, h) + \beta Q V_{l+1} \tag{3.5}$$

or

$$V_{l+1} = (I - \beta Q)^{-1} r(x, h) \tag{3.6}$$

3.3 Exploiting concavity and monotonicity

Fix a state x. Define $v(u; x) = r(x, u) + \beta E[V(x'(u; x))]$ as the objective function in terms of the control variable u.

- When v is concave in the control variable u:

  —▷ if $v(u_i; x) > v(u_{i+1}; x)$, the value u_i is the maximizer

- When the policy function is monotone in the state variable x:

  —▷ if u_i is the maximizer for x_i, consider only values $u > u_i$ for $x > x_i$

Such insights will speed up computations considerably.

3.4 Curse of Dimensionality

—▷ so far we had one state variable with N values

—▷ with k state variables, we have N^k values

—▷ hence: if k is large, algorithms of this kind become computationally infeasible

3.5 Other Approaches

We will briefly introduce other ideas for directly computing dynamic equilibria. These methods become increasingly important when the 2nd WFTHM fails to apply.

3.5.1 Iterating on the Euler Equation

Formulate the Euler equation of either the sequential or recursive problem.[1] For example we have

$$u'(h(x)) - \beta(1+r)E[u'(h(x'))|x] = 0.$$

Contraction Operator:

- Think of the Euler equation as a functional equation.
- Show that the operator T on this equation is a contraction.
- Discretize the state space, fix an initial guess and iterate until convergence.

Example:

In Lucas (1978), the FONC for the consumers problem can be formulated as a contraction mapping, once one takes into account the market clearing conditions (see for example Cooley Ch. 3).

Monotone Operator:

[1] The envelope theorem applied to the derivative of the value function makes both equivalent.

- Discretize the state space and pick an initial guess for the policy function h^0.
- Use the Euler equation to generate a monotone (!) sequence of policy functions $\{h^n\}$ by starting from the initial guess.
- If the sequence of policy functions is uniformly bounded, they will convergence to a candidate solution h^*.
- Show that h^* must be the unique solution to the Euler equation (this can be hard!).

3.5.2 Projection Methods: Approximating the Euler Equation Directly

Idea:

- Recall that we can interpret the Euler equation by $T(h(x)) = 0$, where $T : B_1 \to B_2$ is an operator on a functional space of which we have to find a zero.
- Define an approximation $\hat{h}(x, a) = \sum_{i=1}^{n} a_i \phi(X_i)$, where ϕ is some type of polynomial function of degree n (or the type of approximation) and $a = (a_1, \ldots, a_n)$ are weights.
- We also might have to approximate the operator by some $\hat{T}$ to render the problem finite-dimensional.
- Note that we only have to find a finite number of parameters, the vector a.
- The weights a are chosen so that $\hat{T}(\hat{h}(x, a))$ is close to the zero function.

Problem:

When is $\hat{T}(\hat{h}(x, a))$ close to 0?

Procedure:

- Define a residual function
$$R(x;a) = \hat{T}(\hat{h}(\cdot, a))(x). \tag{3.7}$$
- Define a metric based on the following inner product
$$< f_1, f_2 >= \int_a^b f_1(x) f_2(x) \omega(x) dx$$
where $\omega(x)$ is a weighing function.
- Compute $< R(\cdot; a), p_i(\cdot) >$, where p_i is some test function for $i = 1, \ldots, l$.[2]
- Find the "best" a.
- Finally, check how good the solution is by approximating $T(\hat{h})$.

—▷ Example 1: Minimize Least Squares

Choose a to solve
$$\min_a \int [\hat{T}(\hat{h}(x,a)]^2 dx.$$
This is just minimizing the L^2 norm.

—▷ Example 2: Collocation Method (specific mass points)

Choose a to solve the following system of equations:
$$\hat{T}(\hat{h}(x_i, a)) = 0$$
for $i = 1, \cdots, n$.

—▷ Example 3: Method of Moments

Choose a to solve
$$< |\hat{T}(\hat{h}(x,a))|, x^{i-1} >= 0$$
for $i = 1, \cdots, n$.

<u>Remark:</u> It is quite important to choose an appropriate degree of approximation and method to assess how good the approximation is.

[2] If this inner product is 0, the functions are said to be orthogonal. Hence, the terminology projection method.

3.5.3 Approximate the Economy

There are two general approaches.

1. Approximate the value function with other functions that can easily be described by a finite number of parameters.

   Example:

   A quadratic approximation of the objective function yields linear decision rules.

2. Approximate the equilibrium locally around the steady state.

   Example:

   (Log)-linear approximation or second-order approximations

3.6 Literature

Judd – Ch. 12

Cooley – Ch. 3

Adda & Cooper – Ch. 3

Chapter 4

Competitive Equilibrium

4.1 Arrow-Debreu Equilibrium

Idea: At $t = 0$ there are markets where people trade state-contingent claims. These claims are fulfilled or executed later on as time and uncertainty unfolds. A crucial assumption is that there is perfect enforcement of these claims.

Prices

- $p_{k_0} \in I\!R_+$
- (q_0, r_0, w_0): stochastic processes
- for example: $\{q_t^0(s^t)\}_{t=0}^{\infty}$, where $q_t^0 : S^t \longrightarrow I\!R_+$

4.1.1 Households

$$\max_{\{c_t,\ell_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty}\sum_{s^t} \beta^t \pi_t(s^t) u(c_t(s^t), 1-\ell_t(s^t)) \tag{4.1}$$

$$\begin{aligned}
&\text{subject to} \\
&\quad \sum_{t=0}^{\infty}\sum_{s^t} q_t^0(s^t)c(s^t) \leq \sum_{t=0}^{\infty}\sum_{s^t} w_t^0(s^t)\ell_t(s^t) + p_{k_0}k_0 \\
&\quad c_t \geq 0, \quad \ell_t \in [0,1]
\end{aligned}$$

Remark: Households could also directly hold capital, make state-contingent investment decisions and rent it out in form of state-contingent claims to firms.

Standard FONCs:

$$\tfrac{u_c(s^t)}{u_\ell(s^t)} = \tfrac{q_t^0(s^t)}{w_t^0(s^t)} \tag{4.2}$$

$$\beta^{\tau-t}\tfrac{u_c(s^\tau)}{u_c(s^t)}\pi_t(s^\tau|s^t) = \tfrac{q_\tau^0(s^\tau)}{q_t^0(s^t)} \tag{4.3}$$

4.1.2 Firms

Two types: firms that produce output and firms that produce new capital

Firm I – Final good producers:

$$\max_{\{k_t^I,n_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty}\sum_{s^t} q_t^0(s^t)\left[c_t(s^t)+x_t(s^t)\right] - r_t^0(s^t)k_t^I(s^t) - w_t^0 n_t(s^t) \tag{4.4}$$

$$\begin{aligned}
&\text{subject to} \\
&\quad c_t(s^t)+x_t(s^t) \leq A_t(s^t)F(k_t^I(s^t), n_t(s^t)) \\
&\quad k_t^I \geq 0, \quad n_t \geq 0
\end{aligned}$$

Zero-Profit conditions:

$$q_t^0 A_t(s^t)F_k(s^t) = r_t^0(s^t) \tag{4.5}$$

$$q_t^0 A_t(s^t)F_n(s^t) = w_t^0(s^t) \tag{4.6}$$

Firm II – Capital good producers:

$$\max_{k_0^{II}, \{x_t\}_{t=0}^{\infty}} -p_{k_0} k_0^{II} + \sum_{t=0}^{\infty} \sum_{s^t} r_t^0(s^t) k_t^{II}(s^{t-1}) - q_t^0(s^t) x_t(s^t) \tag{4.7}$$

subject to

$$k_{t+1}^{II}(s^t) = x_t(s^t) + (1-\delta) k_t^{II}(s^{t-1})$$

$$k_0^{II} \geq 0, \quad x_t \geq 0$$

Zero-Profit conditions:

$$q_t^0 = \textstyle\sum_{s_{t+1}|s^t} \left[r_{t+1}^0(s^{t+1}) + (1-\delta) q_{t+1}^0(s^{t+1}) \right] \tag{4.8}$$

$$p_{k_0} = r_0^0(s_0) + (1-\delta) q_0^0(s_0) \tag{4.9}$$

4.1.3 Definition

Definition 4.1.1. *An* *Arrow-Debreu Equilibrium* *is given by prices (and pricing functionals)* $\{\hat{p}_{k_0}, \hat{q}^0, \hat{r}^0, \hat{w}^0\}$ *and allocations* $\{\hat{c}, \hat{l}, \hat{n}, \hat{x}, \hat{k}^I, \hat{k}_0^{II}\}$ *such that*

1. *taking prices as given,* $\{\hat{c}, \hat{\ell}\}$ *solve the households problem*

2. *taking prices as given,* $\{\hat{n}, \hat{k}^I\}$ *solve Firm I's problem and* $\{\hat{k}_0^{II}, \hat{x}\}$ *solve Firm II's problem*

3. *markets clear, i.e.*

$$\hat{c}_t(s^t) + \hat{x}_t(s^t) = A_t(s^t) F(\hat{k}_t^I(s^t), \hat{n}_t(s^t)) \text{ for all } s^t \text{ for all } t \tag{4.10}$$

$$\hat{\ell}(s^t) = \hat{n}(s^t) \text{ for all } s^t \text{ for all } t \tag{4.11}$$

$$\hat{k}_t^{II}(s^{t-1}) = \hat{k}_t^I(s^t) \text{ for all } s^t \text{ for all } t \tag{4.12}$$

$$\hat{k}_0^{II} = k_0 \tag{4.13}$$

4.2 Sequential Equilibrium

Idea: There are markets at every period t after each possible history s^t.

—▷ Markets are complete.

—▷ There is a full set of Arrow securities that "span" the uncertainty next period.

—▷ An Arrow security for state (s_{t+1}, s^t) in period t pays exactly one unit of the consumption good tomorrow (i.e. period $t+1$) if state s_{t+1} occurs.

—▷ It is sufficient to have $\#S$ such Arrow securities to have *dynamically complete markets*.

<u>Prices</u>

—▷ $q_t(s_{t+1}|s^t)$: price of the Arrow security

—▷ $a_{t+1}(s_{t+1}|s^t)$: number of securities held by the household

—▷ $\{w_t(s^t), r_t(s^t)\}_{t=0}^{\infty}$: sequence of factor prices

4.2.1 Households

$$
\begin{aligned}
&\max_{\{c_t,\ell_t,a_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty}\sum_{s^t} \beta^t \pi_t(s^t) u(c_t(s^t), 1-\ell_t(s^t)) && (4.14)\\
&\quad\text{subject to}\\
&\qquad c(s^t) + \sum_{s_{t+1}|s^t} q_t(s_{t+1}|s^t) a_{t+1}(s_{t+1}|s^t) \leq w_t(s^t)\ell_t(s^t) + a_t(s^t)\\
&\qquad a_0 \text{ given}\\
&\qquad a_{t+1}(s_{t+1}|s^t) \geq B(s^t) \text{ s.th. } B<0 \text{ large enough} && (4.15)\\
&\qquad c_t \geq 0, \quad \ell_t \in [0,1]
\end{aligned}
$$

<u>Question:</u>

Why do we need a borrowing limit?

Avoid "Ponzi"-schemes.

Suppose we have no borrowing limit. Then:

—▷ take any state contingent plan of the household $\{c, \ell, a\}$ at prices q and w

—▷ define a new sequence of consumption $\{\tilde{c}_t(s^t)\}_{t=0}^{\infty}$ by $\tilde{c}_0 = c_0 + \epsilon$ and $\tilde{c}_t(s^t) = c_t(s^t)$

—▷ the consumer can afford $\tilde{c}$ and is better off

Why? Borrow more and roll over the principal and interest forever.

<u>Idea:</u> Construct the new sequence of borrowing levels:

$$\tilde{c}_0 + \sum_{s_1} \tilde{a}_1(s_1|s_0)q_0(s_1|s_0) = c_0 + \sum_{s_1} a_1(s_1|s_0)q_0(s_1|s_0) \tag{4.16}$$

Hence, $\epsilon = \sum_{s_1} \left[a_1(s_1|s_0) - \tilde{a}_1(s_1|s_0)\right] q_0(s_1|s_0) > 0$.

Roll over this debt along a specific path $(\tilde{s}_1, \tilde{s}_2, \dots)$. From the sequential budget constraints one obtains

$$\begin{aligned}
\tilde{a}_1(\tilde{s}_1|s_0) &= a_1(\tilde{s}_1|s_0) - \epsilon \tfrac{1}{q_0(\tilde{s}_1|s_0)} \\
\tilde{a}_2(\tilde{s}_2|\tilde{s}^1) &= a_2(\tilde{s}_2|\tilde{s}^1) - \epsilon \tfrac{1}{q_0(\tilde{s}_1|s_0)q_1(\tilde{s}_2|\tilde{s}^1)} \\
&\vdots \\
\tilde{a}_{t+1}(\tilde{s}_{t+1}|\tilde{s}^t) &= a_{t+1}(\tilde{s}_{t+1}|\tilde{s}^t) - \epsilon \tfrac{1}{\Pi_{\tau=0}^{t} q_0(\tilde{s}_{\tau+1}|\tilde{s}^{\tau})} \\
&\vdots
\end{aligned}$$

For any sequence of state-contingent prices that could form an equilibrium, it must be the case that $0 < q_t(s_{t+1}|s^t) < 1$ (why? positive interest rates, see below.)

It cannot be the case that the initial sequence of asset holdings a_{t+1} is unbounded along the path (why? transversality condition, see below).

Hence, $\lim_{t\to\infty} \tilde{a}_{t+1} = -\infty$, i.e. debt along this path grows unboundedly large.

—▷ But: For any $B < 0$ and for any $\epsilon > 0$, we will eventually violate the borrowing constraint. A borrowing limit rules out such behavior.

Conclusion: It is necessary to impose a borrowing limit in order to make the problem economically interesing. However, one can always impose a small enough limit $B < 0$ such that the constraint is never binding given any price processes (q, r, w). Then, markets are still complete and the problem can be solved, "as-if" we neglected the constraint.

Standard FONC:[1]

$$\frac{u_c(s^t)}{u_\ell(s^t)} = \frac{1}{w_t(s^t)} \tag{4.18}$$

$$q_{t+1}(s_{t+1}|s^t) = \beta \frac{u_c(s^{t+1})}{u_c(s^t)} \pi_t(s_{t+1}|s^t) \tag{4.19}$$

$$\lim_{t \to \infty} E_0[\beta^t u_c(s^t) a_t(s_{t+1}|s^t)] = 0. \tag{4.20}$$

The last condition is a transversality condition (TVC) that expresses the fact that the household does not leave any wealth in net present value terms at infinity. It is a necessary condition for a solution of the problem.[2]

4.2.2 Firms

Two types: firms that produce output and firms that produce new capital

[1]The intertemporal Euler equations give an expression for the risk-free interest rate, here corresponding to holding a portfolio of assets that pays out one unit of consumption for sure next period. This rate is given by

$$\frac{1}{1+r} = \sum_{s^{t+1}|s^t} q(s_{t+1}|s^t) = \beta \frac{E\left[u_c(s_{t+1})|s^t\right]}{u_c(s^t)}. \tag{4.17}$$

[2]To my knowledge there is still a debate whether the TVC is a necessary condition for a standard concave problem. It is, however, a sufficient condition for a solution. Also note that people often confuse a TVC with a No-Ponzi game condition and vice versa. The former is a condition on optimality restricting the accumulation of wealth at infinity, while the latter is a constraint that limits the accumulation of debt.

Firm I – Final good producers:

$$\max_{\{k_t^I(s^t), n_t(s^t)\}} c_t(s^t) + x_t(s^t) - r_t(s^t)k_t^I(s^t) - w_t(s^t)n_t(s^t) \tag{4.21}$$

subject to

$$c_t(s^t) + x_t(s^t) \leq A_t(s^t)F(k_t^I(s^t), n_t(s^t))$$

$$k_t^I \geq 0, \quad n_t \geq 0$$

Zero-Profit conditions:

$$A_t(s^t)F_k(s^t) = r_t(s^t) \tag{4.22}$$

$$A_t(s^t)F_n(s^t) = w_t(s^t) \tag{4.23}$$

Firm II – Capital good producers:

—▷ This firm transforms output into capital and vice versa at the rate 1-1.

—▷ It decides to buy old capital and transform output into new capital at the end of the period, rent it out next period and liquidate it again after production.

—▷ The (total) pay-offs from capital are equal to

$$\left[r_t(s^{t+1}) + (1-\delta)\right] k_{t+1}^{II}(s^t). \tag{4.24}$$

—▷ To finance the purchase of capital/output the firm issues one-period state-contingent debt to households against these pay-offs.

—▷ This yields the following problem in state s^t

$$\max_{k_{t+1}^{II}(s^t)} k_{t+1}^{II}(s^t) \left\{ -1 + \sum_{s_{t+1}|s^t} q_{t+1}(s_{t+1}|s^t) \left[r_{t+1}(s^{t+1}) + (1-\delta)\right] \right\}. \tag{4.25}$$

—▷ Zero-profit condition:

$$1 = \sum_{s_{t+1}|s^t} q_{t+1}(s_{t+1}|s^t) \left[r_{t+1}(s^{t+1}) + (1-\delta)\right] \tag{4.26}$$

4.2.3 Definition

Definition 4.2.1. *A* <u>*Sequential Markets Equilibrium*</u> *is a sequence of prices* $\{\hat{q}_t(s_{t+1}, s^t), \hat{w}_t(s^t), \hat{r}_t(s^t)\}_{t=0}^{\infty}$ *and allocations* $\{\hat{c}_t(s^t), \hat{\ell}(s^t), \{\hat{a}(s_{t+1}|s^t)\}_{s_{t+1}}, \hat{n}_t(s^t), \hat{k}_t^I(s^t), \hat{k}_{t+1}^{II}(s^t)\}_{t=0}^{\infty}$ *such that*

1. *taking prices as given,* $\{\hat{c}_t(s^t), \hat{\ell}_t(s^t), \{\hat{a}(s_{t+1}|s^t)\}_{s_{t+1}}\}_{t=0}^{\infty}$ *solve the households problem*
2. *taking prices as given,* $\{\hat{n}_t(s^t), \hat{k}_t^I(s^t)\}_{t=0}^{\infty}$ *solve Firm I's problem and* $\{\hat{k}_{t+1}^{II}(s^t)\}_{t=0}^{\infty}$ *solve Firm II's problem*
3. *markets clear, i.e.*

$$\hat{c}_t(s^t) + \hat{k}_{t+1}^{II}(s^t) = A_t(s^t)F(\hat{k}_t^I(s^t), \hat{n}_t(s^t)) + (1-\delta)\hat{k}_t^{II}(s^{t-1}) \quad \forall s^t, t \tag{4.27}$$

$$\hat{\ell}(s^t) = \hat{n}(s^t) \quad \forall s^t, t \tag{4.28}$$

$$\hat{k}_t^{II}(s^{t-1}) = \hat{k}_t^I(s^t) \quad \forall s^t, t \tag{4.29}$$

$$\left[r_t(s^{t+1}) + (1-\delta)\right] k_{t+1}^{II}(s^t) = a_{t+1}(s_{t+1}|s^t) \quad \forall s^t, t \tag{4.30}$$

$$\left[r_0 + (1-\delta)\right] k_0^{II} = a_0 \tag{4.31}$$

<u>Remark:</u> Note that the initial wealth distribution for households is pinned down by the returns on using the initial capital stock in the first period. The idea is that type II firms have bought it from households in the past and pay them with the receipts. I took account of this as an additional initial condition where the household wealth is equal to the one period present value of the capital stock.[3] Note that in general, however, the sequential equilibrium is defined for an arbitrary initial wealth distribution for households.

4.3 Equivalence of AD and SM Equilibrium

See Proof by TA.

<u>Idea:</u>

—▷ Given an allocation, one can construct prices using the FONC and zero-profit conditions.

[3] Alternatively, households could have just owned capital making investment decisions directly.

—▷ Show that solutions of the household's problems coincide for non-binding borrowing constraints.

—▷ Given the constructed prices, one checks that the feasible sets of consumers and firms are identical. This involves showing that wealth levels (or that the net present values of future income, respectively) are finite given equilibrium prices. Then one can always find non-binding borrowing constraints.

Remark: In the remainder of the course such a general equivalence does not usually hold, as Arrow-Debreu markets describe an economy without distortions such as borrowing limits that are binding or missing markets.

4.4 Recursive Equilibrium

Idea: This concept is closely related to SM equilibrium. Everything is expressed in form of functionals that map from the state space into equilibrium values.

Assumption: Uncertainty is described by a (first-order) finite-state Markov process.

—▷ Markov process: A stochastic process $\{X_t\}_{t=0}^{\infty}$ is (first-order) Markov if for all $k \geq 1$ and all t,

$$\mathcal{P}\{X_{t+1}|X_t, \ldots, X_{t-k}\} = \mathcal{P}\{X_{t+1}|X_t\} \tag{4.32}$$

—▷ It follows that $\pi_t(s^t) = \pi(s_t|s_{t-1})\pi(s_{t-1}|s_{t-2}) \quad \cdots \quad \pi(s_1|s_0)\pi(s_0)$

—▷ Example: $A_t(s^t) = z(s^t)A$ for all t, where z is a Markov process

State space

—▷ aggregate (i.e. economy-wide) state variables

—▷ $X = (K, z)$

—▷ K: aggregate capital stock

—▷ z: productivity

Prices

—▷ functions of the aggregate state

—▷ $(r(X), w(X), q(X'|X))$

4.4.1 Households

—▷ additional, individual state variable: a – wealth level

—▷ restrict domain for a (and, hence, $a'(a, X)$) to capture a No-Ponzi-game condition

—▷ belief about law of motion on aggregate states induces conditional expectation $\tilde{E}$

—▷ $\tilde{E}$ is composed of exogenous stochastic process on z and a forecast about $K' = G(X)$

$$\begin{aligned} J(a, X) = \max_{c,l,a'} u(c, 1-\ell) + \beta \tilde{E}\left[J(a'(X'), X')\right] \qquad (4.33) \\ \text{subject to} \\ c + \sum_{X'} q(X'|X)a'(X') \leq w(X)\ell + a \\ c \geq 0, \ell \in [0,1], a' \in \mathcal{A} \end{aligned}$$

FONC:

$$\frac{u_c(a, X)}{u_l(a, X)} = \frac{1}{w(X)} \quad \text{for all } (a, X) \qquad (4.34)$$

$$\beta \frac{\partial \tilde{E}\left[J(a'(X'), X')\right]}{\partial a'(X')} = \mu(X) q(X'|X), \qquad (4.35)$$

$$(4.36)$$

where $\mu(X)$ is the Lagrange-multiplier on the resource constraint given state X.

Remark: If J is differentiable, use the Envelope Theorem to obtain $\frac{\partial J(a,X)}{\partial a} = \mu(X) = -u_c(a, X)$. Also, note that markets are complete here as long as people's beliefs are such that there is a unique value K' for any given current aggregate state (K, z). Then only the exogenous shock matters for uncertainty next period. Furthermore, if people believe the aggregate productivity shock z evolves according to some Markov process $\tilde{\pi}$, the intertemporal Euler equation can be rewritten as

$$\beta \frac{u_c(a'(a,X);X')}{u_c(a,X)} \tilde{\pi}(X',X) = q(X'|X). \tag{4.37}$$

4.4.2 Firms

—▷ identical to one-period problems in the SM equilibrium

—▷ zero-profit conditions

$$r(X) = zAF_k(k,n) \tag{4.38}$$

$$w(X) = zAF_n(k,n) \tag{4.39}$$

$$1 = \textstyle\sum_{X'} q(X'|X)\left[r(X') + (1-\delta)\right] \tag{4.40}$$

4.4.3 Definition

Definition 4.4.1. *A RE is price functions* $(r^*(X), w^*(X), q^*(X'|X))$*, a value function* J^**, decision rules* $(c^*(a,X), \ell^*(a,X), a'^*(a,X;X'))$*, a law of motion for* K'*,* $G(X)$*, such that*

1. *given price functions and the law of motion* G*, the decision rules* $(c^*(a,X), \ell^*(a,X), a'^*(a,X;X'))$ *and value function* J^* *solve the functional equation of the household*
2. *the zero-profit conditions hold for every* X
3. *(consistency)* $k = K$*,* $a = [r(X) + (1-\delta)]\,K$ *and* $\ell^*(a,X) = n(X)$
4. *the law of motion* G *is induced by the firm's zero-profit conditions and the household's decision rules, i.e.* $K' = G(X)$*, where*

$$G(X) = zAF(K, \ell^*([r(X)+(1-\delta)]\,K, X)) + (1-\delta)K - c^*([r(X)+(1-\delta)]\,K, X) \tag{4.41}$$

and people's belief are given by $\tilde{\pi} = \pi$.

Remark: Solving for a RE – as with any other equilibrium concept – is a fixed point problem.

—▷ decision makers forecast prices and law of motions

—▷ taking these objects as given they make optimal decisions

—▷ decisions must lead to prices and law of motions that are consistent with forecasts

—▷ need two iterations in order to solve for an equilibrium

—▷ one over the value functions and one over the prices/law of motion

Hence: Planning problem is much easier to solve. One can then invoke the Second Welfare Theorem to decentralize the economy and find prices directly from the FONC and optimal allocations.

4.5 Literature

Sargent and Ljungvist, Chapter 8-12

Cooley and Prescott, in: Cooley (Ch.1)

Chapter 5

Welfare Theorems

Main Set-up

—▷ commodity space is infinite-dimensional

—▷ need to define a price system on this commodity space

—▷ price system assigns a value to each element in the commodity space

—▷ value better be finite

5.1 Abstract Definition of an Economy

—▷ people $i \in I$

—▷ technologies $j \in J$

—▷ list of commodities: vector

—▷ commodity space: real-valued normed vector space $(S, \|\cdot\|_S)$

Households

—▷ consumption set $X_i \subset S$

—▷ preferences: $u_i : X_i \longrightarrow I\!R$

—▷ endowment: $\omega_i \in X_i$

Technologies

—▷ production set: $Y_j \subset S$

5.2 Prices

5.2.1 Dual Spaces

Definition 5.2.1. *A price system is a real-valued continuous linear functional $\phi : S \longrightarrow I\!R$.*

—▷ ϕ assigns a real-value to each element in the commodity space

—▷ ϕ is a linear functional if it satisfies

$$\phi(\alpha x + \beta y) = \alpha\phi(x) + \beta\phi(y) \tag{5.1}$$

for all $x, y \in S$ and all $\alpha, \beta \in I\!R$.

—▷ ϕ is continuous if $\|x_n - x\|_S \to 0$ implies $|\phi(x_n) - \phi(x)| \to 0$.

—▷ ϕ is continuous on S, if it is continuous at some $x \in S$.

—▷ ϕ is continuous if and only if ϕ is bounded; i.e., if there exists $M \in I\!R$ such that $|\phi(x)| \leq M\|x\|_S$ for all $x \in S$.

—▷ The space S^* is the space of all continuous linear functionals on S. It is called the dual of S.

—▷ It is a complete normed vector space when one applies the norm given by

$$\|\phi\|_d = \inf\{M \in \mathbb{R}_+ : |\phi(x)| \leq M\|x\|_S, \text{ for all } x \in S\} = \sup_{\|x\|_S \leq 1} |\phi(x)|$$

for every bounded linear functional ϕ on S.

5.2.2 Examples

—▷ ℓ_1 is the space of all sequences that are bounded in the norm $\|x\|_1 = \sum_{i=1}^{\infty} |x_i|$, i.e the series $\sum_{i=1}^{\infty} |x_i|$ converges.

—▷ ℓ_∞ is the space of all sequences that are bounded in the sup-norm.

—▷ ℓ_∞ is the dual of ℓ_1.

—▷ The converse is not true. ℓ_1 is only a subset of the dual of ℓ_∞.

—▷ ℓ_1 is the dual of $c_0 \subset \ell_\infty$, the space of all sequences that converge to 0.

—▷ More generally, for any $p \in (1, \infty)$, the space ℓ_p consists of all sequences that are bounded in the norm

$$\|x\|_p = \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{\frac{1}{p}}.$$

Furthermore, the dual of this space is given by ℓ_q, where $1/p + 1/q = 1$.

—▷ Finally, if one has a general measurable space, one can define the space of all measurable functions on this space that are bounded with respect to a similar norm. These are called the L_p spaces.

5.2.3 Why do we care?

—▷ Total consumption and endowment consists of an infinite sequence of consumption or endowment levels.

—▷ In order to define a budget set, we have to assign a value to any such sequence.

—▷ Prices need to define the value of any commodity in the economy. This value should be finite. Any element of the dual of the commodity space gives us a *price system*.

—▷ An inner-product representation of prices is useful, but not necessary.

—▷ Let ℓ_∞ be the commodity space. Any price system in ℓ_1 gives an inner product representation, i.e. for $\{c_t\}_{t=0}^{\infty} \in \ell_\infty$ and $\{p_t\}_{t=0}^{\infty} \in \ell_1$ we have

$$\sum_{t=0}^{\infty} p_t c_t < \infty. \tag{5.2}$$

—▷ More generally, as ℓ_q is the dual of ℓ_p, we have that any element of $y \in \ell_q$ defines a continuous linear functional on ℓ_p with the inner product representation

$$\phi(x) = y \cdot x = \sum_{i=1}^{\infty} y_i x_i.$$

—▷ An analogue exists for L_p spaces.

5.3 Definitions

Allocation: $(\{x_i\}_{i \in I}, \{y_j\}_{j \in J})$

Definition 5.3.1. *An Allocation is* feasible *if $x_i \in X_i$ and $y_j \in Y_j$ such that*

$$\sum_{i \in I} x_i - \sum_{j \in J} y_j = 0. \tag{5.3}$$

Definition 5.3.2. *A* Pareto-optimal Allocation *is an allocation $(\{x_i\}_{i \in I}, \{y_j\}_{j \in J})$ that is feasible and for which there does not exist another feasible allocation $(\tilde{x}, \tilde{y})$ such that (i) $u_i(\tilde{x}_i) \geq u_i(x_i)$ for all i and (ii) $u_i(\tilde{x}_i) > u_i(x_i)$ for some i.*

Definition 5.3.3. *A* Competitive Equilibrium *is an allocation $(\hat{x}, \hat{y})$ and a continuous linear functional $\hat{\phi} : S \to I\!R$ such that*

(i) $(\hat{x}, \hat{y})$ is feasible

(ii) for all i, $x \in X_i$ and $\phi(x) \leq \phi(\hat{x}) \Rightarrow u_i(x) \leq u_i(\hat{x})$

(iii) for all j, $y \in Y_j \Rightarrow \phi(y) \leq \phi(\hat{y})$.

—▷ The second condition is utility maximization.

—▷ The third condition is profit-maximization.

—▷ We do neither specify endowments nor how profits are distributed across consumers. This is not important for deriving the Welfare Theorems.

5.4 First Welfare Theorem

Assumption: Non-satiation

—▷ For all i and $x \in X_i$, there exists $x_n \to x$, where $x_n \in X_i$ for all n such that $u_i(x_n) > u_i(x)$ for all n.

Theorem 5.4.1. *Suppose preferences are non-satiated. If $(\hat{x}, \hat{y}, \phi)$ is a competitive equilibrium, then the allocation $(\hat{x}, \hat{y})$ is Pareto-optimal.*

Proof. `We first establish a preliminary result. This result shows that -- by using local non-satiation -- we can strengthen the optimality condition for consumer's in the definition of competitive equilibrium.`

Let $(\hat{x}, \hat{y}, \hat{\phi})$ be a competitive equilibrium. We first show that $u_i(x_i) = u_i(\hat{x}_i)$ implies $\phi(x_i) \geq \phi(\hat{x}_i)$.

Suppose to the contrary that for some i, $u_i(x_i) = u_i(\hat{x}_i)$ and $\phi(x_i) < \phi(\hat{x}_i)$. By non-satiation, there exists $x_n \to x_i$ such that $u_i(x_n) > u_i(x_i) = u_i(\hat{x}_i)$ for all n. Since ϕ is continuous, for n large enough, $\phi(x_n) < \phi(\hat{x}_i)$. Hence, $\hat{x}_i$ was not optimal for i and $(\hat{x}, \hat{y}, \phi)$ is not an equilibrium. A contradiction.

`Next, we show that if a CE allocation is not Pareto-optimal, there must be an allocation that is feasible at the equilibrium price and makes someone (here some ‘‘firm’’) better off.`

Suppose that $(\hat{x}, \hat{y})$ is not Pareto-optimal. Then there exists another feasible allocation $(\tilde{x}, \tilde{y})$ such that (i) $u_i(\tilde{x}_i) \geq u_i(\hat{x}_i)$ for all i and (ii) $u_i(\tilde{x}_i) > u_i(\hat{x}_i)$ for some i.

Since $(\tilde{x}, \tilde{y})$ is feasible and $(\hat{x}, \hat{y}, \phi)$ is an equilibrium, we have that $u_i(\tilde{x}_i) > u_i(\hat{x}_i)$ implies $\phi(\tilde{x}_i) > \phi(\hat{x}_i)$ and that $u_i(x_i) = u_i(\hat{x}_i)$ implies $\phi(x_i) \geq \phi(\hat{x}_i)$. Hence, by linearity of ϕ

$$\phi(\sum_{i\in I} \tilde{x}_i) = \sum_{i\in I} \phi(\tilde{x}_i) > \sum_{i\in I} \phi(\hat{x}_i) = \phi(\sum_{i\in I} \hat{x}_i). \tag{5.4}$$

Since both $(\tilde{x}, \tilde{y})$ and $(\hat{x}, \hat{y})$ are feasible, we have by linearity of ϕ

$$\sum_{j\in J} \phi(\tilde{y}_j) = \phi(\sum_{j\in J} \tilde{y}_j) = \phi(\sum_{i\in I} \tilde{x}_i) > \phi(\sum_{i\in I} \hat{x}_i) = \phi(\sum_{j\in J} \hat{y}_j) = \sum_{j\in J} \phi(\hat{y}_j). \tag{5.5}$$

Hence, $\hat{y}_j$ is not optimal for some $j \in J$. A contradiction. □

Remark: The FWThm is very strong. It only requires a very weak assumption on preferences.

5.5 Second Welfare Theorem

Assumptions:

(i) X_i for all $i \in I$ and $Y = \sum_{j\in J} Y_j$ are convex.

(ii) The utility function representing preferences is strictly concave and continuous.[1]

(iii) The set $Y = \sum_{j\in J} Y_j$ has an interior point.

In addition, the next theorem uses a very weak non-satiation assumption that applies only to *one* consumer *at* his particular Pareto-optimal consumption.

Theorem 5.5.1. *Suppose assumptions (i)-(iii) are satisfied. Let $(\hat{x}, \hat{y})$ be a Pareto-optimal allocation. Assume that for some $h \in I$ there is $\tilde{x}_h \in X_h$ such that $u_h(\tilde{x}_h) > u_h(\hat{x}_h)$. Then there exists a continuous linear functional $\phi : S \to I\!R$, $\phi \neq 0$, such that*

1. *for all $i \in I$, $x_i \in X_i$ and $u_i(x_i) \geq u_i(\hat{x}_i) \Rightarrow \phi(x_i) \geq \phi(\hat{x}_i)$*

2. *for all $j \in J$, $y \in Y_j \Rightarrow \phi(y) \leq \phi(\hat{y})$.*

[1] A weaker condition is that preferences are convex and continuous. This is the assumption being used in the proof of Debreu (1959).

The proof is an application of the **Hahn-Banach Theorem**:

Let S be a normed vector space. Let $A, B \subset S$ be convex sets. Assume that either $intB \neq \emptyset$ or S is finite-dimensional, and that $A \cap intB = \emptyset$. Then there exists a continuous linear functional $\phi \neq 0$ and a constant $c \in I\!R$ such that

$$\phi(x) \leq c \leq \phi(y) \tag{5.6}$$

for all $x \in A$ and $y \in B$.

Proof. `We first construct the sum of the upper-contour sets of all consumers at their Pareto-optimal consumption. Points in these sets reflect the points that make consumers better off relative to their Pareto-optimal allocation.`

Let $(\hat{x}, \hat{y})$ be a Pareto-optimal allocation. Define

$$A_i = \{x \in X_i | u_i(x) \geq u_i(\hat{x}_i)\}$$

for all $i \in I$. Let $A = \sum_{i \in I} A_i$. Then, the assumptions above ensure that A and Y are convex and that Y has an interior point.

`We then want to apply the Hahn-Banach Theorem with respect to these two sets. This is done by using local-nonsatiation for one consumer and deriving a contradiction with the original allocation being Pareto-optimal.`

We are left to show that $intY \cap A = \emptyset$. Suppose not. Then there exists $y \in A \cap intY$. Hence, there exists x such that $x_i \in A_i$ for all i and $y = \sum_{i \in I} x_i$. We have

$$u_i(x_i) \geq u_i(\hat{x}) \tag{5.7}$$

for all i.

Define $x_h(\lambda) = \lambda \tilde{x}_h + (1 - \lambda) x_h$ for $\lambda \in (0, 1)$. Then, by convexity of X_h, $x_h(\lambda) \in X_h$ and by concavity of u_h, $u_h(x_h(\lambda)) > u_h(\hat{x}_i)$.

Define $y(\lambda) = \sum_{i \neq h} x_i + x_h(\lambda)$. Since $y \in intY$, there exists $\epsilon > 0$ sufficiently small such that for $\lambda < \epsilon$, $y(\lambda) \in Y$. Then, the allocation $((x_1, \ldots, x_h(\lambda), \ldots, x_I), y(\lambda))$ is feasible and satisfies

$$u_i(x_i) \geq u_i(\hat{x}_i) \text{ for all } i \neq h \text{ and } u_h(x_h(\lambda)) > u_h(\hat{x}_h). \tag{5.8}$$

But then $(\hat{x}, \hat{y})$ was not Pareto optimal. A contradiction.

We then have a pricing functional that separates the two sets. By feasibility, the Pareto-optimal allocation must be on a boundary of both sets. The hyperplane defined by the function intersects the two sets exactly at a point common to the boundary of both sets.

Hence, by the Hahn-Banach Theorem, there exists $c \in I\!R$ and a continuous linear functional ϕ such that $\phi(y) \leq c \leq \phi(x)$ for all $y \in Y$ and $x \in A$.

Then, as $(\hat{x}, \hat{y})$ is feasible, we have $\hat{y} \in A$ and, thus, $\phi(y) \leq \phi(\hat{y})$. Furthermore, for any $x \in A$, $u_i(x_i) \geq u(\hat{x})$ implies $\phi(x) \geq \phi(\hat{x})$, as $\sum_{i \in I} \hat{x}_i = \hat{y} \in Y$. □

This proves only the existence of a quasi-equilibrium (where the household's choice minimizes costs or expenditure). We have not shown, however, that $\hat{x}$ is utility maximizing.

Theorem 5.5.2. *Suppose in addition that for each i there exists $\tilde{x}_i \in X_i$ such that $\phi(\tilde{x}_i) < \phi(\hat{x}_i)$. Then, $(\hat{x}, \hat{y}, \phi)$ is a competitive equilibrium.*

Proof. Let $x_i \in X_i$ such that $\phi(x_i) \leq \phi(\hat{x})$. Define $x_i(\lambda) = \lambda\tilde{x}_i + (1-\lambda)x_i$ for all $\lambda \in (0, 1)$. As X_i is convex and ϕ is linear, we have that $x_i(\lambda) \in X_i$ and that $\phi(x_i(\lambda)) < \phi(\hat{x})$ for all $\lambda \in (0, 1)$.

By the contrapositive of the previous theorem, we have that $u_i(x_i(\lambda)) < u_i(\hat{x}_i)$. Since u_i is continuous, $\lim_{\lambda \to 0} u_i(x_i(\lambda)) = u_i(x_i) \leq u_i(\hat{x}_i)$ which completes the proof. □

Remark: The SWThm requires very strong assumptions that for many models we discuss later are not fulfilled.

5.6 Some Further Comments

—▷ production or consumption sets can be non-convex (see also incomplete markets)

—▷ continuity of u_i depends on $\| \cdot \|_S$

—▷ non-empty interior of the production set also depends on the norm of S

—▷ chosing the norm is also important for having an inner-product representation of prices

5.7 Literature

Stokey, Lucas with Prescott, Chapter 15

Debreu, Theory of Value

Mas-Colell, Whinston and Green, Chapter 16

Chapter 6

Macroeconomics and Asset Pricing – A Primer

We develop now a basic theory of how to price assets. There are three different pricing methods – successively assuming more structure:

- arbitrage-free pricing
- consumption-based pricing
- general equilibrium pricing.

We will only talk mostly about the first two and how they are related. These provide most concepts used in the quantitative macro/finance literature on asset pricing. The last one simply adds market clearing conditions that are either derived from the dividend process or from the production side of the economy.

6.1 Set-up

- $t = 0, \ldots, T$ (where T can mostly interpreted as ∞)
- uncertainty is described by an event tree (or a filtration)

- probability measure $\pi(s^t)$ is well defined
- J long-lived assets with dividend process $x(s^t)$
- assets can be continuously traded ex-dividend
- preferences are described by expected utility
- utility is separable across time and states

6.2 Elementary Concepts

Let $\{a(s^t)\}_{t=0}^{\infty}$ be a *portfolio strategy* of holding the J assets over time. Its payoff is given by

$$z(a,p)(s^t) = \big(p(s_t) + x(s^t)\big)\, a(s^{t-1}) - p(s^t)a(s^t). \tag{6.1}$$

where s^t is a successor event of s^{t-1} – or $s^t \subset s^{t-1}$.

The *asset span* $\mathcal{M}(p)$ is defined by

$$\mathcal{M}(p) = \left\{ z |\text{ there exists } a(s^t) \text{ such that } z(s^t) = z(a,p)(s^t) \text{ for all } s^t, t \right\}. \tag{6.2}$$

In other words, the asset span is all payoffs over time that can be realized with a particular portfolio strategy. For finitely many events over the horizon, we have that $\mathcal{M}(p) \subset I\!R^k$ where k is the total number of events.

Remarks:

1. If $\mathcal{M}(p) = I\!R^k$. markets are complete; otherwise they are incomplete.
2. One does not need the same number of assets as events to have complete markets. It is sufficient to have the same number of assets (with linearly independent payoffs) as the number of successors (or branches). An example is a long-lived bond if there is no uncertainty that can be retraded at any point in time.

3. In continuous time, there is a result that one needs only a small number of assets to have complete markets. The idea is that the continuity of events is matched by trading a few assets continuously. Hence, the continuity of events is matched by the continuity of prices and trading strategies.

A *payoff pricing functional* is a linear functional

$$q : \mathcal{M}(p) \to I\!R \tag{6.3}$$

such that $q(z) = p_0 a_0$, where a generates the payoff z.

A *valuation functional* is a linear functional

$$Q : I\!R^k \to I\!R \tag{6.4}$$

such that $Q(z) = q(z)$ for all $z \in \mathcal{M}(p)$.

6.3 Arbitrage-Free Asset Pricing

An arbitrage is given by a portfolio strategy a such that $p_0 a_0 \leq 0$ and $z(a, p) \geq 0$ with at least one strict inequality.

Theorem 6.3.1. *The payoff pricing functional is strictly positive if and only if there is no arbitrage.*

Remarks:

1. If the utility functions are strictly increasing, then there is no arbitrage in equilibrium. Hence, the equilibrium payoff pricing functional excludes arbitrage and is strictly positive.

2. The valuation functional is unique if and only if securities markets are dynamically complete.

6.4 Portfolio Choice Problem

We consider the following maximization problem:

$$\max_{c,a} E\left[\sum_{t=0}^{\infty} \beta^t u(c_t)\right] \tag{6.5}$$

subject to (6.6)

$$c(s^t) = y(s^t) - \sum_j p_j(s^t)a_j(s^t) + \sum_j (p_j(s^t) + x_j(s^t))a_j(s^{t-1}) \tag{6.7}$$

$$c(s^t) \geq 0 \tag{6.8}$$

$$\lim_{t\to\infty} p_j(s^t)a_j(s^t) \geq 0 \tag{6.9}$$

where $s^t \subset s^{t-1}$.

We have the following first-order conditons:

$$\beta^t \pi(s^t)u'(c(s^t)) = \lambda(s^t) \tag{6.10}$$

$$-p_j(s^t)\lambda(s^t) + \sum_{s^{t+1}\subset s^t} \lambda(s^{t+1})(p_j(s^{t+1} + x_j(s^{t+1})) = 0 \quad \text{for all } j \tag{6.11}$$

$$\lim_{k\to\infty} E_t\left[\beta^k \lambda(s^{t+k}|s^t)a_j(s^{t+k})\right] = 0 \quad \text{for all } j. \tag{6.12}$$

This yields for the condition,

$$p_j(s^t)u'(c(s^t)) = E_t\left[(p_{j,t+1} + x_{j,t+1})\beta u'(c_{t+1})\right] \tag{6.13}$$

for the payoff process of asset or portfolio j or equivalently,

$$E_t\left[p_{j,t+1} + x_{j,t+1}\right] E_t\left[\beta \frac{u'(c_{t+1})}{u'(c_t)}\right] + Cov_t\left[p_{j,t+1} + x_{j,t+1}, \beta \frac{u'(c_{t+1})}{u'(c_t)}\right] = p_{j,t}. \tag{6.14}$$

Written with returns we have

$$u'(c(s^t)) = E_t\left[r_{j,t+1}\beta u'(c_{t+1})\right]. \tag{6.15}$$

and if there is a risk-free asset – i.e., an asset that has the same one-period return across states in $t+1$ –

$$r^*_{j,t+1} = \frac{u'(c(s^t))}{E_t[\beta u'(c_{t+1})]}. \tag{6.16}$$

We obtained again that,

$$\frac{r^*_{j,t+1} - r_{j,t+1}}{r^*_{j,t+1}} = \frac{Cov(r_{j,t+1}, u'(c_{t+1}))}{u'(c_t)}, \tag{6.17}$$

which defines the one-period risk premium for asset or portfolio j.

The covariance term thus expresses a measure of how risk is assessed by the consumer. Hence, every asset can be priced according to a risk-free payoff and an additional risk premium that expresses how well the asset insures the consumer against consumption risk.

6.5 Fundamental Asset Pricing Formula

Question:

How do we price an arbitrary asset that has a payoff given by $\{x(s^t)\}_{t=0}^{\infty}$?

We will look at different concepts:

- event prices
- pricing kernel
- risk-neutral probabilities

6.5.1 Event Prices

We first define event prices through the payoff pricing functional q for complete markets. Consider an Arrow security that pays one unit of the consumption good in event s^t. The *event price* $q(s^t)$ is the price of this security. Then, we have for any payoff process

$$q(z) = \sum_{s^t} q(s^t) z(s^t) = qz. \tag{6.18}$$

More generally, for any given asset structure where markets are complete, event prices are defined by the system of equations

$$q(s^t)p_j(s^t) = \sum_{s^{t+1}\subset s^t} q(s^{t+1})(p_j(s^{t+1}) + x_j(s^{t+1})) \quad \text{for all } j \tag{6.19}$$

where we must have $q(s^t) > 0$, if there is no arbitrage. Of course event prices are linked to the intertemporal marginal rates of substitution in any model where people optimize over asset holdings.

Suppose now asset markets are incomplete. The above system of equations that defines event prices allows for multiple solutions as there are more events than assets. These event prices define a valuation functional, i.e. $q(s^t) = Q(s^t)$. Hence, we have that $Q(z) = \sum_{s^t} q(s^t)z(s^t)$. Of course, the valuation functional is not unique anymore.

How do we value any new asset $\tilde{z}$ that increases the asset span, so that prices are arbitrage free? The value $Q(\tilde{z})$ needs to be in the interval $[q_\ell(\tilde{z}), q_u(\tilde{z})]$, where[1]

$$\begin{aligned} q_u &= \min_a \{p_0 a_0 | z(a,p) \geq \tilde{z}\} && (6.20)\\ q_\ell &= \min_a \{p_0 a_0 | z(a,p) \leq \tilde{z}\}. && (6.21) \end{aligned}$$

6.5.2 Pricing Kernel

We start with a powerful mathematical result.

Riesz Representation Theorem:

Let X be a Hilbert space. If $\phi : X \to I\!R$ is a continuous linear functional, then there exists a unique $y \in X$ such that

$$\phi(x) = E[yx]$$

for all $x \in X$.

[1] Alternatively, we could use a new set of event prices to derive these bounds.

For complete markets, the payoff pricing functional is a continuous linear functional on the entire payoff space. Hence, no arbitrage and a positive payoff pricing functional are equivalent to the existence of a so-called *pricing kernel* which is defined by M_t such that

$$q(z) = \sum_{t=0}^{\infty} E[M_t z_t]. \tag{6.22}$$

for all z_t. Alternatively, M_t is called *stochastic discount factor* or *state-price deflator*. This is the fundamental asset pricing equation in economics.

One can show that

$$M_t p_{j,t} = E_t[M_{t+1}(p_{j,t+1} + x_{j,t+1})] \tag{6.23}$$

for all j and t. Hence, the pricing kernel is a stochastic process that is closely related to event prices and, hence, the intertemporal marginal rates of substitutions.

This implies immediately that as long as there exists a risk-free asset that

$$r^*_{t+1} = \frac{M_t}{E_t[M_{t+1}]} \tag{6.24}$$

and that in dynamically complete markets we have that

$$M(s^t) = \frac{q(s^t)}{\pi(s^t)}, \tag{6.25}$$

so that the pricing kernel is just the event prices rescaled by event prices.[2]

For incomplete markets, exactly the same relationships hold – except for the fact that now the pricing kernel is not unique anymore. The idea is that the asset span $\mathcal{M}(p)$ is a closed linear subspace of the underlying payoff space. One can then decompose the pricing kernel in one component that lies in the asset span $\mathcal{M}(p)$ and in another one that lies in the orthogonal complement of $\mathcal{M}(p)$.

[2]For the special case of risk-neutrality, we obtain that $M(s^t) = \beta^t$, so that the price of any asset is simply its discounted expected net present value of dividends, $q_0 = \sum_{t=0}^{\infty} \beta^t E[x_t]$.

6.5.3 Some Final Remarks

Martingale Theories of Asset Prices

This theory does not predict that asset prices follow a martingale. To obtain a martingale, we would need that the IMRS is constant and that there is no covariance between this term and payoffs.

However a related concept "risk neutral probabilities" can be used to obtain a martingale theory of asset prices. These probabilities are adjusted by the relative weight of the IMRS across states. Under these probabilities, the discounted gains of an asset follows a martingale. Hence, it looks like as if we price assets according to risk neutral agents, but the new probabilities subsume the price of risk.

Ruling out Bubbles

So far, we have ruled out any problems "at infinity" when we priced assets. In general, it need not be the case that the pricing or valuation functionals map into $I\!R$ or converge to infinite series of the event prices times the payoffs of a portfolio strategy. However, if we use equilibrium pricing, it is very hard to obtain bubbles in equilibrium. Consider the Lucas' Asset pricing model. We need to have

$$p_t = \beta E_t \left[\frac{u'(y_{t+1})}{u'(y_t)} (y_{t+1} + p_{t+1}) \right] \tag{6.26}$$

Iterating forward and using the law of iterated expectations, we get that

$$p_t = E_t \left[\sum_{k=1}^{\infty} \beta^k \frac{u'(y_{t+k})}{u'(y_t)} y_{t+k} \right] + E_t \left[\lim_{k \to \infty} \beta^k u'(y_{t+k}) p_{t+k} \right]. \tag{6.27}$$

The last term must be zero in equilibrium, since it would violate households making optimal decisions.

6.6 Literature

Le Roy and Werner (2004) – Principles of Financial Economics

Duffie (2001)

Santos and Woodford (2000) – Econometrica

Part II

Business Cycle Analysis

Chapter 7

Real Business Cycles

7.1 The Canonical Model

The RBC model forms the basis for any Dynamic Stochastic General Equilibrium Model.

A representative household solves the following problem

$$\max_{c_t,k_t,n_t} E_0 \left[\sum_{t=0}^{\infty} \beta^t \left(\frac{c_t^{1-\gamma}}{1-\gamma} + \theta \frac{(1-n_t)^{1-\eta}}{1-\eta} \right) \right] \tag{7.1}$$

subject to

$$c_t + x_t \leq w_t n_t + r_t k_t \text{ for all } t \text{ and } z_t \tag{7.2}$$

$$k_{t+1} = x_t + (1-\delta) k_t \tag{7.3}$$

$$k_0 \text{ and } z_0 \text{ given} \tag{7.4}$$

The production function is Cobb-Douglas and given by

$$z_t A_t k_t^{\alpha} n_t^{1-\alpha} \tag{7.5}$$

where

$$\ln z_t = \rho \ln z_{t-1} + \epsilon_t \tag{7.6}$$

with $\rho \in (0,1)$ and $\epsilon_t \sim \mathcal{N}(0,\sigma)$.

At the core of the model is the neo-classical growth model enriched by a labour-leisure choice.

The only uncertainty in the model arises from the technology shock z_t which is log-normally distributed.

Hence, any fluctuations around the steady state arise from this source; in other words, technology shocks are the sole driver of business cycles.

7.2 Dynamics

The first-order conditions of the firm are given by

$$z_t \alpha \left(\frac{k_t}{n_t}\right)^{\alpha-1} = r_t \tag{7.7}$$

$$z_t(1-\alpha)\left(\frac{k_t}{n_t}\right)^{\alpha} = w_t \tag{7.8}$$

The households first-order conditions are

$$\left(\frac{c_t^{-\gamma}}{\theta(1-n_t)^{-\eta}}\right) = \frac{1}{w(z_t)} \text{ for all } t \text{ and } z_t \tag{7.9}$$

$$1 = E\left[\beta\left(\frac{c_t}{c_{t+1}}\right)^{\gamma}(r_{t+1}+(1-\delta))\,\Big|\, z_t\right] \text{ for all } t \text{ and } z_t \tag{7.10}$$

Market clearing and the process of the technology shock complete the model

$$c_t + k_{t+1} = z_t A k_t^{\alpha} n_t^{1-\alpha} + (1-\delta)k_t \text{ for all } t \text{ and } z_t \tag{7.11}$$

$$\ln z_t = \rho \ln z_{t-1} + \epsilon_t \tag{7.12}$$

The (deterministic) steady state is given by

$$\left(\frac{\bar{c}^{-\gamma}}{\theta(1-\bar{n})^{-\eta}}\right) = \frac{1}{A(1-\alpha)\bar{n}^{-\alpha}}$$

$$1 = \beta\left(A\alpha k_t^{\alpha-1}\bar{n}^{1-\alpha} + (1-\delta)\right)$$

$$\bar{c} + \bar{k} = \bar{z}F(\bar{k},\bar{n}) + (1-\delta)\bar{k}$$

We have three equations in three unknowns that we can solve.

The dynamics of the system are described by a non-linear stochastic second-order difference equation since

- c_t depends on k_t and k_{t+1}
- c_{t+1} depends on k_{t+1} and k_t
- the expectation of z_{t+1} matters for investment.

7.3 Consumption and Leisure Choice

7.3.1 Consumption

Using $1 = \beta\left(\bar{r} + (1-\delta)\right)$, we obtain

$$1 = E_t\left[\left(\frac{c_t}{c_{t+1}}\right)^{\gamma}\left(\frac{r_{t+1} + (1-\delta)}{\bar{r} + (1-\delta)}\right)\right] \tag{7.13}$$

Log-linearizing, we have approximately

$$E_t\left[\ln\left(\frac{c_{t+1}}{c_t}\right)\right] \approx \frac{1}{\gamma}\left(E_t[r_{t+1}] - \bar{r}\right). \tag{7.14}$$

The savings/investment choice (and, thus, consumption growth) depends on the expected changes in the return of capital and the willingness to substitute intertemporally.

The response of current consumption depends on the income and substitution effect (or σ).

7.3.2 Leisure

We have from the first-order condition that

$$\theta(1-n_t)^{-\eta} = \lambda(z_t)w(z_t) \tag{7.15}$$

where $\lambda(z_t)$ is the Lagrange multiplier on the budget constraint in state z_t.

Log-linearizing, we obtain for this equation[1]

$$\theta(1-\bar{n})^{-\eta-1}\bar{n}\hat{n}_t = \bar{\lambda}\bar{w}\lambda(z_t) + \bar{\lambda}\bar{w}\hat{w}(z_t) \tag{7.16}$$

where $\bar{n}$ and $\hat{n}$ are steady state values and (log-)deviations from steady state.

Hence, using the steady state relationship, we have

$$\hat{n}_t = \frac{1}{\eta}\frac{1-\bar{n}}{\bar{n}}\left(\lambda(z_t) + \hat{w}(z_t)\right) \tag{7.17}$$

This shows that both $\bar{n}$ and η matter a lot for how the labour-leisure choice reacts to changes in wages. For low η and $\bar{n}$ one obtains unrealistically large responses of labour supply to wage changes.

7.4 The Way Forward

To solve the model, one usually follows the following steps.

1. Find values for parameters, either through calibration or estimation.
2. Log-linearize the economy around the steady state.
3. Check whether the economy is stable.
4. Solve the "state-space" representation of the log-linearized model.
5. Compute impulse response functions and second moments.
6. Compare the results to actual data, simulated or estimated data.

For most (or all) of these steps, `DYNARE` is your friend.

[1] Try it!

7.5 Appendix – Log-linearization

Define the (approximate) deviation from steady state as

$$\hat{x}_t = \log\left(\frac{x_t}{\bar{x}}\right) = \log\left(\frac{x_t - \bar{x}}{\bar{x}} + 1\right) \approx \frac{x_t - \bar{x}}{\bar{x}}. \tag{7.18}$$

We then approximate any equation $y_t = f(x_t)$ as a linear equation in terms of percentage deviations from steady state values

$$\bar{y}e^{\hat{y}_t} = f(\bar{x}e^{\hat{x}_t}) \tag{7.19}$$

A first-order Taylor expansion around $\hat{y}_t = 0$ and $\hat{x}_t = 0$ yields

$$\bar{y}e^0 + \bar{y}e^0(\hat{y}_t - 0) \approx f(\bar{x}e^0) + f'(\bar{x}e^0)\bar{x}e^0(\hat{x}_t - 0) \tag{7.20}$$

which is

$$\bar{y} + \bar{y}\hat{y}_t \approx f(\bar{x}) + f'(\bar{x})\bar{x}\hat{x}_t \tag{7.21}$$

and since $\bar{y} = f(\bar{x})$

$$\bar{y}\hat{y}_t = f'(\bar{x})\bar{x}\hat{x}_t. \tag{7.22}$$

A general formula is given by

$$y_t = f(x_t, z_t) \Longrightarrow \bar{y}\hat{y}_t = f_x(\bar{x}, \bar{z})\bar{x}\hat{x}_t + f_z(\bar{x}, \bar{z})\bar{z}\hat{z}_t \tag{7.23}$$

Hence, we obtain the following "rules" for log-linearizing equations:

$$x_{t+1} = f(x_t) \Longrightarrow \hat{x}_{t+1} = f'(\bar{x})\hat{x}_t \tag{7.24}$$

$$y_t = x_t z_t \Longrightarrow \hat{y}_t = \hat{x}_t + \hat{z}_t \tag{7.25}$$

$$y_t = \frac{x_t}{z_t} \Longrightarrow \hat{y}_t = \hat{x}_t - \hat{z}_t \tag{7.26}$$

$$y_t = x_t + z_t \Longrightarrow \bar{y}\hat{y}_t = \bar{x}\hat{x}_t + \bar{z}\hat{z}_t \tag{7.27}$$

$$y_t = x_t^{\epsilon} \Longrightarrow \hat{y}_t = \epsilon\hat{x}_t \tag{7.28}$$

$$0 = g(x_t, y_t) \Longrightarrow \hat{y}_t = -\frac{g_x(\bar{x}, \bar{y})\bar{x}}{g_y(\bar{x}, \bar{y})\bar{y}}\hat{x}_t \tag{7.29}$$

Note that after log-linearizing a dynamic system everything is expressed as % deviations from steady state with the parameters expressing elasticities.

7.6 Literature

King and Rebelo (1998) – (Old) Handbook of Macroeconomics Article

Cooley and Prescott (1995) – Frontiers of Business Cycles

Chapter 8

Solving DSGE Models

8.1 Linear Difference Equations

8.1.1 First order

Consider the linear first-order difference equation

$$x_t = \phi x_{t-1} + b \tag{8.1}$$

where $|\phi| < 1$.

Note that this equation is equivalent to

$$x_t = \psi x_{t+1} + c \tag{8.2}$$

where $\psi = 1/\phi$ so that $|\psi| > 1$ and $c = -b/\phi$.

Let's iterate the first difference equation backwards and use the fact that $|\phi| < 1$

$$\begin{aligned} x_t &= \phi^t x_0 + b \sum_{k=0}^{t-1} \phi^k && (8.3) \\ &= \phi^t x_0 + b \frac{1-\phi^t}{1-\phi} && (8.4) \\ &= \phi^t \left(x_0 - \frac{b}{1-\phi} \right) + \frac{b}{1-\phi} && (8.5) \end{aligned}$$

The first part is the particular solution given an initial condition, whereas the second part is the general solution.

With the second difference equation, we would get the same result if we iterate forward with a particular solution being given by a terminal condition at some period $t + j$.

<u>Result:</u>

We iterate stable roots backwards and unstable roots forward to obtain a solution to a linear first-order difference equation. Stability means that we converge to a long-run fixed point (think steady state) and requires that for

$$x_t = \phi x_{t-1} + b \tag{8.6}$$

we have $|\phi| < 1$.[1]

<u>Remark:</u>

Note that by using this convention we tend to make an assumption that is not innocuous. We often assume that we are only looking at non-explosive solutions that converge in the long-run. The time-paths of backward- or forward-looking solutions are different in general so that we pick some sort of “appropriate” solution.

8.1.2 Higher Order

In DSGE models we often have second-order difference equations. Consider then

$$y_{t+1} + ay_t + by_{t-1} = c \tag{8.7}$$

We can then turn this into a system of first-order difference equations by defining

$$x_t = y_{t-1}. \tag{8.8}$$

[1]Note that for $\phi \in (0, 1)$ we get monotone convergence, while for $\phi \in (-1, 0)$ we get dampening oscillations.

We then have

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_{t+1} \\ x_{t+1} \end{pmatrix} = \begin{pmatrix} -a & b \\ 1 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ x_t \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} c \\ c \end{pmatrix} \tag{8.9}$$

which is just a stacked system of first-order difference equations. We can then apply the same techniques (but in matrix form) as with a single difference equations.

We will come back to this below when we talk about state space represenations.

8.1.3 Stochastic

The above can be easily extended to have stochastic elements in the difference equations.

Consider a (forward-looking) difference equation

$$E_t x_{t+1} = \alpha x_t + z_t \tag{8.10}$$

where z_{t+1} – and, hence, x_{t+1} – are random variables. Suppose further we have an unstable root, i.e. $|\alpha| > 1$.

Iterate forward to obtain

$$x_t = \frac{1}{\alpha}(E_t x_{t+1} - z_t) \tag{8.11}$$

$$x_t = -\frac{1}{\alpha} z_t - \sum_{s=1}^{\infty} \left(\frac{1}{\alpha}\right)^{s+1} E_t z_{t+s} \tag{8.12}$$

which is the solution as we did not specify any terminal condition.

8.2 State-Space Representation

8.2.1 Basics

In economics, we deal with systems of difference equations. Hence, it is convenient to write them in the form of the so-called state-space representation.

$$x_{t+1} = Ax_t + Cw_{t+1} \tag{8.13}$$
$$y_t = Gx_t$$
$$x_0 \sim N(\mu_0, \Sigma_0)$$

- state x_t ($n \times 1$ vector)
- iid shocks $w_t \sim \mathcal{N}(0, I)$ ($m \times 1$ vector)[2]
- observations or jump variables y_t (k vector)
- $n \times n$ transition matrix A
- $n \times m$ volatility matrix C
- $k \times n$ output matrix G

Note that drawing shocks and initial conditions pins down sequences for (x_t, y_t). Without those draws, we can still say something about the (long-run) distribution of those variables.

These distributions are described as follows:

- x_t and y_t are normally distributed
- $E[x_{t+1}] = \mu_{t+1} = A\mu_t$
- $Var[x_{t+1}] = \Sigma_{t+1} = A\Sigma_t A' + CC'$ since $x_{t+1} - \mu_{t+1} = A(x_t - \mu_t) + Cw_{t+1}$
- $E[y_t] = G\mu_t$
- $V[y_t] = G\Sigma_t G'$

[2] A less restrictive assumption would be to require martingale difference equations which are martingales with zero conditional expectations.

8.2.2 Example

Let's look once again at the example of an RBC model with a utility function that is linear in labour and full depreciation of capital.

Since the economy is efficient we can simply look at the social planning problem which is given by

$$\max E_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\gamma}}{1-\gamma} - \chi n_t \right] \tag{8.14}$$

subject to

$$c_t + k_{t+1} = z_t k_t^{\alpha} n_t^{1-\alpha} \tag{8.15}$$

$$\ln z_{t+1} = \rho \ln z_t + \epsilon_{t+1} \tag{8.16}$$

where ϵ_{t+1} is $\mathcal{N}(0, \sigma)$.

The solution is described by

$$c_t^{-\gamma} = \beta E_t \left[c_{t+1}^{-\gamma} (\alpha z_{t+1} k_{t+1}^{\alpha-1} n_t^{1-\alpha}) \right] \tag{8.17}$$

$$\chi = c_t^{-\gamma} z_t (1-\alpha) k_t^{\alpha} n_t^{-\alpha} \tag{8.18}$$

$$c_t + k_{t+1} = z_t k_t^{\alpha} n_t^{1-\alpha} + k_t \tag{8.19}$$

$$\ln z_{t+1} = \rho \ln z_t + \epsilon_{t+1} \tag{8.20}$$

How do we solve this now?

We can log-linearize the system to obtain

$$-\gamma \hat{c}_t = E_t[-\gamma + \hat{z}_t + (1-\alpha)\hat{n}_t - (1-\alpha)\hat{k}_{t+1}] \tag{8.21}$$

$$\gamma \hat{c}_t = \hat{z}_t + \alpha \hat{k}_t - \alpha \hat{n}_t \tag{8.22}$$

$$\hat{c}_t \bar{c} + \hat{k}_{t+1} \bar{k} = \left(\hat{z}_t + \alpha \hat{k}_t + (1-\alpha)\hat{n}_t \right) \bar{y} \tag{8.23}$$

$$\hat{z}_{t+1} = \hat{z}_t + \epsilon_{t+1} \tag{8.24}$$

where $\hat{x}_t$ is the deviations in logs from the steady state value $\bar{x}$. Note that to derive the first approximation, we have used the fact that $1 = \beta\alpha\bar{y}/\bar{k}$.

This is a system of first-order difference equations.

8.3 Blanchard and Kahn

We can write this system in state-space form.

We have

$$E_t \begin{pmatrix} 1 & \alpha - 1 & 1 - \alpha & -\gamma \\ \alpha & \bar{k} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{z}_{t+1} \\ \hat{k}_{t+1} \\ \hat{n}_{t+1} \\ \hat{c}_{t+1} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -\gamma \\ \bar{y} & \alpha\bar{y} & (1-\alpha)\bar{y} & -\bar{c} \\ \rho & 0 & 0 & 0 \\ 1 & -\alpha & -\alpha & -\sigma \end{pmatrix} \begin{pmatrix} \hat{z}_t \\ \hat{k}_t \\ \hat{n}_t \\ \hat{c}_t \end{pmatrix} \tag{8.25}$$

or

$$\mathbf{A}E_t[x_{t+1}] = \mathbf{B}x_t \tag{8.26}$$

Hence, the solution is given by

$$x_t = \mathbf{B}^{-1}\mathbf{A}E_t[x_{t+1}] \tag{8.27}$$

provided the matrix B is non-singular and, hence, invertible.[3]

We first group the variables into

- "predetermined" and exogenous (state) variables w_t
- "control" (jump) variables y_t.

Note that the above system has done this already.[4]

[3] If this is not the case, one can still solve this system by using the Generalized Schur Decomposition (see Klein (2000), JEDC and note that `DYNARE` uses that method).

[4] One can also set a system where the exogenous state variables or shocks appear in the form of

$$\mathbf{A}E_t[x_{t+1}] = \mathbf{B}x_t + \mathbf{C}z_t$$

Using the Jordan Decomposition we have

$$\mathbf{B}^{-1}\mathbf{A} = \mathbf{P}^{-1}\mathbf{\Lambda}\mathbf{P} \tag{8.28}$$

where $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues and $\mathbf{P}$ contains the corresponding eigenvectors.

Using this we obtain that

$$E_t \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix} \begin{pmatrix} \tilde{w}_{t+1} \\ \tilde{y}_{t+1} \end{pmatrix} = \begin{pmatrix} \tilde{w}_t \\ \tilde{y}_t \end{pmatrix} \tag{8.29}$$

where

$$\begin{pmatrix} \tilde{w}_{t+1} \\ \tilde{y}_{t+1} \end{pmatrix} = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} \begin{pmatrix} w_t \\ y_t \end{pmatrix} \tag{8.30}$$

The beauty of this approach is that now the two equations are decoupled and can be solved separately. We are of course interested in the law of motion of the state variables w_t, since these give us a full solution to our problem.

8.4 Stability

This approach is useful, because it gives directly conditions for there to be a unique solution to this problem. Consider the matrix of eigenvalues $\mathbf{\Lambda}$.

We have that

- Λ_1 corresponds to backward looking variables so that we need stable roots or $|\lambda_{i1}| > 1$ for all i.
- Λ_2 corresponds to forward looking variables so that we need unstable roots or $|\lambda_{i2}| < 1$ for all i.

Proposition 8.4.1. *If the number of unstable roots is equal to the number of controls, we have a unique solution.*

If the number of unstable roots is larger than the number of controls, we have no solution.

If the number of unstable roots is smaller than the number of controls, we have indeterminancy (multiple solutions or sunspots).

Assume than that we have a unique, stable solution. We then get that the solution requires $\tilde{y}_t = 0$ so that

$$y_t = -P_{22}^{-1} + P_{21}w_t \tag{8.31}$$

$$w_t = (P_{11} - P_{12}P_{22}^{-1}P_{21})^{-1}\tilde{w}_t \tag{8.32}$$

Now, the good news is that `DYNARE` does everything for you: log-linearization, checking B+K condition, solving the difference equations.

8.5 Method of Undetermined Coefficients

An alternative method is to guess and verify the solution.

Consider the system of second-order difference equations given by

$$E_t[Fx_{t+1} + Gx_t + Hx_{t-1} + Lz_{t+1} + Mz_t] \tag{8.33}$$

where $z_{t+1} = Nz_t + \epsilon_{t+1}$ and $E_t[\epsilon_{t+1}] = 0$.

Guess that the solution has the form

$$x_t = Px_{t-1} + Qz_t. \tag{8.34}$$

Use this guess twice in the expectational expression above and use the expression for z_{t+1} to calculate the expectations. P is then the solution to the matrix quadratic equation[5]

$$FP^2 + GP + H = 0 \tag{8.35}$$

The solution for Q is obtain from

$$VQ = -vec(LN + M) \tag{8.36}$$

where V is a matrix obtained from P and the matrices of parameters.

The TA will go over an example that applies this method with "sensitivity" and not brute force.

8.6 Impulse Response Functions

To analyze the localized dynamics around the steady state, one can construct impulse response functions.

1. Start at an initial condition (i.e. steady state) for $(x_0, y_0, z_0) = (\bar{x}, \bar{y}, \bar{z})$.
2. Shock the system with a (one standard deviation or 1 %) shock to w_1 assuming that there are no further shocks.[6]
3. Recursively, calculate x_t, y_t and z_t.

The interpretation is % deviations of a variable from steady state if the model has been rewritten in log-deviations from steady state.

[5]This procedure is equivalent to solve a single second-order difference equation by noting that

$$x_{t+2} + ax_{t+1} + bx_t = x^t \left(x^2 + ax + b\right) = 0$$

[6]One could also use a particular sequence of shocks $\{w_t\}_{t=1}^{T}$.

Stability should mean that a temporary (permanent) shock causes the system to return to the old (new) steady state.

8.7 Second Moments

These can either be obtained by simulation as sample moments or by direct calculations from the solution of the model.

8.8 Appendix – Eigenvalues and Eigenvectors

For a quick reference and interpretation of these concepts, please refer to Sargent & Stachurski (2017).

8.9 Literature

Sargent & Stachurski (2017) – Section on Linear State Space Models

Uhlig (1999) – Ch. 3 in Marimon & Scott, Computational Methods for the Study of Dynamic Economies

Blanchard and Kahn (1980)

Chapter 9

The New Keynesian Model

9.1 Overview

There are two key deviations from the RBC model.

1. Equilibrium will be demand determined. Why? Monopolistic competition.
2. There are some frictions that cause demand to fluctuate.

Policy (monetary/fiscal) matters to reduce the frictions.

The model boils down to a system of three equations in three unknowns (y, i, π); i.e. output, nominal interest rates and inflation.

Policy is effective, since it can (in the short run!) influence the real interest rate.

9.2 Households

There are now many goods indexed by $i \in [0, 1]$.

Households value only aggregate consumption which is assumed to be given by

$$C_t = \left(\int_0^1 C_t(i)^{1-\frac{1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon-1}} \tag{9.1}$$

with $\epsilon > 1$.

<u>Household's Problem:</u>

$$\max_{C_t(i), N_t, B_t} E_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{C_t^{1-\sigma}}{1-\sigma} - \frac{(N_t)^{1+\eta}}{1+\eta} \right) \tag{9.2}$$

subject to

$$\int_0^1 P_t(i) C_t(i) di + Q_t B_t \leq B_{t-1} + W_t N_t - T_t \text{ for all } t \tag{9.3}$$

where all prices are expressed in *nominal* terms.[1]

<u>Question:</u>[2]

How do we choose $C_t(i)$ to achieve the maximum aggregate consumption, holding fixed the total expenditure at some level Z_t?

$$\max_{C_t(i)} C_t \tag{9.4}$$

subject to

$$\int_0^1 P_t(i) C_t(i) = Z_t \tag{9.5}$$

The first-order conditions yield

$$\frac{C_t(i)}{C_t(j)} = \left(\frac{P_t(i)}{P_t(j)} \right)^{-\epsilon}. \tag{9.6}$$

The parameter ϵ is the elasticity of substitution between two goods.

[1] ADD

[2] Note that what follows is being used in many contexts where a variety of goods plays an important role such as IO or international trade.

We can now define the aggregate price index by

$$P_t = \left(\int_0^1 P_t(i)^{1-\epsilon} di \right)^{\frac{1}{1-\epsilon}} \tag{9.7}$$

Plug in $C_t(i)$ in the expenditure constraint to get

$$C_t(j) P_t(j)^\epsilon \int_0^1 P_t(i)^{1-\epsilon} di = Z_t \tag{9.8}$$

$$C_t(j) = \frac{Z_t}{P_t} \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon} \tag{9.9}$$

From the definition of C_t, we have that $Z_t = P_t C_t$. Hence,

$$C_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon} C_t \tag{9.10}$$

This links aggregate demand to the demand for each single good:

- Each good is consumed in proportion to aggregate demand.
- The factor of proportionality decreases in the good's price relative to the price level.
- The price elasticity of demand for any good is given by ϵ.

We can then express the household problem only in terms of aggregate consumption and the aggregate price level.

$$\frac{C_t^\sigma}{(1 - N_t)^\eta} = \frac{W_t}{P_t} \tag{9.11}$$

$$1 = \beta E_t \left[\left(\frac{C_t}{C_{t+1}} \right)^\sigma (1 + i_t) \frac{P_t}{P_{t+1}} \right] \tag{9.12}$$

Hence, the Euler equation remains unchanged so that we have exactly the same micro-foundations as in the RBC model.

Why? $(1 + i_t) P_t / P_{t+1}$ is the real interest rate.

9.3 Firms

Production is linear in labour and there is no capital.

Taking the demand function as given, firms set prices as a monopolist to maximize profits.

$$\max_{P_t(i)} P_t(i)Y_t(i) - W_t N_t(i) \tag{9.13}$$

subject to

$$Y_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\epsilon} C_t \tag{9.14}$$

$$N_t(i) = \left(\frac{Y_t(i)}{A_t}\right)^{\frac{1}{\alpha}} \tag{9.15}$$

The first order condition yields

$$P_t(i)\frac{\partial Y_t(i)}{\partial P_t(i)} + Y_t(i) - W_t \frac{1}{\alpha A_t}\left(\frac{Y_t(i)}{A_t}\right)^{\frac{1}{\alpha}-1} \frac{\partial Y_t(i)}{\partial P_t(i)} = 0 \tag{9.16}$$

where

$$\frac{\partial Y_t(i)}{\partial P_t(i)} = (-\epsilon)\frac{Y_t(i)}{P_t(i)}$$

since ϵ is the price elasticity of demand.

Hence, we obtain the familiar *mark-up condition* of monopoly pricing.

$$P_t(i) = \left(\frac{\epsilon}{\epsilon - 1}\right)\frac{W_t N_t(i)}{\alpha Y_t(i)} \equiv \mu \varphi_t(i).$$

- $\varphi_t(i)$ are the nominal marginal costs when producing $Y_t(i)$
- μ is the mark-up
- $1/\mu$ can be interpreted as the real marginal costs
- ϵ measures the market power of firms

9.4 Sticky Prices

9.4.1 Optimal Price Choice

Since marginal costs are the same across all firms, firms would all set the same price.

Assume instead "Calvo-pricing". Every period, a fairy appears and allows a fraction of $(1-\theta)$ firms to change their price.

This implies that each firm has a probability of $(1-\theta)$ to change its price.

Firms then solve the following problem

$$\max_{P_t(i)} \sum_{k=0}^{\infty} \theta^k E_t \left[Q_{t,t+k} \left(P_t(i) Y_{t+k}(i) - W_{t+k} \left(\frac{Y_{t+k}(i)}{A_{t+k}} \right)^{\frac{1}{\alpha}} \right) \right]$$

subject to

$$Y_{t+k}(i) = \left(\frac{P_t(i)}{P_{t+k}} \right)^{-\epsilon} C_{t+k}$$

where $Q_{t,t+k}$ captures "stochastic equilibrium discounting".

The first-order condition is given by

$$\sum_{k=0}^{\infty} \theta^k E_t \left[Q_{t,t+k} Y_{t+k}(i) \left(P_t(i) - \frac{\epsilon}{\epsilon - 1} \varphi_{t+k}(i) \right) \right] = 0$$

Note that all firms that can change prices today, will chose the same price, P_t^*.

9.4.2 Why Does Inflation Increase Output?

When the price cannot be adjusted in period $t+k$, the firm is stuck at P_t^* and cannot charge its desired mark-up μ so that the actual mark-up is

$$\mu_t = \frac{P_t^*}{\varphi_{t+k}(i)} \neq \mu = \frac{\epsilon}{\epsilon - 1}.$$

The firm thus sets labour demand to satisfy its demand for goods

$$N_{t+k}(i) = \left(\frac{Y_{t+k}(i)}{A_{t+k}}\right)^{\frac{1}{\alpha}} = \left(\frac{P_t^*}{P_{t+k}}\right)^{-\frac{\epsilon}{\alpha}} \left(\frac{C_{t+k}}{A_{t+k}}\right)^{\frac{1}{\alpha}}$$

For fixed P_t^* we have $\partial N_{t+k}(i)/\partial P_{t+k} > 0$, so that nominal marginal costs

$$\varphi_{t+k}(i) = \frac{W_t N_{t+k}(i)}{\alpha Y_{t+k}(i)}$$

are large.

Result:

In other words, mark-ups are depressed, i.e. $\mu_t < \mu$ or, equivalently, real marginal costs are high. Mark-ups can be interpreted as a labour wedge so that an unexpected price increase will depress some mark-ups or, equivalently, increase employment and, hence, output.

Remark:

Note that again there is no unemployment, as in equilibrium income levels will adjust in such a fashion as to clear the labour market. Notwithstanding, shocks to inflation will have real effects as some firms cannot change their prices. Importantly, these output fluctuations are inefficient, as they decrease household's utility beyond the long-run distortion of monopolistic competition.

9.4.3 The New Keynesian Phillips Curve

The aggregate *price level* in period t is given by

$$P_t^{1-\epsilon} = \int_{i|fixed} P_{t-1}(i)^{1-\epsilon} di + (1-\theta) P_t^{*1-\epsilon} \tag{9.17}$$

The distribution of fixed prices corresponds to the distribution of last periods prices with weight θ, or

$$\theta P_{t-1}^{1-\epsilon} = \int_{i|fixed} P_{t-1}(i)^{1-\epsilon} di \tag{9.18}$$

Hence, the inflation dynamics are given by

$$\Pi_t = \left[\theta + (1-\theta) \left(\frac{P_t^*}{P_{t-1}} \right)^{1-\epsilon} \right]^{\frac{1}{1-\epsilon}} \tag{9.19}$$

so that inflation changes less than 1-1 with price changes of individual firms.

Log-linearizing and using the firm's FOC gives[3]

$$\pi_t = \beta E_t[\pi_{t+1}] + \lambda \left(\log \frac{\bar{\varphi}_t}{P_t} - \log \frac{\epsilon - 1}{\epsilon} \right) \tag{9.20}$$

$$= \beta E_t[\pi_{t+1}] + \kappa (y_t - y_t^n) \tag{9.21}$$

and iterating forward

$$\pi_t = \lambda \sum_{t=0}^{\infty} \beta^k E_t \left[\log \frac{\bar{\varphi}_{t+k}}{P_{t+k}} - \log \frac{\epsilon - 1}{\epsilon} \right] \tag{9.22}$$

$$= \lambda \sum_{t=0}^{\infty} \beta^k E_t \left[\left(\sigma + \frac{\eta + (1-\alpha)}{\alpha} \right) \left(y_{t+k} - y_{t+k}^n \right) \right] \tag{9.23}$$

where

- $\frac{\bar{\varphi}_t}{P_t}$ are average real marginal costs for firms
- $\lambda = \frac{(1-\theta)(1-\beta\theta)}{\theta} \frac{\alpha}{\alpha + \epsilon(1-\alpha)}$
- $\kappa = \lambda \left(\sigma + \frac{\eta + (1-\alpha)}{\alpha} \right)$

Result:

Inflation is given by *expected* deviations from steady-state mark-up. Inflation is high (low) whenever firms expect real marginal costs above (below) their steady state values. The expression in the second line is the *output gap* which measures the deviation of the actual output level from the output level associated with flexible prices, y_t^n, which is not (!) the optimal output level.

[3]For details on the algebra, see Gali (2008).

9.5 Monetary Policy

The model then boils down to three equations determining inflation, the output gap and nominal interest rates.

Phillips Curve

$$\pi_t = \beta E_t[\pi_{t+1}] + \kappa(y_t - y_t^n) \tag{9.24}$$

IS equation

$$y_t - y_t^n = -\frac{1}{\sigma}(i_t - E_t[\pi_{t+1}] - r_t^n) + E_t[y_{t+1} - y_{t+1}^n] \tag{9.25}$$

where $r_t^n = \rho + \sigma E_t[y_{t+1}^n - y_t^n]$ is the natural rate of interest which changes due to real (or supply) shocks.

Taylor Rule

$$i_t = \bar{i} + \phi_\pi(\pi_t - \bar{\pi}) + \phi_y(y_t - y_t^n) \tag{9.26}$$

The last equation describes the reaction function of a central bank that sets nominal interest rates.

This rule specifies that the central bank reacts to deviations from an inflation target $\bar{\pi}$ which we can normalize (should?) to 0 and an *output gap* which we call x_t.

It is important to realize that the output gap is specified as deviations in output from the *natural rate of output* which corresponds to the output level with fully flexible prices. Hence, the central bank only reacts to excess deviations relative to the fluctuations of this natural rate.

9.6 Dynamics

The model can then be rewritten as

$$\begin{pmatrix} x_t \\ \pi_t \end{pmatrix} = \mathbf{A} \begin{pmatrix} E_t[x_{t+1}] \\ E_t[\pi_{t+1}] \end{pmatrix} + \mathbf{B}(r_t^n - \bar{r}_t^n - v_t) \tag{9.27}$$

where $\mathbf{A}$ and $\mathbf{B}$ are functions of parameters.

The term v_t is a monetary policy shock and we could extend it to more shocks.

Essentially, we can have three fundamental shocks (one per equation). Modern DSGE models usually have one shock per equation. More about that later on.

For determinacy/stability, we need to have eigenvalues of $\mathbf{A}$ to be less than 1 in modulus. This is the case if and only if

$$\kappa(\phi_\pi - 1) + (1 - \beta)\phi_y > 0. \tag{9.28}$$

This is called the *Taylor principle*. Taylor's original estimates for the reation coefficients are $\phi_\pi = 1.5$ and $\phi_y = 0.5/4$.

A positive shock to the interest rate delivers

- lower inflation
- higher real interest rates
- a lower output gap

which are consider reasonable impulse responses. Why? Equilibrium nominal interest rates increase and, hence, there is a so-called "liquidity effect".

However, one needs a sufficiently high degree of price stickiness for this result.

Note that in response to a positive technology shock, the reaction function implies a decrease in the nominal interest rate. Why? Since prices are sticky, there is a negative output gap.

9.7 Optimal Policy

The optimal monetary policy would simply set $i_t = r_t^n$.

This would remove all fluctuations in the economy that arise from the IS equation.

This is the so-called "divine coincidence of monetary policy" that states that there is no trade-off between inflation and output.

One problem is that such a policy leads to indeterminancy in the model. Why? (see Homework!)

One can still rely on a Taylor rule to obtain determinancy and the optimal policy.

There will be a trade-off with shocks to the Phillips curve (mark-up shocks or cost-push shocks).

9.8 Some Remarks

One can have different frictions built into the model. An important one is a nominal wage rigidity.

It is not clear how sticky prices or wages are. The empirical literature finds different results.

One can endogenize the degree of stickiness in prices. Examples are menu costs or the literature on inattentiveness.

9.9 Literature

Gali (2008)

Woodford (2003)

Chapter 10

Bayesian Estimation of Model Parameters

10.1 The Kalman Filter

In dynamic models, we often have a situation where some variables cannot be directly observed, but others can be measured repeatedly and have information about the non-observable variables.

How can we use repeated measurements to update our forecasts of the unobserved variables?

Consider a spacecraft that needs to land on the moon. We have an estimate of the initial condition and receive a noisy signal about the current location.

We want to predict where the spacecraft is right now from the prior information, its projected flight path and the noisy observation.

Two common applications in macroeconomics:

1. Dynamic models with imperfect information and noisy signals

2. DSGE models and Bayesian Estimation

10.1.1 Set-up

We have a state-space representation with a law of motion

$$x_{t+1} = Ax_t + w_{t+1} \tag{10.1}$$

where $w_{t+1} \sim \mathcal{N}(0, Q)$ and observations

$$y_t = Gx_t + v_t \tag{10.2}$$

where $v_t \sim \mathcal{N}(0, R)$.

What is crucial here is (i) that we have linear relationships and (ii) that noise and shocks are normally distributed random variables.

10.1.2 A Simple Example

Let's assume with have only one state variable and one variable that we observe (i.e. we approach the moon on a line and not in space!).

Hence, A and G are just numbers and the variances of the shock and noise are given by σ_w^2 and σ_v^2.

For convenience, we can normalize $G = 1$.

A Heuristic Solution

We know that x and y have a linear relationship

$$x_t = y_t - v_t \tag{10.3}$$

where v is the noise term in this relationship.

We would like to estimate

$$E[x|y] = \beta_0 + \beta_1 y \tag{10.4}$$

Since we satisfy all assumption for the classic linear regression model, we obtain immediately that the OLS estimates are given by

$$\beta_0 = \bar{x} - \beta_1 \bar{y} \tag{10.5}$$

$$\beta_1 = \frac{Cov(x,y)}{V(y)} = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_v^2} \tag{10.6}$$

Let's assume now, we know that $x \sim \mathcal{N}(\mu_x, \sigma_x^2)$. Thus, the unconditional means are given by $\bar{x} = \bar{y} = \mu_x$, since v is just white noise.

This gives us the estimate for x being

$$E[x|y] = \beta_1 \mu_x + (1 - \beta_1) y. \tag{10.7}$$

Note that for this result, we have (implicitly!) used the fact that the prior information on x is random and has a variance given by σ_x^2. This in a sense is a deviation from the classical regression model where $Cov(x,y)$ is a quantity observed from the data.

Now, we can simply use the law of motion on x_{t+1} and the estimated x to find a forecast for the new state.

<u>Question:</u> Can we actually forecast the variance of x_{t+1} conditional on y with this approach?

We have that

$$V[x_{t+1}|y_t] = a^2 V[x_t|y_t] + V[w_{t+1}] \tag{10.8}$$

so all we need to do is find $V[x_t|y_t]$. This conditional variance tells us how much variance there is left if we use our estimate of $E[x_t|y_t]$ to predict x_t.

$$
\begin{aligned}
V[x_t|y_t] &= E[(x_t - E[x_t|y_t])^2|y_t] \\
&= E[((x_t - (1-\beta_1)\mu_x - \beta_1(x_t - v_t))^2 \, |y_t] \\
&= E[(1-\beta_1)^2(x_t-\mu_x)^2 - 2\beta_1(1-\beta_1)v_t(x_t-\mu_x) + \beta_1^2 v_t^2|y_t] \\
&= (1-\beta)^2 E[(x_t-\mu_x)^2|y_t] + \beta_1^2\sigma_v^2 \qquad (10.9)
\end{aligned}
$$

since v_t is orthogonal to x_t, y_t has no information on v_t and $E[v_t] = 0$.

Importantly, we have

$$E[(x_t-\mu_x)^2|y_t] = E[(x_t-\mu_x)^2] = V(x_t) \qquad (10.10)$$

Why? The observation has information for estimating a new mean for x_t, but it has no information for the deviation of x_t from its old mean.

This yields

$$V[x_t|y_t] = \frac{\sigma_x^2\sigma_v^2}{\sigma_x^2+\sigma_v^2} \qquad (10.11)$$

A Bayesian Solution

Filtering

Goal: Estimate x from the noisy signal y given prior information.

We use Bayes' rule for the prior information $x \sim \mathcal{N}(\mu_x, \sigma_x^2)$ to get

$$p(x|y) = \frac{p(x,y)}{p(y)} = \frac{p(y|x)p(x)}{p(y)} \qquad (10.12)$$

The key is to note that given $X = x$, we have that $y \sim \mathcal{N}(x, \sigma_v^2)$ and that $p(y) = \int p(y|x)p(x)dx$ is just a number.

Introspection implies that $p(x|y) \sim \mathcal{N}(\mu_1, \sigma_1^2)$ where

$$\mu_1 = \alpha\mu_x + (1-\alpha)y \tag{10.13}$$

$$\sigma_1^2 = \frac{1}{\sigma_x^{-2} + \sigma_v^{-2}} \tag{10.14}$$

$$\alpha = \frac{\sigma_v^2}{\sigma_x^2 + \sigma_v^2} \tag{10.15}$$

What do we then think x is? Well, take the maximum likelihood of the posterior probability.

This is intuitive. The best estimate of x is now a weighted average between the prior knowledge and the information contained in the observation (aka Bayesian learning).

The weight depends on the *signal-to-noise ratio.* For example, for very uninformative signals (large σ_v^2) a lot of weight remains on the prior. and little on the observation.

The term $1 - \alpha$ is called the *Kalman Gain* and expresses how much new information is incorporated from the noisy signal.

Forecasting

Goal: Predict tomorrow's state x_{t+1} given today's estimated x_t.

The key insight is that $x_{t+1} = ax_t + w_{t+1}$ is a sum of normally distributed variables and, thus, normally distributed itself.

We then have

$$E[x_{t+1}|y_t] = E[ax_t + w_{t+1}|y_t] = aE[x_t|y_t] = \alpha\mu_x + (1-\alpha)y \tag{10.16}$$

$$V[x_{t+1}|y_t] = V[ax_t + w_{t+1}|y_t] = a^2V[x_t|y_t] + V[w_{t+1}] = a^2\frac{1}{\sigma_x^{-2} + \sigma_v^{-2}} + \sigma_w^2 \tag{10.17}$$

Recursive Procedure

We have now that the forecast x_{t+1} with its distribution becomes the new prior.

Iterate now the filtering and forecasting steps.

What about convergence?

- The variance of the forecast only depends on $(\sigma_x^2, \sigma_v^2, \sigma_w^2)$.
- We will never figure out x for sure.
- However, the variance converges provided $|a| < 1$.
- The limit is given by σ_w^2.

The estimate of the state x successively incorporates all the information.

10.1.3 The General Case

The general dynamics for the Kalman Filter are described by

$$\hat{x}_{t+1} = A\hat{x}_t + K_{\Sigma_t}(y_t - G\hat{x}_t) \tag{10.18}$$

$$\Sigma_{t+1} = A\Sigma_t A' - K_{\Sigma_t} G\Sigma_t A' + Q \tag{10.19}$$

where

$$K_{\Sigma_t} = A\Sigma G'(G\Sigma G' + R)^{-1} \tag{10.20}$$

is the Kalman gain.

Note that the convergence properties for the variance depends once again on the eigenvalues of the matrix A. A sufficient condition is once again that they are all less than 1 in modulus.

10.2 Bayesian Estimation of Parameters

An alternative to calibration is to use time series data to estimate the deep parameters of the model.

The idea here is that instead of starting out with a single value we start out with a prior distribution on these parameters, use actual data in the model to derive posterior estimates for the parameters.

10.2.1 Example

To get the idea, we look at a very simple example from statistics.

Suppose we have a simple coin, but we do not know the probability of getting heads or tails.

We can observe data of n coin tosses.

Question: What can we infer about the probability from these coin tosses?

We know that the coin generates data according to a binomial distribution given by

$$g(y|\theta) = \begin{pmatrix} n \\ y \end{pmatrix} \theta^y (1-\theta)^{n-y} \tag{10.21}$$

This is our model. We know what the likelihood is observing data and assuming a particular parameter value.

Question: How can we estimate the unknown parameter?

Assume now that we have an initial belief (prior distribution) about the probability being distributed according to a β distribution so that

$$h(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1}(1-\theta)^{\beta-1} \tag{10.22}$$

We use Bayes' Theorem to get

$$k(\theta|y) = \frac{k(y,\theta)}{k_1(y)} = \frac{g(y|\theta)h(\theta)}{\int k(y,\theta)d\theta} = \frac{g(y|\theta)h(\theta)}{\int g(y|\theta)h(\theta)d\theta} \tag{10.23}$$

The distribution of $k(\theta|y)$ is our updated belief (posterior distribution) about the probability having observed the data y.

It depends on the joint distribution $k(\theta, y)$ and is scaled by the marginal distribution of y, $k_1(y)$.

If we get another observation of n coin flips, we can use this posterior as our prior and update our beliefs again, and so forth.

For our example, it turns out that the posterior $k(\theta|y)$ is again a beta distribution, but now with updated parameters $(y + \alpha, n - y + \beta)$ instead of (α, β). This is very convenient!

Question: What should we use for our estimate?

One can either pick the posterior mean or the posterior mode depending on whether one would like to minimize the mean squared error or the absolute value of the error.

In the former case, we would have that

$$E(\theta|y) = \frac{\alpha + y}{\alpha + n + \beta} \tag{10.24}$$

Question: How can we think about this example in terms of our DSGE model?

- The prior distribution concerns all the parameters we would like to estimate.
- The state space representation gives us our model.
- The likelihood of our model will be derived by using the Kalman Filter or some simulation technique.
- The posterior distribution, however, can not be analytically analyzed.

10.2.2 Procedure

Basic Idea

1) Draw different parameter values from the initial distribution.

2) Use data and the solution of the model to determine the likelihood function of the model.

3) Use the likelihood function to generate a posterior distribution of the model.

4) Use the posterior to find estimates of the parameters and the impulse response functions.

We can rely on linear state-space systems and Kalman Filtering to obtain the likelihood of the model when using first-order approximations and normally distributed shocks.[1]

Problem: The number of observables needs to be smaller or equal to the number of shocks in the model. Why? Extra observables would be simply deterministic functions of the others ones.[2]

Problem: The approach gives us only the likelihood of the approximated solution, and not the likelihood of the full model. Also, we lose information associated with higher-order uncertainty such as precautionary behavior.

The Likelihood Function

Consider again our state space system. We can write the likelihood function of our model given a sequence of data $y_{1:T}$ as

$$\begin{aligned} \mathcal{L}(y_{1:T};\theta) &= \prod_{t=1}^{T} \mathcal{L}(y_t|y_{1:t-1};\theta) && (10.25) \\ &= \int \mathcal{L}(y_1|x_0;\theta)dx_0 \prod_{t=1}^{T} \int \mathcal{L}(y_t|y_{1:t-1};\theta)p(x_t|y_{1:t-1};\theta)dx_t && (10.26) \end{aligned}$$

[1]Otherwise other techniques such as the particle filter need to be used.

[2]One could use more observables and assume that there are measurement errors. However, this will lead to identification issues.

This is the probability that the data was generated by the model given parameters θ.

How can we compute the likelihood for given parameters θ?

- $\mathcal{L}(y_t|x_t;\theta)$ follows from the measurement equation given S_t
- $p(x_0;\theta)$ is either the steady state together with the covariance matrix ...
- ... or is specified as draws from a simulation of the model
- $p(x_t|y_{1:t-1};\theta)$ is obtained from the Kalman Filter[3]

The last step is to form the posterior, or

$$p(\theta|y_{1:T}) \propto \mathcal{L}(y_{1:T};\theta)p(\theta) \tag{10.27}$$

where $p(\theta)$ is our prior distribution.

However, we cannot characterize the posterior distribution in closed form. Why? We only have an evaluation of the likelihood, not its functional form.

If we can evaluate it, then we can get point estimates for parameters, posterior distributions for impulse responses to unanticipated shocks and the marginal likelihood of the model among others.

Evaluating the Posterior

One can rely on Markov Chain Monte Carlo (MCMC) simulation to evaluate the posterior.

Can we generate a Markov chain on parameters θ such that its ergodic distribution is given by $p(\theta|y_{1:T})$?[4]

[3]This is not applicable for higher-order approximations. One would then need to use different methods.

[4]In a sense, we are working backwards starting out with the resulting distribution and finding some process that leads us to this distribution.

If so, we can draw from this Markov chain to generate and approximate the distribution of the posterior by the empirical frequency we obtain.

To generate the Markov chain one can rely on the so-called Metropolis-Hastings algorithm.

Step 1: Choose initial values for the parameters θ_0, use $p(\theta_0)$ and $p(y_{1:T}, \theta_0)$ to evaluate the posterior.

Step 2: Sample a new value for θ_i^* according to $\theta_i^* = \theta_{i-1} + \epsilon$ where $\epsilon \sim \mathcal{N}(0, \Sigma)$. This gives us a density $q(\theta_{i-1}, \theta_i^*)$.

Step 3: Solve again the model for θ_i^* and draw $\chi \sim U(0, 1)$.

If $\xi < \frac{p(\theta_i^*|y_{1:T})q(\theta_{i-1},\theta_i^*)}{p(\theta_{i-1}|y_{1:T})q(\theta_i^*,\theta_{i-1})}$, use θ_i^*.

If not, stay with θ_{i-1}.

Step 4: Go back to Step 2 and repeat.

One can generate several such sequences. These sequences should behave as if they were coming from the posterior $p(\theta|y_{1:T})$. This means that they should be the same across different parts of the sequence and they should look the same across different sequences for the same part of the sequence.

10.2.3 What Priors?

Choosing priors depends mainly on three considerations:

- limits on parameter values
- efficient use of available information
- priors such that the posterior distribution maintains the prior distribution (conjugate)

A few pointers.

- Beta distribution on $[0, 1]$: a probability of a binary variable
- Gamma distribution on $[0, \infty)$: arrival rate for a Poisson process
- Normal distribution: for the mean of a normally distributed variable
- Inverse Gamma distribution: for the variance of a normally distributed variable
- Pareto distribution: for a uniformly distributed variable
- Uninformative (improper) prior

10.2.4 Implementation in `DYNARE`

`DYNARE` does – once again – all the work.

It generates posterior distributions using MCMC methods with the Metropolis-Hastings method.

One can specify all the parameters of the algorithm.

There is also output for checking convergence of the MCMC.

finally, one can obtain impulse response functions with confidence intervals based on the posterior distributions of parameters.

10.3 Application: Smets and Wouters (2003) & (2007)

10.4 Literature

Sargent & Stachurski (2017) – Section on Kalman Filter

Smets & Wouters (2007) – AER

Villaverde, Gueroon-Quintana, Rubio-Ramirez (2009) – The New Macroeconometrics: A Bayesian Approach

Villaverde (2010) – The Econometrics of DSGE Models

Part III

Taxes

Chapter 11

Ricardian Equivalence

11.1 Model

Government:

- builds useless pyramids: g_t
- lump-sum taxes: τ_t
- borrowing: B_t (B_0 given)
- feasible policy satisfies

$$(1 + r_t)B_t + g_t = \tau_t + B_{t+1} \tag{11.1}$$

Households:

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

subject to

$$c_t + b_{t+1} \leq y_t - \tau_t + (1 + r_t)b_t$$

$$b_t \geq \underline{B}$$

$$b_0 \text{ given.}$$

Definition 11.1.1. *An equilibrium for a given feasible policy* $\{\tau_t, g_t, B_{t+1}\}_{t=0}^{\infty}$ *and an initial debt level* B_0 *is given by an allocation* $\{c_t, b_{t+1}\}_{t=0}^{\infty}$ *and interest rates* $\{r_t\}_{t=0}^{\infty}$ *such that*

1. *Given interest rates and lump-sum taxes, the allocation solves the household problem for* $b_0 = B_0$*.*

2. *Markets clear*

$$b_{t+1} = B_{t+1} \quad \text{for all } t \tag{11.2}$$

$$c_t + g_t = y_t \quad \text{for all } t. \tag{11.3}$$

11.2 Main Result

Suppose the sequence of government expenditures $\{g_t\}_{t=0}^{\infty}$ is fixed. The timing of financing these expenditures by lump-sum taxes is irrelevant for the real allocation in the economy. In other words, the type of financing (borrowing vs. taxation) does not matter.

Proposition 11.2.1. *(Ricardian Equivalence) Let* $\{c_t^*, b_{t+1}^*\}_{t=0}^{\infty}$ *and* $\{r_t^*\}_{t=0}^{\infty}$ *be an equilibrium for a feasible tax policy* $\{\tau_t^*, g_t^*, B_{t+1}^*\}_{t=0}^{\infty}$*. Then* $\{c_t^*, \hat{b}_{t+1}\}_{t=0}^{\infty}$*, where* $\hat{b}_{t+1} = \hat{B}_{t+1}$*, and* $\{r_t^*\}_{t=0}^{\infty}$ *is an equilibrium for any feasible tax policy* $\{\hat{\tau}_t, g_t^*, \hat{B}_{t+1}\}_{t=0}^{\infty}$*.*

11.2.1 Argument

The idea is to show that the household's FONC and budget constraint do not depend on how the government finances its expenditure.

Step 1:

FONC

$$\beta^t u'(c_t) = \lambda_t \tag{11.4}$$

$$(1 + r_{t+1})\lambda_{t+1} - \lambda_t = 0 \tag{11.5}$$

$$\lim_{T\to\infty} \lambda_T b_{T+1} = 0 \tag{11.6}$$

—▷ Feasibility requires $c_t + g_t = y_t$ in equilibrium.

Result:

Equilibrium interest rates are independent of lump-sum taxes and borrowing and depend only on government expenditure.

Step 2:

The agent's TVC implies

$$\begin{aligned}
\lim_{T\to\infty} \lambda_T b_{T+1} &= \lim_{T\to\infty} \lambda_t \prod_{j=0}^{T-t} \frac{1}{(1+r_{t+j})} b_{T+1} \\
&= u'(c_t)\beta^t \lim_{T\to\infty} \left(\prod_{j=0}^{T-t} \frac{1}{(1+r_{t+j})} \right) b_{T+1} \\
&= 0
\end{aligned} \tag{11.7}$$

—▷ This implies a "No-Ponzi-Game" condition[1] on the government.

Result:

In equilibrium,

$$\lim_{T\to\infty} \left(\prod_{j=0}^{T-t} \frac{1}{(1+r_{t+j})} \right) b_{T+1} = \lim_{T\to\infty} \left(\prod_{j=0}^{T-t} \frac{1}{(1+r_{t+j})} \right) B_{T+1} = 0 \tag{11.8}$$

—▷ The government cannot hold positive wealth "at infinity" (in net present value terms), as this implies a negative wealth position for the household "at infinity".

—▷ The government cannot have debt "at infinity" (in net present value terms), as this implies that the household has some wealth left over "at infinity".

Step 3:

[1] There is a crucial difference between TVC and No-Ponzi-game conditions. The former is a necessary condition for a solution to an optimization problem in order to prevent the overaccumulation of wealth. The latter is an (externally) imposed constraint on a problem in order to prevent overaccumulation of debt. Also, notice that TVC are always an equality representing some form of boundary conditions on the solution of the problem, while No-Ponzi are inequalities restricting the choice.

PV budget constraint for the government at any t is derived as follows

$$B_t = [(\tau_t - g_t) + B_{t+1}] \frac{1}{(1+r_t)} \tag{11.9}$$

$$B_t = \left[\sum_{j=0}^{T-t} (\tau_{t+j} - g_{t+j}) \left(\prod_{k=t}^{t+j} \frac{1}{(1+r_k)} \right) \right] + B_{T+1} \left(\prod_{j=0}^{T-t} \frac{1}{(1+r_{t+j})} \right) \tag{11.10}$$

$$B_t = \left[\sum_{j=0}^{\infty} (\tau_{t+j} - g_{t+j}) \left(\prod_{k=t}^{t+j} \frac{1}{(1+r_k)} \right) \right] \tag{11.11}$$

where the last step follows from above.

—▷ Using the TVC, the PV budget constraint for the household at any t is given by

$$c_t + \tau_t + \sum_{j=1}^{\infty} \left(\prod_{k=t+1}^{t+j} \frac{1}{(1+r_k)} \right) (c_{t+j} + \tau_{t+j}) \leq y_t + \sum_{j=1}^{\infty} \left(\prod_{k=t+1}^{t+j} \frac{1}{(1+r_k)} \right) y_{t+j} + (1+r_t) b_t. \tag{11.12}$$

—▷ Using $b_t = B_t$, in equilibrium

$$c_t + \sum_{j=1}^{\infty} \left(\prod_{k=t+1}^{t+j} \frac{1}{(1+r_k)} \right) c_{t+j} \leq y_t - g_t + \sum_{j=1}^{\infty} \left(\prod_{k=t+1}^{t+j} \frac{1}{(1+r_k)} \right) (y_{t+j} - g_{t+j}). \tag{11.13}$$

Result:

The household's budget constraint only depends on the PV of government spending, but not on the sequence of taxes and government borrowing.

11.3 Further Remarks

—▷ Ricardian equivalence fails whenever prices are influenced by changes in policy and/or the NPV of the agent's budget constraint changes with policy changes.

- This can be the case with incomplete markets.
- Example I: Borrowing constraints that are binding.
- Example II: Government debt completes the market.

—▷ Ricardian Equivalence depends crucially on the assumption of lump-sum taxation. With distortionary taxation it will fail – unless taxes are *history-dependent* (see Basetto and Kocherlakota (2004)). For example when labour taxes are changed in some period – and, hence, the surplus –, the NPV of the budget constraint can be left unchanged if one can tax labour input for that period at a later stage.

—▷ With uncertainty, the results hold as long as one uses the appropriate stochastic discount factor/state price process for evaluating the government's intertemporal budget constraint and its No-Ponzi condition. Again, as long as markets are complete, this is independent of the debt instrument the government chooses; i.e., there is no difference between risky vs. risk-free or long-term vs. short-term debt.

—▷ There is a general question whether a-priori the government has to satisfy an infinite-horizon budget constraint or only in equilibrium. If the second holds, the government by choosing its policy implicitly determines the initial price level such that an infinite-horizon budget constraint holds in equilibrium. This is called the "Fiscal Theory of the Price Level" (for details see Kocherlakota and Phelan (1999)).

11.4 Literature

S+L, Ch. 10 and Ch. 13.10

Basetto and Kocherlakota (2004)

Kocherlakota and Phelan (1999)

See also the work by Leeper and Cochrane

Chapter 12

Long-run Effects of Fiscal Policies

Idea:

Take a standard Neoclassical Growth Model and look at the effects of distortionary taxes on permanent deviations in output levels from trend growth.[1]

12.1 Model

Households:

$$\max_{\{c_t,n_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(c_t, 1-n_t) \tag{12.1}$$

subject to

$$(1+\tau_{ct})c_t + (1+\tau_{xt})x_t \leq (1-\tau_{kt})r_t k_t + (1-\tau_{nt})w_t n_t + T_t \tag{12.2}$$

$$k_{t+1} = (1-\delta)k_t + x_t \tag{12.3}$$

$$k_0 \text{ given} \tag{12.4}$$

$$c_t, k_{t+1} \geq 0, n_t \in [0,1] \tag{12.5}$$

Technology:

[1] Other questions ask whether taxes influence growth rates or the economy in the short-run around the steady-state.

—▷ production function

$$y_t = F(k_t, n_t) \tag{12.6}$$

—▷ standard assumptions (e.g. Cobb-Douglass)

—▷ constant returns to scale implies zero profits in equilibrium

Government:

—▷ policy $\{z_t\}_{t=0}^{\infty} = \{g_t, \tau_{ct}, \tau_{xt}, \tau_{kt}, \tau_{nt}, T_t\}$

—▷ policy is feasible if it satisfies a flow budget constraint

$$g_t = \tau_{ct}c_t + \tau_{xt}x_t + \tau_{kt}r_tk_t + \tau_{nt}w_tn_t - T_t \tag{12.7}$$

—▷ households do not derive utility from government expenditure

12.2 Equilibrium

Definition 12.2.1. *A competitive equilibrium given a feasible government policy* $\{g_t, \tau_{ct}, \tau_{xt}, \tau_{kt}, \tau_{nt}, T\}$ *is an allocation* $\{(c_t, n_t, k_{t+1})\}_{t=0}^{\infty}$ *and prices* $\{(r_t, w_t)\}$ *such that*

1. *given the policy, prices and* k_0*, the allocation solves the household's problem*
2. *given prices,* $\{n_t, k_t\}$ *solves the firm's problem*
3. *markets clear, i.e.*

$$c_t + g_t + k_{t+1} - (1-\delta)k_t = F(k_t, n_t) \quad \textit{for all } t. \tag{12.8}$$

Equilibrium conditions

—▷ Firms:

$$r_t = F_k(k_t, n_t) \tag{12.9}$$

$$w_t = F_n(k_t, n_t) \tag{12.10}$$

—▷ Households:

$$\beta^t u_c(c_t, 1-n_t) - \lambda_t(1+\tau_{ct}) = 0 \tag{12.11}$$

$$\beta^t u_n(c_t, 1-n_t) - \lambda_t w_t(1-\tau_{nt}) = 0 \tag{12.12}$$

$$-\lambda_t(1+\tau_{xt}) + \lambda_{t+1}\left[(1-\delta)(1+\tau_{xt+1}) + r_{t+1}(1-\tau_{kt+1})\right] = 0 \tag{12.13}$$

$$\lim_{T\to\infty} \lambda_T(1+\tau_{xt})k_{T+1} = 0 \tag{12.14}$$

—▷ Market Clearing

12.3 Effects

12.3.1 Tax Wedges

There are two *tax wedges*:

- intratemporal

$$\frac{(1-\tau_{nt})}{(1+\tau_{ct})} = \frac{u_n(c_t, 1-n_t)}{u_c(c_t, 1-n_t)F_n(k_t, n_t)} \tag{12.15}$$

- intertemporal

$$\frac{u_c(c_t, 1-n_t)}{\beta u_c(c_{t+1}, 1-n_{t+1})} = \frac{(1+\tau_{ct})}{(1+\tau_{ct+1})}\left[(1-\delta)\frac{(1+\tau_{xt+1})}{(1+\tau_{xt})} + F_k(k_{t+1}, n_{t+1})\frac{(1-\tau_{kt+1})}{(1+\tau_{xt})}\right] \tag{12.16}$$

Remark: If there is no tax/subsidy on investment and taxes are constant, we obtain for the intertemporal wedge the much easier expression

$$\left[(1-\delta) + F_k(k_{t+1}, n_{t+1})(1-\tau_k)\right]. \tag{12.17}$$

<u>General Idea:</u>

—▷ These wedges can be important to understand relative levels of income/output across countries with similar characteristics.

—▷ Remove trend growth from data (relative to a benchmark). Then, look at relative performance of countries and their tax policies (see e.g. Prescott (2002)). Hence, equilibrium levels relative to a benchmark matter.[2]

—▷ One can use data to assess which factors cause the difference in performance (e.g. taxes vs. "detrended TFP")

12.3.2 Steady State

The steady state is given by the solution (c^{SS}, n^{SS}, k^{SS}) to

$$1 = \beta \left[(1-\delta) + \frac{(1-\tau_k)}{(1+\tau_x)} F_k(k^{SS}, n^{SS}) \right] \tag{12.18}$$
$$\frac{u_c(c^{SS}, 1-n^{SS})}{u_n(c^{SS}, 1-n^{SS})} = \frac{(1-\tau_n)}{(1+\tau_c)} F_n(k^{SS}, n^{SS}) \tag{12.19}$$
$$g + c^{SS} + \delta k^{SS} = F(k^{SS}, n^{SS}) \tag{12.20}$$

Result:

—▷ Suppose $u(c, 1-n) = u(c)$, i.e. labour is inelastically supplied. Then, $\tau_c \neq 0$ does not influence the steady-state value of capital.

—▷ Why? Taxing consumption is non-distortionary if labour is inelastically supplied.

—▷ Hence, $\tau_x = \tau_k = 0$ is optimal and k^{SS} is first-best.

12.3.3 Transition

Dynamics are described by a second-order difference equation in capital k treating the exogenous[3] variables z as given parameters: $H(k_t, k_{t+1}, k_{t+2}; z_t, z_{t+1}) = 0$.

Question: How do we compute the equilibrium path given a sequence of tax policies?

[2]This is in contrast to RBC studies that analyze the effects of policy stabilizing the economy in response to shocks. There, statistical properties of time-series of variables matter.

[3]These are in general variables describing government policy as well as exogenous shocks.

Shooting-Algorithm:

1. Fix government policy z. Assume there exists T such that $z_t =$ const. for all $t \geq T$.

2. Find steady state given z_T.

3. Let $T_{max} >> T$. Assume that in T_{max} we have reached a steady state.

    - Guess c_0 and use k_0 to find n_0 and k_1 from the FONC in $t = 0$ and from market clearing.

    - Use k_1 to find (c_1, n_1) from the Euler equation and the FONC in $t = 1$.

    - Iterate forward until period T_{max}.

4. Calculate $k_{T_{max}} - k^{SS} = \epsilon$.

5. If $\epsilon > 0 (< 0)$, increase (decrease) c_0. Go back to Step (3). If $\epsilon \simeq 0$, stop.

Remark:

For a different algorithm to calculate the equilibrium in an economy with tax distortions, see Coleman (1991). The algorithm is based on iterating directly on the intertemporal Euler equation to obtain the law of motion for capital.

Remark:

`DYNARE` can solve for transitions. One needs to define an initial and a final steady state together with shocks and a length of transition. Try it!

12.4 Literature

Sargent and Ljunqvist, Ch. 11

Cooley, Ch. 2 and Ch. 3

Coleman, Econometrica (1991)

Chapter 13

Optimal Taxation under Commitment

13.1 Preliminaries

Question:

What is the optimal (i.e., welfare maximizing) tax policy?

Themes:

- long-run distortionary taxes on productive factors
- tax smoothing vs. response to cyclical fluctuations
- importance of commitment
- importance of full information for the government (New Public Finance)

"Ramsey"-Problem:

- *Primal* approach: eliminate all prices and choose equilibrium allocation directly (Ramsey allocation problem)
- *Dual* approach: choose taxes and, hence, the after-tax factor prices as well as interest rates on debt

13.2 Model

Technology:

- CRS: $F(k(s^t), l(s^t), s_t)$
- factor prices: $r(s^t)$ and $w(s^t)$

Government:

- expenditure $\{g(s^t)\}_{t=0}^{\infty}$ exogenously given
- chooses policy $\pi = \{\pi(s^t)\}_{t=0}^{\infty} = \{\tau_l(s^t), \tau_k(s^t), (1 + r_b(s^t))\}_{t=0}^{\infty}$
- flow budget constraint:

$$g(s^t) + (1 + r_b(s^t))b(s^{t-1}) = \tau_l(s^t)w(s^t)l(s^t) + \tau_k(s^t)[r(s^t) - \delta]k(s^{t-1}) + b(s^t) \quad (13.1)$$

- expenditure and choice of policy pins down the sequence of debt levels

Households:

$$\max_{\{c(s^t),b(s^t),k(s^t),l(s^t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \mu(s^t) u(c(s^t), 1 - l(s^t)) \quad (13.2)$$

subject to

$$c(s^t) + k(s^t) + b(s^t) \leq k(s^{t-1}) + (1 - \tau_k(s^t))[r(s^t) - \delta]k(s^{t-1}) +$$
$$+(1 - \tau_l(s^t))w(s^t)l(s^t) + (1 + r_b(s^t))b(s^{t-1}) \quad (13.3)$$

$$(k_{-1}, b{-}1) \text{ given} \quad (13.4)$$

$$c(s^t), k(s^t) \geq 0, l(s^t) \in [0, 1] \quad (13.5)$$

—▷ plus some No-Ponzi-Game condition

—▷ households are identical here

—▷ hence: absence of private contingent claims is irrelevant

13.3 Definition of Ramsey Equilibrium

- Allocation and Price rules: given policy π an allocation $x = (c, b, k, l)$ and prices (r, w) are realized in equilibrium
- $x(\pi) = \{x(s^t|\pi)\}_{t=0}^{\infty}$, $r(\pi) = \{r(s^t|\pi)\}_{t=0}^{\infty}$, $w(\pi) = \{w(s^t|\pi)\}_{t=0}^{\infty}$
- k_{-1} and b_{-1} are given (i.e. inelastically "supplied")
- hence: we take the tax rate and the interest rate on these variables as given

Definition 13.3.1. *A Ramsey Equilibrium for a given initial tax rate on capital $\tau_k(s_0)$ and a given initial interest rate $r_b(s_0)$ is a policy π and allocation and price rules (x, w, r) such that*

1. *π maximizes the household's utility subject to the government's flow budget constraint*
2. *for all policies π', $x(\pi')$ solves the household's problem taking π', $r(\pi')$ and $w(\pi')$ as given*
3. *for all policies π',*

$$w(s^t|\pi') = F_l(k(s^t|\pi'), n(s^t|\pi'), s^t)$$
$$r(s^t|\pi') = F_k(k(s^t|\pi'), n(s^t|\pi'), s^t)$$

for all s^t, for all t.

—▷ The second and third condition impose that allocation and price rules form competitive equilibria.

—▷ The planner chooses then a *second-best* being restricted by the equilibrium choices of the agents.

Remark: We impose that households and firms behave optimally *for all* possible government policies. Suppose this was not the case. Then, take any policy and find an allocation and price rule that is a competitive equilibrium given this policy. The policy is optimal for an allocation rule that specifies that people supply zero labour for all policies except for the one policy we look at. Obviously, under appropriate assumptions on u, zero labour supply cannot be an equilibrium except for a 100% tax on labor income.

13.4 Long-run Capital Taxes should be Zero

We will use the Dual Problem.

- no uncertainty and $\{g_t\}_{t=0}^{\infty}$ given
- for simplicity, we assume $\tau_{ct} = 0$
- government chooses new factor prices

$$\bar{r}_t = (1-\tau_{kt})(r_t - \delta) \tag{13.6}$$

$$\bar{w}_t = (1-\tau_{lt})w_t \tag{13.7}$$

- equilibrium characterized by

$$\frac{u_l(c_t, 1-l_t)}{u_c(c_t, 1-l_t)} = \bar{w}_t \tag{13.8}$$

$$\lambda_t = \lambda_{t+1}(1+\bar{r}_{t+1}) = \lambda_{t+1}(1+r_{bt+1}) \tag{13.9}$$

$$F(k_t, n_t) = g_t + c_t + k_{t+1} - (1-\delta)k_t \tag{13.10}$$

- in equilibrium, we have:

$$\begin{aligned} g_t + (1+r_{bt})b_{t-1} &= \tau_{lt}w_t l_t + \tau_{kt}(r_t - \delta)k_t + b_t \\ &= (w_t - \bar{w}_t)l_t + (r_t - \delta - \bar{r}_t)k_t + b_t \\ &= F_l(k_t, l_t)l_t + F_k(k_t, l_t)k_t - \bar{w}_t l_t - \bar{r}_t k_t - \delta k_t + b_t \\ &= F(k_t, l_t) - \bar{w}_t l_t - \bar{r}_t k_t - \delta k_t + b_t \end{aligned}$$

Ramsey Problem:

$$\max_{\{\bar{r}_t,\bar{w}_t,c_t,l_t,k_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(c_t, 1-l_t) \tag{13.11}$$

$$\text{subject to} \tag{13.12}$$

$$F(k_t,l_t) - \bar{w}_t l_t - \bar{r}_t k_t - \delta k_t + b_t = g_t + (1+\bar{r}_t)b_{t-1} \tag{13.13}$$

$$F(k_t,l_t) = g_t + c_t + k_{t+1} - (1-\delta)k_t \tag{13.14}$$

$$\frac{u_l(c_t,1-l_t)}{u_c(c_t,1-l_t)} = \bar{w}_t \tag{13.15}$$

$$u_c(c_t,1-l_t) = \beta u_c(c_{t+1},1-l_{t+1})(1+\bar{r}_{t+1}) \tag{13.16}$$

where the first constraint is the government flow budget constraint taking into account equilibrium conditions for the stand-in firm.[1]

FONC (w.r.t. k_{t+1}):

$$\nu_t = \beta\big\{\nu_{t+1}\left[F_k(k_{t+1},l_{t+1}) + (1-\delta)\right] + \psi_{t+1}\left[F_k(k_{t+1},l_{t+1}) - \bar{r}_{t+1} - \delta\right]\big\} \tag{13.17}$$

—▷ $\nu_t > 0$ is the Lagrange multiplier on the resource constraint at t

—▷ $\psi_t \geq 0$ is the Lagrange multiplier on the government budget constraint at t

—▷ hence: MC of capital in t = MV of capital in $t+1$

—▷ second term consists of (i) increase in resources and (ii) decrease in tax burden

Use intertemporal Euler-equation for household ($1 = \beta(1+\bar{r})$) to obtain in SS

$$\begin{aligned} \nu\left(\frac{1}{\beta} - 1\right) &= \nu(r-\delta) + \psi(r-\delta-\bar{r}) \\ \nu\bar{r} &= \nu(r-\delta) + \psi(r-\delta-\bar{r}) \\ 0 &= (\nu+\psi)[(r-\delta)-\bar{r}]. \end{aligned} \tag{13.18}$$

[1] It seems one would also have to include the household's budget constraint. This is redundant, however, as it is always fulfilled when the market clearing condition and the government's budget constraints hold.

Proposition 13.4.1. *(Chamley (1986)) Suppose the economy converges to a steady-state for any given government policy. Then, it is optimal to set* $\tau_k^{SS} = 0$.

Remarks:

1. Judd (1985) shows that is never optimal to tax capital in steady-state in order to redistribute wealth among agents.
2. Note that this result applies only for the steady state. It can be optimal to tax capital over the transition to the steady state.

13.5 Primal Approach to the Ramsey Problem

13.5.1 Ramsey Allocation Problem

Optimization Problem for the Planner:

$$\max_{\{c(s^t),l(s^t),k(s^t)\}} \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \mu(s^t) u(c(s^t), 1 - l(s^t)) \tag{13.19}$$

subject to

$$c(s^t) + g(s^t) + k(s^t) = F(k(s^{t-1}), l(s^t), s^t) + (1 - \delta)k(s^{t-1}) \tag{13.20}$$

$$\sum_{t=0}^{\infty} \sum_{s^t} \beta^t \mu(s^t) \left[u_c(s^t)c(s^t) + u_l(s^t)l(s^t)\right] = \tag{13.21}$$

$$u_c(s_0) \left[k_{-1} + (1 - \tau_{k0})(r(s^t) - \delta)k_{-1} + (1 + r_{b0})b_{-1}\right] \tag{13.22}$$

—▷ The last condition is an "implementability" condition.

—▷ It is the household's intertemporal budget constraint...

—▷ ...and can be obtained by using the FONC to eliminate prices.[2]

[2]Again, the government's budget constraint is implied by the resource constraint and the household's budget constraint. The TA will derive this equation.

Proposition 13.5.1. *The allocation (c, k, l) associated with the policy of a Ramsey equilibrium solves the Ramsey Allocation Problem.*

Proof. We show that all restrictions imposed in the definition of the Ramsey equilibrium can be summarized by the two constraints in the Ramsey allocation problem.

Step 1: Add the feasibility constraint and the government flow budget constraint to obtain the household's budget constraint. This uses CRS in production. This shows that we can use any 2 of these three constraints.

Step 2: Describe all necessary conditions for the household's problem given a policy π. These are

$$\beta^t \mu(s^t) u_c(s^t) = \lambda(s^t) \tag{13.23}$$

$$\beta^t \mu(s^t) u_l(s^t) = -\lambda(s^t)(1 - \tau_l(s^t)) w_t \tag{13.24}$$

$$\lambda(s^t) b(s^t) - \left[\sum_{s_{t+1}|s^t} \lambda(s^{t+1})(1 + r_b(s^{t+1})) \right] b(s^t) = 0 \tag{13.25}$$

$$\lambda(s^t) k(s^t) - \left[\sum_{s_{t+1}|s^t} \lambda(s^{t+1})(1 + (1 - \tau_k(s^t))(r(s^t) - \delta)) \right] k(s^t) = 0 \tag{13.26}$$

$$\lim_{t \to \infty} \lambda(s^t) b(s^t) = 0 \tag{13.27}$$

$$\lim_{t \to \infty} \lambda(s^t) k(s^t) = 0. \tag{13.28}$$

Multiply the budget constraint by $\lambda(s^t)$ for all s^t and sum until infinity to obtain a life-time budget constraint for the household. Using the definition of $w(s^t)$, $r(s^t)$ and $\lambda(s^t)$ from the FONC of firms and households plus the no-arbitrage condition on bonds and capital, one obtains the "implementability condition".

Step 3: Take any solution to the Ramsey Allocation Problem. Then, it satisfies feasibility and the "implementability condition". One can construct Lagrange multipliers, bond and capital holdings so that the solution satisfies the FONC for a competitive equilibrium as described above and the sequential budget constraints. □

13.5.2 Necessary Conditions

Goal: decompose the problem into an initial period and a dynamic part that (under certain conditions) evolves endogenously as a Markov process

—▷ Incorporate the implementability constraint into the objective function

$$W(c(s^t), l(s^t), \nu) = u(c(s^t), l(s^t)) + \nu \left[u_c(s^t)c(s^t) + u_l(s^t)l(s^t)\right] \tag{13.29}$$

Ramsey allocation problem in Lagrangian form:

$$\begin{aligned}\max_{\{c(s^t),l(s^t),k(s^t)\}} \sum_{t=0}^{\infty}\sum_{s^t} &\beta^t \mu(s^t) W(c(s^t), l(s^t), \nu) + \\ &- \nu u_c(s_0)\left[k_{-1}(1 + (1-\tau_{k0})(r(s_0)-\delta) + b_{-1}(1+r_{b0})\right] \\ &+ \nu(s^t)\left[F(k(s^{t-1}), l(s^t), s^t) + (1-\delta)k(s^{t-1}) - c(s^t) + g(s^t) + k(s^t)\right]\end{aligned} \tag{13.30}$$

FONC for $t \geq 1$:

$$-\frac{W_l(s^t)}{W_c(s^t)} = F_l(s^t) \tag{13.31}$$

$$W_c(s^t) = \sum_{s_{t+1}|s^t} \beta\mu(s_{t+1}|s^t) W_c((s_{t+1}, s^t))\left[1 - \delta + F_k((s_{t+1}, s^t))\right] \tag{13.32}$$

FONC for $t = 0$:

$$\begin{aligned}-\frac{W_l(s_0) - \nu\big(u_{cl}(s_0)\left[k_{-1}(1+(1-\tau_{k0})(r(s_0)-\delta) + b_{-1}(1+r_{b0})\right] - u_c(s_0)[1-\tau_{k0}]F_{kl}(s_0)\big)}{W_c(s_0) - \nu u_{cc}\left[k_{-1}(1+(1-\tau_{k0})(r(s_0)-\delta) + b_{-1}(1+r_{b0})\right]} \\ = F_l(s_0)\end{aligned} \tag{13.33}$$

$$\begin{aligned}W_c(s_0) - \nu u_{cc}(s_0)&\left[k_{-1}(1+(1-\tau_{k0})(r(s_0)-\delta) + b_{-1}(1+r_{b0})\right] \\ &= \sum_{s_1|s_0} \beta\mu(s_1|s_0) W_c((s_1, s_0))\left[1 - \delta + F_k((s_1, s_0))\right]\end{aligned} \tag{13.34}$$

13.5.3 Solving for a Ramsey Equilibrium

—▷ The above system of 4 equations describes the necessary conditions for a Ramsey equilibrium given by $(c(s^t), l(s^t), k(s^t))$ and ν. The Ramsey equilibrium allocation together with the multiplier ν has to satisfy these equations, the feasibility constraints and the implementability constraint.

—▷ To compute an equilibrium, take ν as fixed and solve the system of equations without the implementability constraint. Then check wether the implementability constraint is satisfied. If it is not, adjust the multiplier (or "price") ν accordingly.

—▷ Given the Ramsey Allocations, we can use the household's and the firm's equilibrium conditions to easily recover prices and tax policies $(r, w, \tau_l, \tau_k, r_b)$.

13.6 Ramsey Policies

13.6.1 Indeterminancy of Capital Taxes

—▷ Ramsey policies need to implement an equilibrium. We take the Ramsey allocation as given.

—▷ static Euler equation gives $\tau_l(s^t)$ for all s^t

$$-\frac{u_l(s^t)}{u_c(s^t)} = (1-\tau_l(s^t))F_l(s^t) \tag{13.35}$$

—▷ $r_b(s^t)$ and $\tau_k(s^t)$ are described by intertemporal Euler equation and the budget constraints

$$u_c(s^t) = \sum_{s_{t+1}|s^t} \beta\mu(s_{t+1}|s^t)u_c(s_{t+1}|s^t)(1+r_b(s_{t+1}|s^t)) \tag{13.36}$$

$$u_c(s^t) = \sum_{s_{t+1}|s^t} \beta\mu(s_{t+1}|s^t)u_c(s_{t+1}|s^t)\big(1+(1-\tau_k(s_{t+1}|s^t))(F_k(s_{t+1}|s^t)-\delta)\big) \tag{13.37}$$

$$\begin{aligned} c(s_{t+1}|s^t) + b(s_{t+1}|s^t) + k(s_{t+1}|s^t)) &= (1-\tau_l(s_{t+1}|s^t))F_l(s_{t+1}|s^t)l(s_{t+1}|s^t) \\ &+(1+r_b(s_{t+1}|s^t))b(s^t) + \big(1+(1-\tau_k(s_{t+1}|s^t))(F_k(s_{t+1}|s^t)-\delta)\big)k(s^t) \end{aligned} \tag{13.38}$$

—▷ If we have N states, there are $2N$ variables to be determined, but we have only $N+2$ equations given today's state is s^t.

—▷ Accounting for one linear dependency, we have $N-1$ degrees of freedom for implementing the equilibrium.

Intuition:

- need to "design the right asset" for the agents
- have to span N states – or, *state-dependent* net deficit
- need to offer the correct ex-ante returns on two assets, capital and government bonds

Result: There are $N-1$ degrees of freedom (or indeterminancy) to set capital taxes and interest rates on government debt. This implies that *either* capital taxes *or* interest rates can be state-independent.

Result: The **ex-ante tax rate on capital income** is uniquely defined by

$$\tau_k^e(s^t) = \frac{\sum_{s_{t+1}|s^t} q(s_{t+1}|s^t)\tau_k(s_{t+1}|s^t)\left[F_k(s_{t+1}|s^t) - \delta\right]}{\sum_{s_{t+1}|s^t} q(s_{t+1}|s^t)\left[F_k(s_{t+1}|s^t) - \delta\right]} \tag{13.39}$$

where $q(s_{t+1}|s^t)$ is the Arrow-Debreu price of consumption in state $(s_{t+1}|s^t)$.

13.6.2 Optimal Taxation

—▷ We can decompose Ramsey policies in two parts.

Results (Dynamic Part):

- Suppose s_t follows a first-order Markov process. The Ramsey policies can then be described by time-invariant policy rules $\tau_l(k, s|\nu)$, $\tau_k^e(k, s|\nu)$ and the state-dependent net deficit which also follow a first-order Markov process.

- Note that the solution of the dynamic part depends on the parameter ν, which is the multiplier on the implementability constraint.

Results (Initial Period):

- The period $t = 0$ policies are different from the stationary ones.
- The ex-ante tax on capital, τ_k^e, is forward-looking. Hence, capital taxes for period $t = 1$ are different.

Remark:

The initial conditions $(\tau_{k0}, r_{b0}, k_{-1}, b_{-1})$ pin down the value of the multiplier ν, which is interpreted as the utility cost of financing government expenditure through distortionary taxation. Note that τ_{k0} and r_{b0} are exogenously given. If not, one would simply set $\tau_{k0} = 1$ and finance all government expenditures up-front through interest earned on renting out the capital stock.

Why?

—▷ The derivative of the planner's welfare function w.r.t. to τ_{k0} is given by

$$\nu u_c(s_0)(F_k(s_0) - \delta)k_{-1} \geq 0. \tag{13.40}$$

Given any $\nu > 0$, by raising τ_{k0}, the government reduces utility costs from distorting taxation as initial capital is a fixed factor.

—▷ When $\nu = 0$, the government has enough tax revenue from τ_{k0} to finance its future expenditure.

13.6.3 When are Zero Taxes on Capital Optimal?

—▷ We restrict attention to utility functions of the form

$$u(c, 1 - l) = \frac{c^{1-\sigma}}{1 - \sigma} + V(1 - l) \tag{13.41}$$

where $\sigma > 0$ and V is strictly concave.

Proposition 13.6.1. *With the above utility function, the optimal ex-ante tax rate on captial income is given by*

$$\tau_k^e(s^t) = 0 \tag{13.42}$$

for all $t \geq 1$.

Proof. For $t \geq 1$, the intertemporal Euler-equation for the Ramsey problem is given by

$$W_c(s^t) = \sum_{s_{t+1}|s^t} \beta\mu(s_{t+1}|s^t) W_c((s_{t+1}, s^t)) \left[1 - \delta + F_k((s_{t+1}, s^t))\right]. \tag{13.43}$$

With the above utility function, we have that

$$\frac{W_c(s^{t+1})}{W_c(s^t)} = \frac{c^{-\sigma}(s^{t+1})\big[1 + \nu(1-\sigma)\big]}{c^{-\sigma}(s^t)\big[1 + \nu(1-\sigma)\big]} = \frac{u_c(s^{t+1})}{u_c(s^t)}. \tag{13.44}$$

The intertemporal Euler equation for the household is given by

$$u_c(s^t) = \sum_{s_{t+1}|s^t} \beta\mu(s_{t+1}|s^t) u_c(s_{t+1}|s^t)\big(1 + (1 - \tau_k(s_{t+1}|s^t))(F_k(s_{t+1}|s^t) - \delta)\big). \tag{13.45}$$

Hence, subtracting the households intertemporal Euler equation from the one of the Ramsey planner, we obtain

$$\begin{aligned} 0 = \sum_{s_{t+1}|s^t} \beta\mu(s_{t+1}|s^t) \frac{u_c(s^{t+1})}{u_c(s^t)} \tau_k(s_{t+1}|s^t) \left[F_k(s_{t+1}|s^t) - \delta\right] = \\ \sum_{s_{t+1}|s^t} q(s_{t+1}|s^t) \tau_k(s_{t+1}|s^t) \left[F_k(s_{t+1}|s^t) - \delta\right], \end{aligned} \tag{13.46}$$

where $q(s_{t+1}|s^t)$ is the Arrow-Debreu price. This implies that the ex-ante tax rate on capital income is 0 for all $t \geq 1$. □

13.7 Tax Smoothing

13.7.1 Environment

- no capital

- labor is transformed into output 1-1
- government expenditure given by an exogenous stochastic process $\{\{g_t\}_{t=0}^{\infty} | g : S^t \to I\!R\}$
- the probability distribution induced by government spending is given by $\mu(s^t)$
- financed by one-period state-contingent debt $b(s^t)$ and taxes on labor income $\tau_l(s^t)$

In equilibrium, government will be restricted by a present-value budget constraint. Hence, any current government deficits must be eventually financed by taxes.

13.7.2 Ramsey Allocation Problem

$$\max_{\{c(s^t),l(s^t)\}} \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \mu(s^t) u(c(s^t), 1 - l(s^t)) \tag{13.47}$$

subject to

$$c(s^t) + g(s^t) = l(s^t) \tag{13.48}$$

$$\sum_{t=0}^{\infty} \sum_{s^t} \beta^t \mu(s^t) \left[u_c(s^t) c(s^t) + u_l(s^t) l(s^t) \right] = u_c(s_0)(1 + r_{b0}) b_{-1} \tag{13.49}$$

FONC for $t \geq 1$:

$$\begin{aligned} &(1+\nu) u_c(s^t) + \nu \left[c(s^t) u_{cc}(s^t) - (c(s^t) + g(s^t)) u_{lc}(s^t) \right] = \\ &\quad (1+\nu) u_l(s^t) + \nu \left[c(s^t) u_{cl}(s^t) - (c(s^t) + g(s^t)) u_{ll}(s^t) \right] \end{aligned} \tag{13.50}$$

Result:

Ramsey allocation depends only on the *current* government expenditure $g(s^t)$ and, hence, is *history-independent* and *time-independent.*

13.7.3 Some Examples

Constant Government Expenditure

- $g_t = g$ for all t

—▷ FONC and market clearing imply that $c_t = c$ and $l_t = l$ for all t.

—▷ From the household problem we have

$$\frac{u_l(s^t)}{u_c(s^t)} = 1 - \tau_l(s^t). \tag{13.51}$$

Result: $\tau_l = \tau$ for all t.

—▷ Using this FONC from the households labor choice, the implementability constraint is given by

$$\sum_{t=0}^{\infty} \sum_{s^t} \beta^t \mu(s^t) u_c(s^t) \left[\tau_l(s^t) l(s^t) - g(s^t)\right] - u_c(s_0)(1 + r_{b0}) b_{-1} = 0. \tag{13.52}$$

Result:

- If $b_{-1} = 0$, the government's budget is balanced for all t, i.e. $\tau l = g$.
- If $b_{-1} > 0$, the initial outstanding debt plus interest is redeemed over time by taxes exceeding government spending according to

$$\frac{1}{1-\beta} (\tau l - g) = (1 + r_{b0}) b_{-1} > 0. \tag{13.53}$$

<u>A Perfectly Foreseen One-period War</u>

- $g_T > 0$ and $g_t = 0$ for all $t \neq T$.
- Assume: $b_0 = 0$.

—▷ For $t \neq T$, the allocation and taxes are constant.

—▷ From the FONC one can derive the following condition[3]

$$\begin{aligned}(1+\nu)\left[c(s^t)u_c(s^t) - l(s^t)u_l(s^t)\right] + \nu\left[c(s^t)^2 u_{cc}(s^t) - 2l(s^t)c(s^t)u_{lc}(s^t) + l(s^t)^2 u_{ll}(s^t)\right] \\ - \nu(s^t)(c(s^t) - l(s^t)) = 0.\end{aligned} \tag{13.54}$$

[3] See e.g. S+L, p.510.

Result:

- For $t \neq T$, $c = l$. Also, $\nu > 0$ as the government must resort to distortionary taxation.
- One can show that the second term is negative for any u strictly concave. It then follows that $c(s^t)u_c(s^t) - l(s^t)u_l(s^t) > 0$.
- We have then for $t \neq T$ that

$$0 < c + l\frac{u_l}{u_c} = c - l(1-\tau) = \tau l. \tag{13.55}$$

- For $t = T$, $c_T < l_T$ and the above argument does not go through.

—▷ This implies that the government first runs a surplus. It uses the surplus to build up assets by buying bonds from the consumer until $T-1$ (in other words, the consumer borrows from the government). Then, in period T it sells all its bonds and issues more bonds, which it roles over forever at a constant interest rate.

—▷ The household needs to finance its consumption partly by issuing debt to the government until $T-1$. In T it settles all outstanding debt and buys government bonds. Later on, it uses the interest income to finance its tax obligations.

—▷ Why? Except for period T, aside from distortionary taxation there is no direct impact of the government on the economy, so that consumption equals output.

<u>A Chance for a One-Period War</u>

- Same set-up as before except for $g_T > 0$ with probability α and $g_T = 0$ otherwise.

The optimal debt policy is now state-contingent:

- For $t = 0, \ldots, T-2$, the gov't buys bonds from the consumer using its tax revenue.
- At $t = T-1$, the gov't sells all bonds and uses the proceeds plus tax revenue to create the following portfolio of assets:

1. It buys state-contingent bonds from the consumer that only pay a return if $g_T > 0$, i.e. if there is a war.
2. It issues non-contingent bonds. This ensures to perfectly smooth taxes, even if there is no war in period T.

- At T, if there is no war, the gov't pays interest only on the non-contingent bonds. If there is a war, the gov't receives pay-offs from the contingent bonds to finance $g_T > 0$.
- After T, the gov't rolls over debt irrespective of history.

13.8 Literature

Sargent and Ljunqvist, Ch. 15

Chamley, Econometrica (1986)

Chari, Christiano and Kehoe, Journal of Political Economy (1994) and in Cooley (ed.) (1995)

Judd, Journal of Public Economics (1985)

Stokey and Lucas, Journal of Monetary Economics (1983)

Chapter 14

New Public Finance

14.1 A 2-period Moral Hazard Problem

14.1.1 Model

Risk-neutral principal:

- pays a wage w to the agent
- discounts the future according to $\beta = \frac{1}{1+r} = \frac{1}{R}$

Risk-averse agent:

- choses action $a \in A$
- preferences: $u(w) - c(a)$ with discounting according to β
- outcomes: $\{x_1, \ldots, x_N\}$ with probability $\pi_i(a)$
- strategy: $(s_0, s_1, \ldots s_N)$ where s_0 is the period 1 action and s_i is the action taken in period 2 conditional on the observed outcome x_i in period 1

Private Information:

- Principal cannot observe action a.
- Realized Outcome x is publicly observable.

Contract:

- state-contingent path of wages for the agent
- finite N implies a contract is a list of $N + N^2$ wages
- in utility terms, we have $z = \left(\{u(w_i)\}_{i=1}^N, \{u(w_{ij})\}_{i,j=1}^N\right)$

14.1.2 Main Results

Proposition 14.1.1. *The Pareto-optimal contract satisfies*

$$\frac{1}{u'(w_i)} = \sum_{j=1}^{N} \pi_j(s_i) \frac{1}{u'(w_{ij})} \tag{14.1}$$

for all $i = 1, \ldots, N$.

Proof. Consider any contract z and the optimal strategy s given z. Change the contract for only one state i in the first period to

$$\tilde{z}_i = z_i - y \tag{14.2}$$

$$\tilde{z}_{ij} = z_{ij} + \frac{y}{\beta} \text{ for all } j = 1, \ldots, N. \tag{14.3}$$

Then, the original strategy is still optimal, since (i) for $j \neq i$ the contract doesn't change, (ii) the relative pay-offs for ij do not change and (iii) the NPV of the contract remains unchanged.

It must then be the case that the optimal contract minimizes the costs of the principal at $y = 0$. The costs for the principal is given by

$$u^{-1}(z_i + y) + \beta \sum_{j=1}^{N} \pi_j(s_i) u^{-1}\left(z_{ij} - \frac{y}{\beta}\right). \tag{14.4}$$

The FONC needs to be 0 at $y = 0$ which gives

$$\frac{1}{u'(w_i)} = \sum_{j=1}^{N} \pi_j(s_i) \frac{1}{u'(w_{ij})} \tag{14.5}$$

which completes the proof. □

14.1.3 Properties of the Optimal Contract

1. Memory: If $w_i \neq w_j$, then there exist k such that $w_{ik} \neq w_{jk}$.

2. "Martingale Property": If $\frac{1}{u'}$ is convex (concave/linear), $w_i \geq (\leq / =) \sum_{k=1}^{N} \pi_k(s_i) w_{ik}$.

3. The agent is "savings-constrained": Since the agent bears income risk in the 2nd period, he would like to save some of his wage in the first period for additional consumption in the second period.

Proposition 14.1.2. *In the optimal contract, the agent is constrained in his savings.*

Proof. Consider the problem

$$\max_b u(w_i - b) + \beta \sum_{j=1}^{N} \pi_j(s_i) u(w_{ij} + (1+r)b) \tag{14.6}$$

which yields a necessary condition equal to

$$-u'(w_i - b) + \sum_{j=1}^{N} \pi_j(s_i) u'(w_{ij} + (1+r)b) = 0. \tag{14.7}$$

At $b = 0$, this FONC must be positive, since we have from the optimal contract

$$u'(w_i) = \frac{1}{\sum_{j=1}^{N} \frac{\pi_j(s_i)}{u'(w_{ij})}} \leq \sum_{j=1}^{N} \pi_j(s_i) u'(w_{ij}) \tag{14.8}$$

where the last inequality follows from the weighted arithmetic mean being larger than the weighted harmonic mean (by Jensen's inequality).

Hence, with the optimal contract the agent would like to set $b > 0$ if given the opportunity to save. □

14.1.4 Digression: Martingales

Definition: Let $X_1, X_2, \ldots$ be a sequence of random variables on a probability space $(\Omega, \mathcal{F}, P)$ for which $E[|X_n|] < \infty$ and let $\mathcal{F}_1, \mathcal{F}_2, \ldots$ be a filtration on $\mathcal{F}$ such that X_n is measurable $\mathcal{F}_n$. The sequence $\{(X_n, \mathcal{F}_n) : n = 1, 2, \ldots\}$ is a martingale (supermartingale) (submartingale), if a.s.

$$E[X_{n+1}|\mathcal{F}_n] = (\leq)(\geq)X_n.$$

—▷ A martingale reflects a fair gamble. A gambler with wealth X_n given his information after the n-th play expects his wealth after the next round to be the same as his current wealth.

Supermartingale Convergence Theorem

Let $\{X_t\}$ be a non-negative supermartingale. Then $X_t \to X$ a.s. such that $E|X| < +\infty$.

—▷ Any non-negative supermartingale converges to a random variable whose mean is finite (which implies that the distribution of the random variable does not have any mass on $+\infty$).

14.2 Generalizing the Inverse Euler Equation

14.2.1 Model

- measure one of agents
- preferences
$$\sum_{t=1}^{T} \beta^{t-1} \left[u(c_t) - v(l_t)\right] \tag{14.9}$$
where u strictly concave, v strictly convex and both are bounded
- idiosyncratic shocks: θ^T drawn from μ_Θ
- effective labour: $y_t(\theta^T) = \phi_t(\theta^T) l_t(\theta^T)$

- open economy: $\beta R = 1$

Assumptions:

1. People privately learn θ_t at the beginning of period t.
2. Output y_t and consumption c_t are publicly observed.

Hence, allocations in period t are only θ_t measurable.

Remark: All shocks are drawn at the start of time. Hence, all variables in period t are functions of the shocks drawn, but are measurable only with respect to the history of shocks revealed up to period t.

Remark: Note that the agents can chose a particular (c, y), once they have observed their labour productivity ϕ. After reporting ϕ, the planner instructs them to deliver output y which is associated with utility $u(c(\phi)) - v(y(\phi)/\phi^*)$, where ϕ^* is the true realized idiosyncratic productivity shock.

14.2.2 Pareto Problem

Let ω be the utility level promised to a group of people. A Pareto optimal allocation (c^*, y^*) solves for some ω^*

$$\max_{c,y} \sum_{\theta^T} \sum_t \beta^{t-1} \mu(\theta^T) \left[u(c_t(\omega^*, \theta^T)) - v(y_t(\omega^*, \theta^T)/\phi_t(\theta^T)) \right] \tag{14.10}$$

subject to

$$\sum_{\theta^T} \sum_t \beta^{t-1} \mu(\theta^T) \left[u(c_t(\omega, \theta^T)) - v(y_t(\omega, \theta^T)/\phi_t(\theta^T) \right] \geq \omega \text{ for all } \omega \neq \omega^* \tag{14.11}$$

$$\sum_{\omega} \sum_{\theta^T} \sum_t R^{-t} \mu(\omega) \mu(\theta^T) \left[c_t(\omega, \theta^T) - y_t(\omega, \theta^T) \right] \leq 0 \tag{14.12}$$

$$V(\sigma_{TT}; c, y, \omega) \geq V(\sigma; c, y, \omega) \text{ for all } \sigma, \omega \tag{14.13}$$

The constraints are ex-ante promised utility, intertemporal feasibility and truthtelling, respectively.

<u>Step 1 – Perturbation</u>

Consider any incentive feasible allocation (c^*, y^*). Then, for some time t and some group with utility ω^*, change the allocation to (c', y^*) according to

$$u(c'_t(\omega^*, \theta^T)) = u(c^*_t(\omega^*, \theta^T)) + \Delta + \epsilon(\theta^t) \text{ for all } \theta^T \tag{14.14}$$

$$u(c'_{t+1}(\omega^*, \theta^T)) = u(c^*_{t+1}(\omega^*, \theta^T)) - \beta^{-1}\epsilon(\theta^t) \text{ for all } \theta^T \tag{14.15}$$

$$\sum_{\theta^T}[c'_t(\omega^*, \theta^T) - c^*_t(\omega^*, \theta^T)]\mu(\theta^T) + R^{-1}\sum_{\theta^T}[c'_{t+1}(\omega^*, \theta^T) - c^*_{t+1}(\omega^*, \theta^T)]\mu(\theta^T) = 0 \tag{14.16}$$

This perturbation is incentive feasible, since

- it leaves all other utilities ω untouched
- it scales utilities V by Δ for all reporting strategies σ
- it is resource feasible.

Note that the perturbation happens at specific dates t and $t+1$ and only across all paths with initial history θ^t.

<u>Step 2 – Pareto Problem Rewritten</u>

The optimal allocation solves the problem

$$\max_{\Delta, \epsilon, c'_t, c'_{t+1}} \Delta \tag{14.17}$$

subject to

(14.14) - (14.16)

The solution must be $\Delta = 0$, $\epsilon = 0$, and $c' = c^*$.

<u>Step 3 – FONC at $(0, 0, c^*_t, c^*_{t+1})$</u>

Denote the Lagrange multiplier on the first two constraints $\eta_t(\cdot)$ and $\eta_{t+1}(\cdot)$. The multiplier on the resource constraint is given by λ.

$$\sum_{\theta^T} \eta_t(\theta^T) = 1 \tag{14.18}$$

$$-\sum_{\theta^T \geq \theta^t} \eta_t(\theta^t) + \beta^{-1} \sum_{\theta_{t+1}} \sum_{\theta^T \geq (\theta_{t+1}, \theta^t)} \eta_{t+1}(\theta^t) = 0 \tag{14.19}$$

$$u'(c_t^*(\theta^T)) \sum_{\theta^T \geq \theta^t} \eta_t(\theta^T) = \lambda \sum_{\theta^T \geq \theta^t} \mu(\theta^T) \tag{14.20}$$

$$u'(c_{t+1}^*(\theta^T)) \sum_{\theta^T \geq (\theta_{t+1}, \theta^t)} \eta_{t+1}(\theta^T) = \lambda R^{-1} \sum_{\theta^T \geq (\theta_{t+1}, \theta^t)} \mu(\theta^T) \tag{14.21}$$

Rewriting, we obtain the result

$$\frac{1}{u'(c_t^*(\omega^*, \Theta^T))} = E\left[\frac{1}{u'(c_{t+1}^*(\omega^*, \Theta^T))}|\theta^t\right] \tag{14.22}$$

where we have used the fact that $\beta R = 1$.

—▷ The inverse of the marginal utility follows thus a martingale. Any change on the inverse of marginal utility today has the same expected change on the inverse of marginal utility in the future. Hence, all shocks have permanent effects.

—▷ Why does it work? The key here is that both consumption and the marginal utility of consumption are publicly observable for the planner. This allows us to use the perturbation method as in the two-period moral hazard model to characterize Pareto-optimal allocations.

—▷ Again, we have that there is a wedge in the standard Euler equation,

$$u'(c_t^*(\omega^*, \Theta^T)) < E\left[u'(c_{t+1}^*(\omega^*, \Theta^T))|\theta^t\right] \tag{14.23}$$

which implies that people are savings-constrained.

—▷ What is the intuition? Suppose the agent is not savings-constrained so that the above equality holds with equality. Then, there is a small second-order loss for reducing consumption smoothing, but a first-order gain from offering better insurance when $u'(c_{t+1})$ is not constant across θ^{t+1}.

14.3 Dynamic Mirrlees Taxation

14.3.1 General Idea

Ramsey Taxation:

- planner needs to use linear taxes
- minimize distortions (deadweight loss) from linear taxes
- cannot choose lump-sum taxes

Mirrless Taxation:

- planner can choose any tax system he wants
- but faces frictions (information, enforcement, etc.)
- optimal tax system achieves a *constrained* Pareto optimal allocation
- need to balance insurance vs. incentives
- can choose lump-sum taxes, but does not want to

14.3.2 Model

- measure one of agents
- preferences
$$\sum_{t=1}^{T} \beta^{t-1} \left[u(c_t) - v(l_t)\right] \tag{14.24}$$
where u strictly concave, v strictly convex and both are bounded
- aggregate shock: z^T drawn from μ_Z
- idiosyncratic shocks: θ^T drawn from μ_Θ

- aggregate shock z_t and θ_t learned at the beginning of period t
- effective labour: $y_t(\theta^T, z^T) = \phi_t(\theta^T, z^T) l_t(\theta^T, z^T)$
- effective labour is publicly observed; labor input and skills are private information
- aggregate production function CRS

Assumption: Again all shocks are drawn at the start of time. Hence, all variables in period t are functions of the shocks drawn, but are measurable only with respect to the history of shocks revealed up to period t.

14.3.3 The Inverse Euler Equation Once More

Feasible allocation:

$$\sum_{\theta^T} c_t(\theta^T, z^T)\mu(\theta^T) + K_{t+1}(z^T) + G(z^T) \leq F(K_t, Y_t, z^T) + (1-\delta)K_t(z^T) \tag{14.25}$$

where $Y_t(z^T) = \sum_{\theta^T} y_t(\theta^T, z^T)\mu(\theta^T)$ and $G(z^T)$ is government expenditure.

Incentive Compatibility:

—▷ strategy: $\sigma : \theta^T \times Z^T \to \theta^T \times Z^T$

—▷ pay-off: $V(\sigma; c, y) = \sum_{t=1}^{T} \beta^{t-1} \sum_{z^T} \sum_{\theta^T} \left[u(c_t(\sigma)) - v(l_t(\sigma))\right] \mu(\theta^t)\mu(z^t)$

—▷ truthtelling strategy σ^*

An allocation is incentive compatible, if

$$V(\sigma^*; c, y) \geq V(\sigma; c, y) \tag{14.26}$$

for all σ.

A Pareto-optimal allocation maximizes ex-ante expected utility subject to being resource feasible and incentive compatible.

We again use the fact that there cannot be any way to redistribute consumption between today and tomorrow's states to save costs, while leaving the expected utility of any agent the same at any point in time for any shock (θ^T, z^T) – which implies incentive compatibility.

We solve a perturbed problem given by

$$\min_{c_t, c_{t+1}, K_{t+1}, \xi} \sum_{\theta^T} c_t(\theta^T)\mu(\theta^T) + K_{t+1} \tag{14.27}$$

subject to

$$u(c_t(\theta^T)) = u(c_t^*(\theta^T, z^t)) + \beta \sum_{z_{t+1}} \xi(\theta^T, z_{t+1})\mu(z_{t+1}|z^t) \text{ for all } \theta^T \tag{14.28}$$

$$u(c_{t+1}(\theta^T, z_{t+1})) = u(c_{t+1}^*(\theta^T, z_{t+1})) - \xi(\theta^T, z_{t+1}) \text{ for all } \theta^T, z_{t+1} \succ z^t \tag{14.29}$$

$$\sum_{\theta^T} c_{t+1}(\theta^T)\mu(\theta^T) - F_{t+1}(K_{t+1}, Y_{t+1}(z^t), z^t) - (1-\delta)K_{t+1} = -K_{t+2}(z_{t+1}, z^t) - G_{t+1} \text{ for all } z_{t+1} \succ z^t \tag{14.30}$$

The first-order necessary conditions are given by

$$\mu(\theta^t) - \eta_t(\theta^t)u'(c_t) = 0 \tag{14.31}$$

$$-u'(c_{t+1})\eta_{t+1}(\theta^{t+1}) + \gamma(z_{t+1}|z^t)\mu(\theta^{t+1}) = 0 \text{ for all } z_{t+1} \tag{14.32}$$

$$1 - \sum_{z_{t+1}} \gamma(z_{t+1}|z_t)\left[1 - \delta + MPK(z_{t+1}|z^t)\right] = 0 \tag{14.33}$$

$$\beta\eta_t(\theta^t)\mu(z_{t+1}|z^t) - \sum_{\theta_{t+1}} \eta_{t+1}(\theta^{t+1}) = 0 \text{ for all } z_{t+1} \tag{14.34}$$

where – slightly abusing notation – η's and μ's are understood where appropriate to be the sum of all probabilities and Lagrange multipliers across future paths given a history θ^t. Define $\lambda_{t+1} = \frac{\gamma(z^{t+1}|z^t)}{\mu(z_{t+1}|z^t)}$ which yields the following result.

Proposition 14.3.1. *Suppose (c^*, y^*, K^*) is an optimal allocation. Then, there exists a z_{t+1}-measurable function $\lambda_{t+1}^* : Z^T \to R_+$ such that*

$$\lambda_{t+1}^* = \beta \frac{1}{E\left[\frac{u'(c_t^*)}{u'(c_{t+1}^*)}|\theta^t, z^{t+1}\right]} \tag{14.35}$$

$$E\left[\lambda_{t+1}^*(1 - \delta + MPK(z_{t+1}|z^t))|z^t\right] = 1 \tag{14.36}$$

Again, we get a wedge in the intertemporal Euler equations. To see this, use first Jensen's inequality to obtain

$$\lambda^*(z_{t+1}) < \beta E\left[\frac{u'(c^*_{t+1})}{u'(c^*_t)}|\theta^t, z^{t+1}\right] \tag{14.37}$$

for all z_{t+1} succeeding z^t. Plugging into the second equation and using the law of iterated expectations, we obtain

$$\beta E\left[u'(c^*_{t+1})(1-\delta+MPK_{t+1})|\theta^t, z^t\right] > u'(c^*_t). \tag{14.38}$$

Later on, the idea for implementing the optimal allocation will be to choose taxes on capital so that the intertemporal Euler equation holds with equality for any agent in the decentralized economy.

14.3.4 Interpreting λ^*_{t+1}

We call λ^*_{t+1} the *social discount factor*.

—▷ The Lagrange multiplier λ^*_{t+1} is the shadow value of a unit of more resources tomorrow. It expresses the discounted value of an additional amount of resources next period in event z_{t+1} taking into account the probability of the event. The shadow value of today's resources has been normalized to 1.

—▷ The proposition states that the social discount factor is equal to the harmonic mean of the MRS conditional on θ^t and is independent of individual histories θ^t. That is all agent's harmonic mean of the MRS has to be equal to $\lambda^*_{t+1}(z_{t+1})$ after history z^t.

—▷ The social discount factor then determines how much capital should optimally be accumulated.

—▷ The social discount factor takes into account that an extra unit of consumption needs to be split in such a fashion as to keep the utility level (!) fixed across different histories θ^t. This is very different from raising everyone's consumption by some amount in order to equate marginal utilities.

14.3.5 Decentralization through a Tax System

We restrict ourselves to

- non-linear labour taxes $\psi : \mathbb{R}_+^T \times Z^T \to \mathbb{R}^T$
- linear capital taxes $\tau : \mathbb{R}_+^T \times Z^T \to \mathbb{R}^T$

Hence, the agent pays taxes on new and old capital according to $\tau_t(y(\theta^t, z^t), z^t)(1 - \delta + r_t(z^t))k_t(\theta^t, z^t)$ where I have slightly abused notation with respect to states.

How do capital taxes look like?

Proposition 14.3.2. *Capital taxes equate agent's after-tax MRS with the social discount factor, or*

$$(1 - \tau_{t+1}(y^T, z^T))\frac{\beta u'(c_{t+1}^*(y^T, z^T))}{u'(c_t^*(y^T, z^T))} = \lambda_{t+1}^*(z^T). \tag{14.39}$$

Note that taxes depend on *observable* output and not directly on the announcement of skills. I assume here that there is a 1-1 mapping between the two.

—▷ Taxes are *history-dependent* (θ^t matters) and are *state-contingent* (c_{t+1} matters) as they have to depend on next periods labour income through consumption. Hence, there is uncertainty for the household about tomorrow's capital tax rate.

—▷ Capital taxes are high when future consumption is low and vice versa. This deters a deviation which includes saving more, work too little when skilled and claim to be unskilled tomorrow.

Results:

1) At the optimal tax, the intertemporal Euler equation of the agent is satisfied for the optimal allocation. Why?

$$\beta E[(1-\tau_{t+1})u'(c^*_{t+1})(1-\delta+r_{t+1})|\theta^t, z^t] - u'(c^*_t) = \beta u'(c^*_t)\left[E[\lambda^*_{t+1}(1-\delta+r_{t+1})|\theta^t, z^t] - 1\right] \tag{14.40}$$

2) Conditional on (θ^t, z^{t+1})), tomorrow's *expected* individual capital tax is zero.

$$E[(1-\tau^*_{t+1})|\theta^t, z^{t+1}] = \lambda^*_{t+1}\beta^{-1}u'(c_t)E\left[\frac{1}{u'(c_{t+1})}|\theta^t, z^{t+1}\right] \tag{14.41}$$

3) Aggregate capital taxes are zero for any history z^{t+1}.

$$\sum_{\theta^T} \tau^*_{t+1}k^*_{t+1}(1-\delta+MPK^*_{t+1})\mu(\theta^T) =$$
$$= (1-\delta+MPK^*_{t+1})E[\tau^*_{t+1}k^*_{t+1}|z^{t+1}] = (1-\delta+MPK^*_{t+1})E[E[\tau^*_{t+1}|\theta^t, z^{t+1}]k^*_{t+1}|z^{t+1}] = 0$$

Hence, capital taxes do not raise revenue and are purely redistributive.

4) Labour taxes are lump-sum and thus are chosen to satisfy the budget constraints at the optimal allocation.

5) Current capital taxes are a decreasing function of people's consumption/skills (see above).

Remark: We have assumed throughout that agents cannot engage in side trades that can undermine the planner's incentive scheme.

14.4 Literature

Rogerson, Econometrica (1985)

Kocherlakota, Econometrica (2005)

Kocherlakota, The New Dynamic Public Finance (2010)

Part IV

Introduction to Search Theory

Chapter 15

Search and Unemployment

15.1 The Mortensen-Pissarides Model

15.1.1 Set-up

- measure 1 of workers (w)
- large measure of firms (f)
- linear preferences and discounting at β
- states: $e \in \{0, 1\}$ – unemployed, employed
- if employed, firm produces y and pays wage w
- random matching with probability λ^j
- a match is separated with probability δ

Assumption 15.1.1. *1. Firms post vacancies at cost k and there is free entry.*

2. Workers get a benefit b when unemployed.

15.1.2 Bellman Equations

$$W_1^w(t) = w_t + \beta W_1^w(t+1) + \beta\delta(W_0^w(t+1) - W_1^w(t+1)) \tag{15.1}$$

$$W_0^w(t) = b + \beta W_0^w(t) + \beta\lambda^w(t)(W_1^w(t+1) - W_0^w(t)) \tag{15.2}$$

$$W_1^f(t) = y - w + \beta W_1^f(t+1) + \beta\delta(W_0^f(t+1) - W_1^f(t+1)) \tag{15.3}$$

$$W_0^f(t) = \max\{-k + \beta W_0^f(t+1) + \beta\lambda^f(t)(W_1^f(t+1) - W_0^f(t+1)); \beta W_0^f(t+1)\} \tag{15.4}$$

Using free entry, it must be the case that firms make zero profits from posting a vacancy.

$$W_0^f(t+1) = 0 \tag{15.5}$$

or

$$k = \beta\lambda^f(t)W_1^f(t+1) \tag{15.6}$$

for all t.

The law of motion for unemployed people is given by[1]

$$u(t+1) = u(t)(1 - \lambda^w(t)) + \delta(1 - u(t)) \tag{15.7}$$

We need now to figure out two things.

—▷ First, the wage w is determined by Generalized Nash Bargaining when a match occurs.

—▷ Second, the matching probabilities (λ^w, λ^f) are determined by a matching technology and the behaviour of the agents; i.e., the number of vacancies posted which depends on wages which in turn depends on the (endogenous) outside options in the bargaining problem.

[1] The number of vacancies ensures that – given the matching probabilities – firms make zero profits. More below.

15.2 Steady State Equilibrium

15.2.1 Nash Bargaining

The Bellman equations are now given by

$$W_1^w = w + \beta W_1^w + \beta\delta(W_0^w - W_1^w) \tag{15.8}$$

$$W_0^w = b + \beta W_0^w + \beta\lambda^w(W_1^w - W_0^w) \tag{15.9}$$

$$W_1^f = y - w + \beta(1-\delta)W_1^f \tag{15.10}$$

$$0 = -k + \beta\lambda^f W_1^f \tag{15.11}$$

We solve the following problem

$$\max_w \left(S_f\right)^\eta \left(S_w\right)^{1-\eta} \tag{15.12}$$

where

$$S_f = W_1^f - W_0^f \tag{15.13}$$

$$S_w = W_1^w - W_0^w \tag{15.14}$$

Note that

$$\frac{\partial S_i}{\partial w} = 1 - \beta + \beta\delta(\equiv \rho) \tag{15.15}$$

Solution:

$$\eta S_w = (1-\eta)S_f \tag{15.16}$$

or

$$\eta = \frac{S_f}{S_f + S_w} \tag{15.17}$$

This implies that

$$S_f = \frac{y-w}{\rho} \tag{15.18}$$

$$S_w = \frac{w-b}{\rho + \beta\lambda^w} \tag{15.19}$$

where the denominators are to be enterpreted as the discount factor of a match. In a sense, separation is worse for a firm since it has to reincur the fix cost of posting a vacancy.

After some algebra we arrive at

$$w = \frac{\eta\rho b + (1-\eta)(\rho + \beta\lambda^w)y}{\rho + (1-\eta)\beta\lambda^w} \tag{15.20}$$

$$k = \frac{\eta\beta(y-b)\lambda^f}{\rho + (1-\eta)\beta\lambda^w} \tag{15.21}$$

So everything depends on the matching probabilities.

15.2.2 Equilibrium Market Tightness

Assume a CRS matching function $M(u,v)$ that satisfies

$$M(u,v)/u = M(1,\theta) \text{ and } M(u,v)/v = M(1,\theta)/\theta \tag{15.22}$$

The variable $\theta = \frac{v}{u}$ is called market tightness.

Since $M(u,v)$ describes the number of matches, we have

$$\lambda^w = M(1,\theta) = \theta\lambda^f. \tag{15.23}$$

Definition 15.2.1. *An equilibrium is given by value functions (W_1^f, W_1^w, W_0^w), an equilibrium wage w, a market tightness θ and levels of unemployment and vacancies (u,v) such that*

1. *The free entry condition holds for θ.*
2. *Given market tightness and the value functions, the wage solves the bargaining problem between firms and workers.*
3. *The number of unemployed workers and vacancies satisfy the law of motion for unemployment given market tightness θ.*

Equilibrium condition

$$k = \left(\frac{M(1,\theta)}{\theta}\right)\left(\frac{\eta\beta(y-b)}{\rho+(1-\eta)\beta M(1,\theta)}\right) \tag{15.24}$$

From θ^*, one can now calculate $u^* = \delta/(\delta + M(1,\theta^*))$, v^*, w^*, etc.[2]

To show that – for k sufficiently small – an equilibrium exists and is unique, note first that the RHS is a continuous, strictly decreasing function of θ.[3]

Why? The denominator is given by

$$\frac{\theta}{M(1,\theta)}\rho + (1-\eta)\beta\theta \tag{15.25}$$

which is increasing in θ, since $M_1(u,v) > 0$.

Furthermore, for $\theta \to \infty$, we have that the RHS goes to 0.

Why? $M(1,\theta)$ is increasing in θ.

Hence, there exists a unique θ^* satisfying this equation.

15.2.3 Dynamics

Interestingly, the dynamics do not look really any different from steady-state.

Why?

—▷ One can show that the Nash bargaining problem is independent of t.

—▷ To do show, one can express the value functions in terms of net present value of wages and incomes.

[2]For some basic insights in comparative statics, see homework and the discussion with the TA.

[3]Note that

$$M(1,\theta) - M_2(1,\theta)\theta = M_1(u,v)$$

and that

$$M_2(u,v) = M_2(1,\theta).$$

Prove it!

—▷ This yields a first-order-difference equation in market tightness that is fulfilled for level of market tightness in steady state. This might not be the unique solution though.

Hence, for any sequence $\{u_t\}_{t=0}^{\infty}$, we can compute the time path of the economy.

15.3 Efficiency

Let's think about how the planner is constrained.

- matching technology
- free entry condition
- exogenous separation

We have transferable utility, since people are risk-neutral. Hence, the planner can simply maximize joint surplus by choosing market tightness θ taking as given the law of motion on unemployed people and that – given # of unemployed u – a total number of vacancies $v = \theta u$ needs to be created.

The recursive planner's problem is given by

$$P(u) = \max_{\theta} ub + (1-u)y - kv + \beta P(u') \tag{15.26}$$

subject to

$$u' = (1 - M(1,\theta))u + (1-u)\delta \tag{15.27}$$

$$v = \theta u \tag{15.28}$$

In the homework, you verify that the (unique) solution to this problem is described by a linear function with slope a_1.

The FOC is given by

$$-ku + \beta P'(u')(-1)M_2(1,\theta)u = 0 \tag{15.29}$$

Using the envelope condition, $P'(u) = a_1$ for all u and the value of a_1, we obtain[4]

$$k = \frac{\beta(y-b)M_2(1,\theta)}{\rho + \beta M_1(1,\theta)}. \tag{15.30}$$

Comparing this condition with the equilibrium condition, we obtain the following result.

Proposition 15.3.1 (Hosios Condition)**.** *The equilibrium is efficient if and only if*

$$\eta = \frac{M_2(1,\theta)\theta}{M(1,\theta)} \tag{15.31}$$

Intuition:

—▷ This is the classic case of a search externality.

—▷ Firms do not necessarily take into account the social value of creating a vacancy.

—▷ Since the posting cost k is sunk, the bargaining problem does not take it into account. Hence, conditional on being in a match, the worker bargains as if there are no fixed cost. In other words, there is a hold-up problem for the firm.

—▷ Unless the bargaining power is just right, firms will not invest the socially optimal amount.

—▷ More generally, individuals in search problems do not necessarily take into account their actions on the aggregate market tightness. For example, searching for a job will decrease someone else's probability of finding a match.

15.4 Competitive Search

15.4.1 Set-up

—▷ market makers set up *submarkets*

—▷ each submarket is characterized by a posted wage

[4]See previous footnote.

—▷ firms and workers "direct" their search to a submarket and are randomly match according to $M(u, v)$ in that submarket

—▷ market makers can charge workers a fee for entering the submarket

—▷ firms will enter any submarket as long as they can recover their fixed cost k through expected profits

15.4.2 Main Idea

Workers will choose the submarket with the highest expected (!) payoff.

This payoff depends on (i) the promised wage and (ii) on the probability of finding a job in the submarket – or the market tightness.

The market maker thus proposes (w, θ) to the worker and ensures that he can satisfy this market tightness by attracting enough firms.

Since there is competition among submarkets, the fees charged by market makers are 0.

Market Maker's problem:

$$\max_{(w,\theta)} M(1,\theta)\beta W_1^w + (1 - M(1,\theta))\beta W_0^w \tag{15.32}$$

subject to

$$k = \left(\frac{M(1,\theta)}{\theta}\right)\left(\frac{\beta(y-w)}{\rho}\right) \tag{15.33}$$

$$\tag{15.34}$$

Note that the last condition is the free entry condition for firms with the second term on the RHS being βW_1^f.

The constraint pins down an iso-profit line in the space (w, θ) for all market makers. That is, it gives all combinations of θ and w that are consistent with zero profits on the firm side.

Hence, there is a relationship $\theta = \theta(w)$ that defines all *feasible* submarkets. It is any *iso-profit curve* that arises from the free entry condition and ties the hands of market makers to deliver any particular market tightness θ.

Totally differentiating the constraint yields

$$\frac{d\theta}{dw} = \frac{M(1,\theta)}{(y-w)\left(M_2(1,\theta) - \frac{M(1,\theta)}{\theta}\right)} < 0 \tag{15.35}$$

Hence, if a market offers a higher wage, it is getting tighter, that is there are less vacancies v given a fixed number of workers u. Hence, workers will trade off a higher wage with a lower probability of a match.

In equilibrium, we have that all workers are indifferent between all submarkets and we can assume that there is only a single submarket open.

15.4.3 Efficiency

We can now solve for the competitive search equilibrium.

To do so, we first solve for the value functions to obtain

$$W_1^w - W_0^w = \left(\frac{1}{\rho + \beta M(1,\theta)}\right)(w-b) \tag{15.36}$$

$$W_0^w = \left(\frac{1}{1-\beta}\right) b + \left(\frac{\beta}{1-\beta}\right)\left(\frac{M(1,\theta)}{\rho + \beta M(1,\theta)}\right)(w-b) \tag{15.37}$$

The market maker's problem is thus (up to some normalizing constant)

$$\max_w \left(\frac{M(1,\theta(w))}{\rho + \beta M(1,\theta(w))}\right)(w-b) \tag{15.38}$$

The first-order condition is given by

$$(y-w)\left(\frac{M(1,\theta)}{\theta} - M_2(1,\theta)\right) = M_2(1,\theta)(w-b)\frac{\rho}{\rho + \beta M(1,\theta)}. \tag{15.39}$$

Using the constraint $\theta(w)$ we have once again

$$k = \frac{\beta(y-b)M_2(1,\theta)}{\rho + \beta M_1(1,\theta)}. \tag{15.40}$$

Proposition 15.4.1. *Competitive search equilibrium is efficient.*

The intuition is that the set-up forces agents to consider the impact of wages they earn on the creation of vacancies and the costs associated with it. Market markers take into account that higher wages lower firms profits and, hence, there are less vacancies, markets are more tight and workers finding a job with lower probability.

15.5 Matching Functions

15.5.1 Basics

Consider a matching function $M(u, v)$.

In discrete times, this expresses the # of matches during a time interval given the # of unemployed people u looking for work and the # of vacancies v.

In continuous time, it is the instantaneous rate of job matching given the instantaneous stocks u and v.

Basic properties:

- $M(0, v) = M(u, 0) = 0$
- $M(u, v) \leq \min(u, v)$
- M is increasing in u and v and concave.
- M/u and M/v are probabilities of matching.
- The inverse of these are the mean duration of unemployment and vacancies.

15.5.2 The Role of Elasticities

The elasticity is defined as

$$\frac{\partial M(u,v)}{\partial x}\frac{x}{M} = \eta_x. \tag{15.41}$$

η_u measures the positive externality of workers on firms (thick market).

Why? $M_1(u,v) > 0$.

$\eta_u - 1$ measures the negative externality of workers on other workers (congestion).

Why? $\partial\lambda^w/\partial u \propto \eta_u - 1$.

The same for η_v and $\eta_v - 1$.

15.5.3 Constant vs. Increasing Returns to Scale

The most common matching function is based on an urn-ball experiment. Workers send applications to exactly one vacancy and one application is randomly selected for the vacancy.

The resulting matching function is

$$M = v\Big(1 - (1 - 1/v)^u\Big) \tag{15.42}$$

which can be approximated by

$$M = v\left(1 - e^{-\frac{u}{v}}\right). \tag{15.43}$$

They are homogenous of degree 1 and, hence, exhibit constant returns to scale.

An alternative is the Cobb-Douglas matching function

$$M = \mu u^\alpha v^{1-\alpha} \tag{15.44}$$

where $\alpha \in (0,1)$.

If thick market effects outweigh congestion effects, we get a matching function that exhibits increasing-returns-to-scale. This is the case when

$$\eta_v + \eta_u - 1 > 0. \tag{15.45}$$

There are some microfoundations that give rise to such elasticities which point to larger markets being better for matching.

15.6 Literature

Mortensen & Pissarides (1994), REStud

Burdett, Shi & Wright (2001), JPE

Petrongolo & Pissarides (2001), JEL

Shimer (2005), AER – Quantitative critique of the basic search model based on the Beveridge curve

Chapter 16

Search and Liquidity

16.1 Set up

- time is continous
- measure 1 of people
- $M \in (0, 1)$ have one unit of money
- money is indivisible and no one can hold more than one unit
- people can produce for other people a specific good
- people cannot consume their own good
- $u(q)$ strictly increasing and strictly concave
- $c(q)$ strictly increasing and convex
- they discount time according to e^{-rt}

Each agent faces a probability of a meeting according to a Poisson process with arrival rate αM and $\alpha(1 - M)$ conditional on having money or not.

There is only a trade opportunity if a potential buyer (money) meets a potential seller (no money).

16.2 Value Functions in Continuous Time

16.2.1 Heuristic Derivation

We first derive the value function by approximating the continuous time stochastic process.

For a small time interval Δ, we have that

$$e^{-r\Delta} \approx \frac{1}{1+r\Delta} \tag{16.1}$$

The value of a seller is thus given by

$$V_s(t) = \frac{1}{1+r\Delta}\Big[\alpha M\Delta\left(V_b(t+\Delta) - c(q_{t+\Delta})\right) + (1-\alpha M\Delta)V_s(t+\Delta) + o(\Delta)\Big] \tag{16.2}$$

where $o(\Delta)$ captures all higher order terms. The notation means that $o(\Delta)$ goes faster to 0 that Δ goes.

Rewrite the equation to obtain

$$r\Delta V_s(t) = \alpha M\Delta\left(V_b(t+\Delta) - V_s(t+\Delta) - c(q_{t+\Delta})\right) + V_s(t+\Delta) - V_s(t) + o(\Delta) \tag{16.3}$$

Divide now by Δ and consider $\Delta \to 0$ to obtain

$$rV_s(t) = \alpha M(V_b(t) - V_s(t) - c(q_t)) + \dot{V}_s(t) \tag{16.4}$$

Similarly, for buyers we obtain

$$rV_b(t) = \alpha(1-M)(u(q_t) + V_s(t) - V_b(t)) + \dot{V}_b(t) \tag{16.5}$$

16.2.2 Mathematical Derivation

We now derive the value functions precisely. Note that we take into account the exact form of the Poisson matching process.

The sellers value is given by

$$V_s(t) = E_t\left[\int_t^\tau \bar{u}e^{-r(s-t)}ds + (V_b(\tau) - c(q_\tau))e^{-r(\tau-t)}\right] \tag{16.6}$$

where τ is the stopping time of having a first arrival and where I have added a flow component $\bar{u} \equiv 0$ for generality.[1]

We now use the fact that τ is distributed according to a Possion process to get

$$V_s(t) = \int_t^\infty \left[\int_t^\tau \bar{u}e^{-r(s-t)}ds + (V_b(\tau) - c(q_\tau))e^{-r(\tau-t)}\right]\alpha M e^{-\alpha M(\tau-t)}d\tau \tag{16.7}$$

and differentiating with respect to time t we obtain

$$\dot{V}_s(t) = \frac{d}{dt}\int_t^\infty \left[\int_t^\tau \bar{u}e^{-r(s-t)}ds + (V_b(\tau) - c(q_\tau))e^{-r(\tau-t)}\right]\alpha M e^{-\alpha M(\tau-t)}d\tau \tag{16.8}$$

$$= -\alpha M(V_b(t) - c(q_t)) + \int_t^\infty \frac{d}{dt}\left(\left[\int_t^\tau \bar{u}e^{-r(s-t)}ds + (V_b(\tau) - c(q_\tau))e^{-r(\tau-t)}\right]\alpha M e^{-\alpha M(\tau-t)}\right)d\tau \tag{16.9}$$

$$= -\alpha M(V_b(t) - c(q_t)) + \int_t^\infty \left[\int_t^\tau \bar{u}e^{-r(s-t)}ds + (V_b(\tau) - c(q_\tau))e^{-r(\tau-t)}\right](\alpha M)^2 e^{-\alpha M(\tau-t)}d\tau + \tag{16.10}$$

$$\int_t^\infty \left[\int_t^\tau r\bar{u}e^{-r(s-t)}ds - \bar{u} + r(V_b(\tau) - c(q_\tau))e^{-r(\tau-t)}\right]\alpha M e^{-\alpha M(\tau-t)}d\tau \tag{16.11}$$

$$= -\alpha M(V_b(t) - c(q_t)) + \alpha M V_s(t) + rV_s(t) + \int_t^\infty \bar{u}\alpha M e^{-\alpha M(\tau-t)}d\tau \tag{16.12}$$

$$= rV_s(t) - \alpha M\left(V_b(t) - V_s(t) - c(q_t)\right) + \bar{u} \tag{16.13}$$

where we first apply the Leibniz rule at $t = \tau$ and then apply the Leibniz rule again for the integral with respect to t at t and use the expression for $V_s(t)$.

[1] Note that here we can even allow $\bar{u}(t)$.

16.3 Steady State

We close the model now once again with bargaining.

The quantity q solves in steady state

$$\max_q (u(q) + V_s)^\eta (V_b - c(q))^{1-\eta} \tag{16.14}$$

$$\text{subject to} \tag{16.15}$$

$$V_b - c(q) \geq V_s \tag{16.16}$$

$$V_s + u(q) \geq V_b \tag{16.17}$$

Note that the value functions depend on the optimal q. However, the buyer and seller only take into account the quantity that is produced in their match.

The seller incurs an immediate cost, while the buyer incurs an immediate benefit. Hence, when the seller participates, the buyer must have an incentive to participate as well. The second constraint is therefore never binding.

The FOC is given by

$$\left(\frac{1-\eta}{\eta}\right)\left(\frac{u(q) + V_s}{V_b - c(q)}\right) = \frac{u'(q)}{c'(q)} \tag{16.18}$$

The value functions are given by

$$rV_b = \alpha(1 - M)(u(q) + V_s - V_b) \tag{16.19}$$

$$rV_s = \alpha M(V_b - V_s - c(q)) \tag{16.20}$$

Proposition 16.3.1. *There exists a unique monetary equilibrium. If $\eta = 1/2$, we have that $u'(q) > c'(q)$.*

Proof. Use the value functions to rewrite the FOC as a function of q only.

This function $T(q)$ is zero at $q = 0$, increasing at $q = 0$, continuous in q and eventually becomes negative for large q. This shows existence of an unconstrained monetary equilibrium with $q > 0$ by the intermediate value theorem.

For uniqueness, observe that the RHS of the FOC is strictly increasing and the LHS of the FOC is strictly decreasing in q on some interval $(0, \bar{q})$ where $T(\bar{q}) < 0$.

Finally, at the first best $u'(q^*) = c'(q^*)$, one can show that if $\eta = 1/2$, we have that $T(q^*) < 0$. □

Remark: There is always a non-monetary equilibrium with $q = V_s = V_b = 0$. This equilibrium is not interesting.

Remark: We can define a price equal to $p = 1/q$ because one unit of money purchases q goods.

16.4 Liquidity and Efficiency

We assume for now that $\eta = 1/2$.

Let's define welfare by

$$W = MV_b + (1 - M)V_s \tag{16.21}$$

which can be seen as an ex-ante expected utility of a representative agent.

From the value functions, we get

$$rW = M(1 - M)(u(q) - c(q)). \tag{16.22}$$

A social planner would choose

$$u'(q) = c'(q) \tag{16.23}$$

so that the equilibrium is inefficient.

The inefficiency arises again from the fact that sellers cannot cash in immediately for the production.

One can show that as $r \to 0$ or $\alpha \to \infty$, we have $q \to q^*$. Hence, greater discounting or search frictions increase the inefficiency.

Question: Does changing the money supply improve efficiency?

Differentiating the welfare function with respect to M yields

$$\frac{\partial W}{\partial M} = \frac{1}{r}\left((u(q) - c(q))(1 - 2M) + M(1 - M)(u'(q) - c'(q))\frac{\partial q}{\partial M}\right) \tag{16.24}$$

Hence, there is a trade-off between a *liquidity effect* and a *price effect*.

—▷ Suppose q were fixed. Then, it is optimal that $M = 1/2$. Why? The number of meetings is maximized.

—▷ Suppose M is very small. Then, $\partial q/\partial M > 0$. Why? Upon a meeting, sellers have a chance to generate surplus for themselves in the future. Hence, conditional on being in a meeting, they are "desperate" to acquire money.

—▷ For $M = 1/2$, however, we have that $\partial q/\partial M < 0$. Hence, the price level increases when money increases. This implies that $M^* < 1/2$.

Remark: We could also look at how bargaining power affects the quantity produced. It turns out to be the case that $\partial q/\partial \theta > 0$ with a range from 0 to $q_1 > q^*$.

16.5 Dynamics

We look at some primitive dynamics. To do so, set $\eta = 1$; i.e., the buyer makers a take-it-or-leave-it-offer. Furthermore, we can set $c(q) = q$ without loss of generality.

From the bargaining problem, we thus have immediately that

$$q(t) = V_b(t) - V_s(t) = V_b(t) \tag{16.25}$$

since the buyer extracts all surplus.

Thus, $\dot{q} = \dot{V}_b$ and the dynamics are described by an ordinary first-order difference equation

$$\dot{q} = rq - \alpha(1 - M)u(q) \tag{16.26}$$

which yields two steady states[2], $q = 0$ and $\tilde{q}$ defined by

$$\frac{u(\tilde{q})}{\tilde{q}} = \frac{r}{\alpha(1 - M)}. \tag{16.27}$$

Note that the RHS of the differential equation is a continuous, convex function that decreases for $q \to 0$.

Proposition 16.5.1. *The non-monetary steady state is stable, while the monetary steady state is unstable. There is a continuum of paths that converge to the non-monetary steady state.*

Not that all these paths are purely driven by self-fulfilling expectations (or sunspots!). In essence, we can pick any $q_0 \in (0, \tilde{q})$ for a non-stationary equilibrium that converges to $q = 0$.

16.6 Literature

Trejos and Wright (1995) – JPE

Duffie, Garleanu, Pedersen (2005) – Econometrica

[2]Since u is strictly concave, we have that $u'(q) > u(q)/q$.

Part V

Information in Macroeconomic Models

Chapter 17

News Shocks

17.1 Introduction

What are "news shocks"?

Direct Shocks to Expectations:

A random variable follows a process given by

$$x_t = \rho x_{t-1} + \xi_t + \epsilon_{t-s} \tag{17.1}$$

Here ξ_t are surprises and ϵ is information received s periods ago about the random variable in period t.

This implies that we have to offset an anticipated shock that does not happen by setting $\xi_t = -\epsilon_{t-s}$.

Noise signals:

People receive imperfect signals about the true state of the world

$$s_{it} = x_t + \epsilon_{it} \tag{17.2}$$

Note that the signals could be public or private ones. If there are private ones, issues of higher order beliefs and common knowledge arises.

Expectations are automatically revised upon receiving new signals or information (e.g. through prices).

17.2 Co-movement puzzle

There is an old idea that business cycles are driven by expectations about the future and disappointments related to these expectations.

One would expect that consumption, investment and labour supply would move in the same direction if people receive "news" that changes their expectations about the future.

Consider the following RBC model:

$$\max_{C,N,K} E_0 \left[\sum_{t=0}^{\infty} \beta^t \left(\log C_t + \psi \frac{(1-N_t)^{1-\eta}}{1-\eta} \right) \right] \tag{17.3}$$

$$\text{subject to} \tag{17.4}$$

$$C_t + K_{t+1} \leq e^{a_t} K_t^{\alpha} N_t^{1-\alpha} + (1-\delta) K_t \tag{17.5}$$

$$x_t = \rho x_{t-1} + \xi_t + \epsilon_{t-4} \tag{17.6}$$

The impulse functions from `DYNARE` are shown below.

The issue is that with a positive news shock about future productivity the labour supply and investment falls.

Why?

- The labour demand curve does not shift at all since productivity has not increased yet.

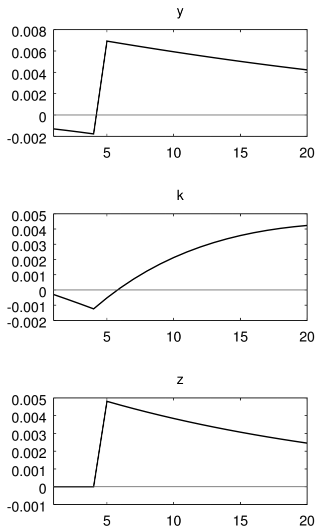

Figure 17.1: RBC Model – IRF w.r.t to news shock

- There is a positive income effect as future expected income has increased. The consumer smoothes consumption by both (!) consuming some of the capital stock and increasing its leisure.
- Hence, the labour supply curve shifts and we have lower equilibirum employment.
- Once the productivity has increased, there is an investment boom and the labour demand shifts. Hence, there is higher equilibrium employment starting in period 5.

When we increase the elasticity of substitution, we can shift the labour supply curve due to a strong substitution effect to generate positive investment. However, consumption falls at the same time.

Beaudry and Portier (2007) JET have shown that in *any* one-sector growth model consumption and investment must move in opposite direction in response to news shocks.

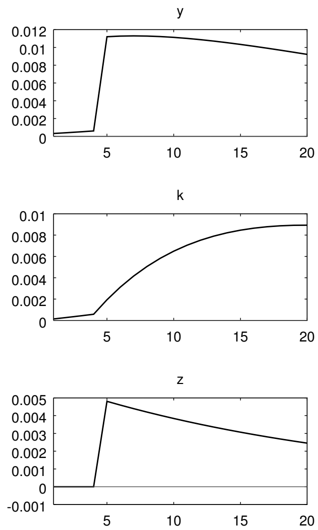

Figure 17.2: RBC Model with $\gamma = 1/4$ – IRF w.r.t to news shock

17.3 Moving Labour Supply

17.3.1 Multi-Sector Growth Models

Consider the following production function

$$C_t = (X_t^\nu + K_t^\nu)^{\frac{1}{\nu}} \tag{17.7}$$

that makes intermediate inputs and capital imperfect substitutes.

Suppose further that intermediate goods and capital are produced according to

$$X_t = e^{a_t} N_{X,t}^\alpha \tag{17.8}$$

$$K_{t+1} - (1-\delta)K_t = N_{i,t}^\alpha \tag{17.9}$$

$$N_t = N_{X,t} + N_{i,t} \tag{17.10}$$

Suppose there is a news shock about an increase in a_{t+1}, the productivity in the intermediate goods sector. Then, people will increase capital investment in period t already so that they can take advantage of increased productivity tomorrow. This requires more labour. Hence, we have a positive shift in labour supply.

Why? X_{t+1} and K_{t+1} are complementary in production.

Remark: We need linear disutility of labour. Otherwise, we see a relocation from the intermediate goods sector to the investment goods sector. This can be solved by separating households into workers in different sectors.

17.3.2 Consumption Habit

Consider now a difference preference structure given by

$$E_0 \left[\sum_{t=0}^{\infty} \beta^t \left(\log(C_t - hC_{t-1}) + \psi \frac{(1-N_t)^{1-\eta}}{1-\eta} \right) \right] \tag{17.11}$$

These preference imply that households do not want consumption fall too much. Note that households internalize their effect of today's consumption on tomorrow's consumption.

What is the marginal utility of consumption today?

$$\frac{\partial U}{\partial C_t} = \frac{\beta^t}{C_t - hC_{t-1}} - E_t \left[\frac{\beta^{t+1} h}{C_{t+1} - hC_t} \right] \tag{17.12}$$

Suppose the consumer expects higher consumption C_{t+1}. Then, his marginal utility of consuming more today increases. Hence, he will work more and we have – once again – a positive shift in the labour supply.

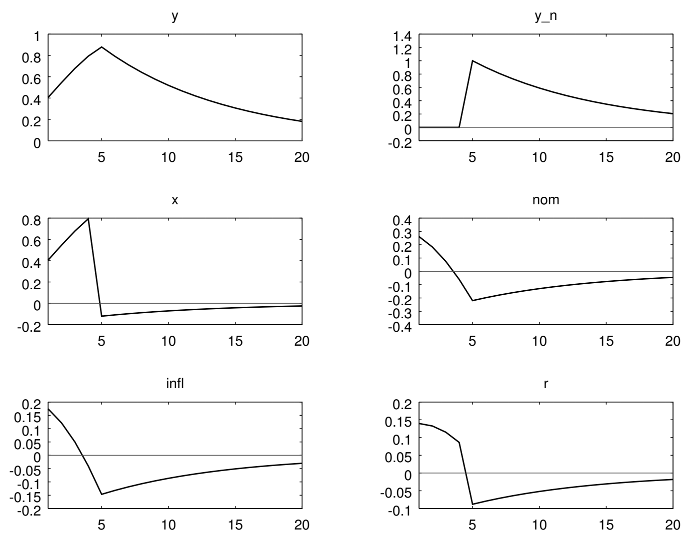

Figure 17.3: NK model – IRF w.r.t to news shock in 4 periods

17.4 Moving Labour Demand

17.4.1 Basic NK Model

Suppose we have the basic NK model with sticky prices and log utility

$$x_t = E_t[x_{t+1}] - (i_t - E_t[\pi_{t+1}] - r_t^n)] \tag{17.13}$$

$$\pi_t = \kappa x_t + \beta E_t[\pi_{t+1}] \tag{17.14}$$

$$i_t = \phi \pi_t \tag{17.15}$$

where we have normalized the rate of time preference to $\rho = 0$.

Productivity follows an AR(1) process

$$a_t = \rho_a a_{t-1} + \xi_t + \epsilon_{t-s} \tag{17.16}$$

so that the news shock is about tomorrow's level of productivity.

The impulse response functions are shown below.

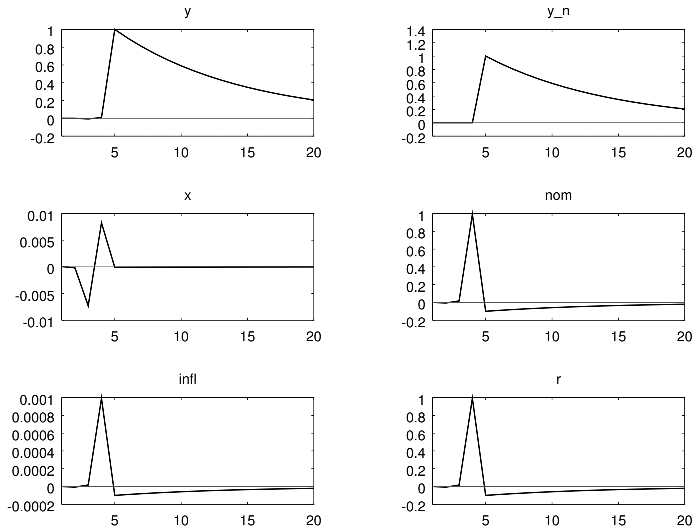

Figure 17.4: NK model with "inflation nutter" – IRF w.r.t to news shock

The key idea here is again that output is demand determined. Hence, we have that the income effect increases consumption and, hence, output. In other words, the news shock positively shifts the labour demand function.

Note that there is a positive output gap. Indeed, this implies that there is an inefficiency. The economy is producing too much before productivity increases in period 4.

Suppose we increase the reaction coefficient of monetary policy from $\phi = 1.5$ to $\phi = 1000$. This means that the central bank does everything to counteract sticky prices. As the impulse response functions show, we have that the news shock has zero impact on the economy.

<u>Remark:</u> The problem here is that the positive reaction to news about future productivity arises from the fact that monetary policy reacts insufficiently to inflation. Setting $\phi \to \infty$ would fully stabilize the inflation and the output gap. This can be fixed, however, by assuming that the central bank cannot observe the shocks directly or that there are shocks (e.g. mark-up shocks to the Phillips curve) that imply a trade-off for the central bank when reacting to inflation.

One can see this also analytically. Set $\rho_a = 1$. Since productivity follows a random walk, there is no forward-looking behaviour in terms of inflation and $E_t[\pi_{t+1}] = 1$.

Hence, we have that

$$x_t = E_t[x_{t+1}] - (\phi\kappa x_t - r_t^n) \tag{17.17}$$

$$y_t = a_t + E_t[y_{t+1} - a_{t+1}] - \phi\kappa(y_t - a_t) + E_t[a_{t+1}] - a_t \tag{17.18}$$

$$y_t = \frac{1}{1+\phi\kappa}E_t[y_{t+1}] + \frac{\phi\kappa}{1+\phi\kappa}a_t \tag{17.19}$$

$$y_t = a_t + \frac{1}{\phi\kappa}\epsilon_t \tag{17.20}$$

since $E_t[y_{t+1}] = a_t + \epsilon_t$. Therefore, as $\phi \to \infty$, we have that the economy shows no reaction to the news shock which is efficient here.

Remark: We have that inflation is given by

$$\pi_t = \kappa(y_t - a_t) = \frac{\kappa}{1+\phi\kappa}\epsilon_t \tag{17.21}$$

so that inflation moves 1-1 with the news shock. Monetary policy thus dampens the inflationary response to a news shock, as it dampens aggregate demand.

17.4.2 Dispersed Information and Noise

See below.

17.5 Literature

Lorenzoni (2011) – Annual Review of Economics

Beaudry & Portier (2004), (2006), (2007)

Christiano et al. (2010)

Chapter 18

Belief Driven Business Cycles

18.1 Noisy Signal

See notes by Frederic Tremblay on Lorenzoni (2009) AER.

18.2 Sentiments and Higher Order Beliefs

See Angeletos & La'O (2014) Econometrica

Chapter 19

Information Choice

19.1 Inattentiveness

See Reis and Mankiw & Reis

19.2 Rational Inattention

Sims & Information Theory

19.3 Literature

See Veldkamp (2011) – Information Choice in Macroeconomics and Finance

Chapter 20

Near Rational Expectations

Woodford